



**A University of Sussex PhD thesis**

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

University of Sussex

# **THE EFFECTIVENESS OF POLICY EVALUATION**

## **Insights from the health care sector in Mexico and Chile**

Blanca Odille López Rodríguez

A thesis submitted to the Department of Politics  
of the School of Law, Politics and Sociology  
of the University of Sussex  
for the degree of Doctor of Philosophy in Politics  
Brighton, December 2017

### **Declaration**

I certify that the thesis I have presented for examination for the degree of Doctor of Philosophy in Politics of the University of Sussex is solely my own work.

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent.

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 87,676 words.

## **Abstract**

Over time, the predominant tendency of many governments' agencies has been to evaluate a programme or policy investing large amount of resources in supporting policy evaluation. However, recommendations suggested by policy evaluators are not always taken up. Moreover, there is relatively little evidence of the extent of policy evaluation effectiveness (i.e. the influence of evaluation on the programme evaluated) and the factors which have significant impact on it. This dissertation aims to shed light on this issue by focusing on the Mexican and Chilean experiences of policy evaluation in the health care sector.

It provides a detailed analysis of the extent to which evaluations have led to changes in policies and programmes and reveals a rather limited effectiveness of policy evaluation in these countries. I argue that shortcomings in the effectiveness of policy evaluation can be explained by institutional and political factors, primarily the nature of Intra Governmental Relations (IGR), but also the quality of bureaucracy, the level of democracy, the autonomy of policy evaluators, and the type of policy evaluation framework. While all of these factors seem to have some influence, the relationship between the executive and legislature is clearly the key determinant of the take up of recommendations.

Thus, the findings of this thesis suggest that strengthening coordination between the different parts of government is needed to enhance the effectiveness of policy evaluation. In addition, the analysis also suggests that policy evaluation is likely to be more effective when it incorporates budgetary incentives.

### **Acknowledgements**

This research it would not have been possible without the kind support and help of many individuals. I would like to extend my sincere thanks to all of them.

I would like to express my gratitude towards the Mexican government's Consejo Nacional de Ciencia y Tecnología (CONACYT) for the financial support and assistance to pursue this doctoral project.

I take this as an opportunity to express my greatest regards to my family for their patience, support and understanding to pursue this dissertation.

I also express my gratitude to my friends for their valuable suggestions and constant encouragement throughout my research period.

My thanks and appreciations also go to my colleagues and people who have willingly helped me out with their abilities and comprehensive advice in the conduct of this research.

To all the respondents who were very cooperative in giving information which facilitate completion of this project.

This thesis has indeed helped me to explore more knowledgeable avenues related to my topic and will help me in my future.

## Table of contents

|   |    |
|---|----|
| Abstract  | 3  |
| Acknowledgements  | 4  |
| <b>Introduction</b>   |    |
| Why is effectiveness important?   | 9  |
| Contribution of the thesis  | 9  |
| The central discussion of the thesis  | 10 |
| Chapter outline   | 12 |
| <b>Chapter 1. Broadening understanding of policy evaluation and effectiveness</b> |    |
| 1.1 Introduction  | 14 |
| 1.2 The roots of evaluation   | 14 |
| 1.3 Evaluation in emerging countries and the New Public Management                | 17 |
| 1.4 A re-conceptualisation of the theory of evaluation use                        | 20 |
| 1.5 What explains the effectiveness of policy evaluation?                         | 23 |
| 1.6 Factors influencing the effectiveness of evaluation                           | 25 |
| 1.7 Concluding remarks  | 33 |
| <b>Chapter 2. Methodological approach</b>   |    |
| 2.1 Introduction  | 34 |
| 2.2 Research strategy or how the research is to be carried out?                   | 34 |
| 2.2.1 Dependent variable and operationalisation of evaluation effectiveness       | 35 |
| 2.2.2 Independent variables   | 39 |
| Intra Governmental and Interagency Relations 'IGR'                                | 40 |
| Quality of the bureaucracy  | 43 |
| Level of democracy  | 44 |
| Degree of autonomy  | 45 |
| Policy evaluation framework   | 46 |
| 2.3 The comparative dimension of policy evaluation                                | 47 |
| 2.4 Documentary and analysis sources  | 51 |
| 2.4.1 Elite interviews  | 52 |
| <b>Chapter 3. Policy evaluation in Mexico</b>                                     |    |
| 3.1 Introduction  | 55 |
| 3.2 Historical perspective  | 55 |
| 3.3 The current state of affairs of policy evaluation                             | 71 |
| <b>Chapter 4. Policy evaluation in Chile</b>                                      |    |
| 4.1 Introduction  | 73 |
| 4.2 Historical perspective  | 73 |
| 4.3 The current state of affairs of policy evaluation                             | 80 |
| <b>Chapter 5. The effectiveness of policy evaluation in Mexico</b>                |    |
| 5.1 Introduction  | 82 |
| 5.2 Results-based Budgeting 'PbR/SED'   | 82 |
| 5.2.1 How is the federal budget made?   | 82 |
| 5.3 Results of the programmes assessed  | 83 |
| 5.4 The effectiveness of evaluation in the health sector                          | 84 |

|  |     |
|--|-----|
| 5.5 The effectiveness of evaluation in the Results-based Budgeting (PbR/SED) | 87  |
| 5.5.1 What explains the effectiveness in policy evaluation?                  | 91  |
| 5.6 Intra-Governmental Relations (IGR)                                       | 91  |
| 5.7 Quality of bureaucracy   | 94  |
| 5.8 Level of democracy   | 96  |
| 5.9 Policy evaluation framework  | 99  |
| 5.10 Degree of autonomy  | 104 |

## **Chapter 6. The effectiveness of policy evaluation in Chile**

|  |     |
|--|-----|
| 6.1 Introduction   | 108 |
| 6.2 The budget process “Results-based Budgeting”                   | 109 |
| 6.3 The effectiveness of evaluation in the health sector           | 110 |
| 6.4 The effectiveness of evaluation in the Results-based Budgeting | 114 |
| 6.4.1 What explains the effectiveness in policy evaluation?        | 116 |
| 6.5 Intra-Governmental Relations (IGR)                             | 117 |
| 6.6 Quality of bureaucracy   | 120 |
| 6.7 Level of democracy   | 124 |
| 6.8 Policy evaluation framework                                    | 125 |
| 6.9 Degree of autonomy   | 128 |

## **Chapter 7. A comparative analysis of the institutional factors influencing the effectiveness of evaluation**

|   |     |
|---|-----|
| 7.1 Introduction  | 131 |
| 7.2 Categories of recommendations   | 131 |
| 7.3 Evaluation in the Mexican health sector                                   | 132 |
| 7.3.1 <i>Status quo</i> programmes  | 132 |
| 7.3.2 Amended programmes  | 134 |
| 7.3.3 Cancelled programmes  | 135 |
| 7.4 Evaluation in the Chilean health sector                                   | 138 |
| 7.4.1 <i>Status quo</i> programmes  | 138 |
| 7.4.2 Amended programmes  | 139 |
| 7.4.3 Cancelled programmes  | 140 |
| 7.5 The influence of institutional factors on the effectiveness of evaluation | 142 |

## **Chapter 8. Conclusions**

|   |     |
|---|-----|
| 8.1 Theoretical contributions and policy implications of the thesis | 164 |
| 8.2 Limitations and recommendations                                 | 166 |
| 8.3 Final statements  | 168 |

|                        |     |
|------------------------|-----|
| <b>List of acronym</b> | 170 |
|------------------------|-----|

|                          |     |
|--------------------------|-----|
| <b>List of interview</b> | 171 |
|--------------------------|-----|

|                   |     |
|-------------------|-----|
| <b>References</b> | 172 |
|-------------------|-----|

## **List of annex, charts, figures and tables**

|   |     |
|---|-----|
| Annex 1. Measurement of independent variables                   | 194 |
| Annex 2. Questionnaire “The effectiveness of policy evaluation” | 195 |
| Annex 3. Office letter no. 1761 DIPRES in Chile                 | 197 |
| Annex 4. Evaluation systems and monitoring practices in Mexico  | 203 |
| Annex 5. Healthcare programmes of Maule                         | 204 |
| Annex 6. Healthcare programmes of Osorno                        | 206 |
| Annex 7. Evaluation systems and monitoring practices in Chile   | 208 |
| Annex 8. Categories of recommendations across countries         | 209 |

|   |     |
|---|-----|
| Chart 1. Effectiveness of evaluation in the Mexican health sector 2007-2014               | 85  |
| Chart 2. Resources provided to coordinators of evaluation in Mexico                       | 93  |
| Chart 3. Effectiveness of evaluation in the Chilean health sector 2007-2014               | 111 |
| Chart 4. Resources provided to evaluations across the Chilean public administration       | 119 |
| Chart 5. Remunerations (real) in the Chilean economy and public sector 1978-2000          | 121 |
| Figure 1. The effectiveness of policy evaluation  | 21  |
| Figure 2. Elements of the policy evaluation framework                                     | 46  |
| Figure 3. Planning and evaluation functions in the SPP and SECOGEF                        | 59  |
| Figure 4. Modernization of the Mexican Federal Public Administration                      | 61  |
| Figure 5. Results-based Budgeting/Performance Evaluation System                           | 69  |
| Figure 6. Annual Evaluation Programme 'PAE' 2007 to 2013                                  | 70  |
| Figure 7. System of evaluation and management control 'SECG'                              | 78  |
| Figure 8. Budget cycle in Mexico  | 83  |
| Figure 9. Budget cycle in Chile   | 110 |
| Figure 10. Dimensions of political relationships  | 144 |
| Table 1. Planning, control, monitoring and evaluation practices in Mexico                 | 217 |
| Table 2. Sources of elite interviews  | 53  |
| Table 3. The effectiveness of policy evaluation in Mexico                                 | 223 |
| Table 4. Allocation of financial resources based on evaluations in Budget (PEF) FY 2011   | 88  |
| Table 5. Government effectiveness (quality of bureaucracy) in Mexico                      | 95  |
| Table 6. The effectiveness of policy evaluation in Chile                                  | 232 |
| Table 7. Financial resources of the Chilean health sector                                 | 115 |
| Table 8. Comparison of resources to perform evaluations                                   | 120 |
| Table 9. Evolution of incentives in Chile 1998-2010 through the PMG                       | 122 |
| Table 10. Government effectiveness (quality of bureaucracy) in Chile                      | 122 |
| Table 11. Categories of recommendations of sample of programmes assessed across countries | 137 |
| Table 12. Recommendations achieved of sample of programmes assessed across countries      | 141 |



## Introduction

In recent decades, evaluation has become a byword within public management discourse for most governments. There has been considerable interest in using evaluation for policy improvement, enlightenment and learning. Given this interest it might be assumed that the findings of evaluations would be utilised for further improvements on policies and programmes. The use or utilisation of evaluations highlights the issue of the effectiveness of evaluation as the extent to which the evaluation of a policy or programme leads to changes in that policy. In practice, however, the evidence indicates that this rarely takes place.

For the most part, the existence of recommendations resulting from evaluations does not ensure that they will be taken into consideration in the policy process. The gap between evaluation and policy change is the main focus for this research. Indeed, the central research question for this dissertation is *whether evaluation is effective across the Mexican and Chilean policies*, in other words whether or not evaluation is utilised in policy making process. Moreover, this research builds on the examination of evaluations to analyse the institutional factors that might affect the effectiveness of evaluation.

The effectiveness of policy evaluation has to be understood in the context of attempts to improve the performance of the State as an entity that aims to develop policies to address public problems. When such policies do not achieve the expected results the credibility of the State is eroded. This was particularly the case in the 1980s, when many States (e.g. Canada, the UK, the USA, Mexico, Brazil, Argentina, etc.) faced a crisis due to inefficiencies and a lack of capability to tackle these public problems. As a result, over the following years many countries engaged in a process of administrative modernisation according to the paradigm of the New Public Management 'NPM' (see Osborne & Gaebler, 1992). This has been the basis of the State modernization agenda, including administrative reforms and focus on performance evaluation.

Initially, these reforms were introduced in developed countries such as Australia, New Zealand and the United Kingdom (Rhodes, 1991) amongst others. Later on, the emerging economies followed a similar process of public sector reforms (see Aguilar, 2013, Cabrero, 2003, and Uvalle, 1994). This research focuses on the experience of reforms in such economies, conducting a comparative study of the effectiveness of evaluation in the Chilean and Mexican policy processes. Although some scholars have argued that these strategies seemed more as a "paradigm" than administrative reforms<sup>1</sup> (Arellano, 2002, Cejudo, 2008, Hood & Peters, 2004, Jones, 2002, Jones & Kettl, 2003 and Manning, 2001), the NPM brought the Results-based Budgeting<sup>2</sup> to the epicentre of policy evaluation. In this context, evaluation is seen as a tool, which aims to improve government performance across all sectors (energy, security, education, health, etc.) and to administer efficiently the resources allocated by the government.

---

<sup>1</sup> More focused on government outputs than outcomes (Jones, 2002), especially in developing countries, where hierarchical bureaucracies had not been downsized or replaced by market mechanisms (Cejudo, 2008 and Manning, 2001).

<sup>2</sup> Essentially, the Results based Budgeting systems (RbB or PbR acronym in Spanish) and the Performance Evaluation System (SED) developed under this model, are based on the principle that the public sector needs an administrative culture that emphasizes the management and measurement of inputs, activities, and outcomes. Although, references to RbB might raise discussion on budgeting this is not provided in this dissertation.

In order to do this, evaluation enables agencies to assure whether their programmes are coherent and congruent with their goals. As Howlett et al. (2009: 186) suggest, the goal of evaluation is to find out whether the programme – policy, plan or project – is doing what it is supposed to do or whether it needs to be adjusted. Looking at the policy process as a whole, evaluation is often regarded as a final stage in assessing the State's actions. It 'should be' closely linked to the political process that determines the selection, funding, modification, extension and termination of projects (Bamberger, 1991: 326).

### **Why is effectiveness important?**

Over time, the predominant tendency of many governments' agencies has been to evaluate a programme, project, plan or policy, on the basis of investing large amount of resources in supporting evaluation and developing evaluators' expertise and data collection, and then making the results publicly available. However, the making of recommendations as a result of evaluations does not guarantee that they will be adopted by the policy process. The importance of this topic (effectiveness) is highlighted by the fact that, despite several evaluations and official data performed, there is relatively little evidence of their effectiveness.

In other words, there is limited information indicating whether and to what extent policy-makers make decisions based on results and recommendations from the evaluation process. Besides, much of the existing body of research on the effectiveness of evaluation has been conducted by evaluators based on their experience. Such research is often prone to bias and usually offers a limited description of the method used (Herbert, 2014: 412) to determine the effectiveness of evaluation. Thus, this research is an empirical study of the relationship between evaluations of policies and programmes performed on the one hand and the extent to which recommendations of evaluations are taken into account by decision-makers in terms of policy change on the other.

### **Contribution of the thesis**

In order to develop a comprehensive narrative, this research aims to address the core questions of effectiveness influencing the evaluation process from an objective perspective and the factors which have significant impact on it. This dissertation also aims to contribute to the debate on effectiveness by looking at the problems which policy evaluation has encountered in Chile and Mexico.<sup>3</sup> It also will attempt to explain the broad spectrum of the effectiveness to contribute to the theory of evaluation use. In general terms, the evaluation is considered fragmented, not timely, and disconnected from the policy process and the performance of the government agencies. Moreover, there is little evidence about how results are used in the budgetary process and there is rarely a systematic connection between evaluation and policy improvement. This may be a reminder that evaluation policy is also a political process (Dye, 2011: 336, Hill, 2009: 281-82 and Howlett & Ramesh, 1995: 169).

The dissertation also seeks to explore *whether or not evaluation is effective in supporting decision-making* to improve programmes, policies and services, identifying factors (negative and positive), providing a complement to other programmes, and

---

<sup>3</sup> These problems are mainly a consequence of the ambiguity in the roles performed, lack of autonomy of evaluators and absence of leadership and coordination amongst agencies, ministries and branches. Besides, the lack of experience in evaluation and the quality of civil service also affect the capacity of other actors (e.g. academics, non-government organisations, consultancy firms, audit agencies and citizenry) to analyse information of the government's performance. Finally, there are few incentives for officials (bureaucrats), ministers or legislators to encourage learning and to improve decision-making on the basis of evaluations.

in the case of the budgetary process, offering a rationale for increasing or reducing finances. While effective evaluation is needed to strengthen the public sector, especially for countries characterised by limited resources such as Mexico and Chile, it is necessary to ask whether in such emerging economies, public resources have been wasted on an ineffective evaluation process. This research aims to provide insights into the way evaluations are done, to explain *whether policy evaluation matters for the public policy and budgetary process*.

This research goes far beyond to determine whether the policy evaluation is effective or not exploring the potential influence of institutional factors and whether these are the most pressing issues for closing the gap in the effectiveness. Lastly, the contribution of this research to the literature is empirical rather than conceptual to prove whether the recommendations are used or not and whether the achievement of these recommendations is linked either to the complexity of the task or to the costs of implementing changes. Moreover, to enhance our theoretical understanding of the evaluation process. Indeed, the effectiveness of policy evaluation is central for understanding its value.<sup>4</sup>

### **The central discussion of the thesis**

From the perspective of political science, this project will argue that shortcomings in the effectiveness of policy evaluation can be explained by institutional factors. The following chapters attempt to answer *What are the most influential institutional (political) factors shaping the effectiveness of evaluation?* In a way, the debate seems as a puzzle due to a variety of approaches and perspectives regarding the question of effectiveness. Recent studies focus on the human, contextual and evaluation factors (Alkin, 1985 cited in Fleischer & Christie, 2009) while others argue that the evaluation capacity building, the role of evaluator, stakeholders' involvement, and intentionality facilitate the use of evaluation findings (Fleischer & Christie, 2009). A further group of studies tends to assume that the organisational and political context, and the intentions of those requiring evaluation (Teirlinck et al., 2013: 374) are the factors which influence the adoption of recommendations.

This dissertation will, on the basis of qualitative case study research, explore the potential influence of a number of institutional factors: the Intra Governmental Relations (IGR), quality of bureaucracy, level of democracy and autonomy, and the policy evaluation framework, which are relatively under-explored by scholars. Indeed, little has been said in the evaluation literature about how these factors might influence the evaluation process. The thesis also explores the political 'arenas' in which evaluation becomes a powerful source of information for those actors who have the responsibility for achieving results, in order to conclude the policy process and decide whether to modify, eliminate, reallocate resources to or "learn" from policies.

In exploring the influence of these factors, the dissertation seeks to test the following statement: *The political factor in the IGR is decisive for the effectiveness of policy*

---

<sup>4</sup> One of the most relevant examples of how programme improvements can be attributed to evaluation results is the social programme PROSPERA. Numerous scholars (see Essama-Nssah, 2013, Rubio, 2012) and international organisations (WB, OECD, and IADB) refer to this programme as an international practice of conditional cash transfer programme in developing countries. The programme PROGRESA was launched by the Mexican government in 1997 and renamed as Oportunidades in 2001 and Prospera in 2014, with the aim to develop human capital and reduce both present and future poverty. To determine whether and to what extent this programme achieved its intended outcomes, have had performed several evaluations. Its innovative model of impact evaluation was pioneering in the social sector, having a relevant role due to the positive findings to enlighten the policy process, increase financial resources and coverage to more households, and eventually influenced the spread of evaluation across the Mexican public administration.

*evaluation.* As explained in the methodology chapter, the IGR are those patterns of interaction among actors, agencies and branches of government. A first assumption is that the IGR are the condition either present or absent during the decision-making process and the way the IGR are arranged reflects the political factors which influence the decision made by actors involved.

Though the comparative case study, the analysis of programmes evaluated will demonstrate they (actors) have the power to achieve evaluation. Comparatively to other factors, this explanation seems stronger for the purpose of utilisation than other potential explanations mentioned above. To answer the above questions, a central task of the research is to find alternative explanations and cases, which can be compared in terms of the factors identified above.

Studying results across countries allows some conclusions to be drawn on how institutional factors influence the effectiveness of policy evaluation. The emphasis in most studies of the utilisation of evaluation findings is on developed rather than emerging economies (Askim, 2007, Dahler-Larsen, 2000, Fleischer & Christie, 2009, Ledermann, 2011 and Teirlinck et al., 2003). The focus of this research is a comparative analysis of the effectiveness of policy evaluation in Chile and Mexico, taking as a case study the use of evaluation in each country's health sector as the focus for an examination of the effectiveness of evaluation. It examines the extent to which evaluations have led to changes in the policies or programmes carried out by the government. The decision for the comparative analysis of these countries is because these along with many Latin American countries have a tradition of policy-transfer<sup>5</sup> of administrative reforms, policies and programmes.

In addition there are a number of similarities between the cases including economic, language, cultural and most importantly administrative features. In this context, Dussauge (2013) claims that political systems based on presidential principles, transitions from authoritarianism to democracy are relevant features possess in common. Moreover, in recent administrations both the Management Improvement Programme and results-based budgeting travelled from Chile to Mexico associated with cross-national policy learning (Dussauge, 2012b; 2013) as a familiar approach considering political, economic and geographic circumstances. Whereas the most crucial difference found were in banking and bankruptcy procedures (Bergoeing et al., 2002).

Indeed, the Chilean policy evaluation is key for this analysis due to its characteristics such as an emerging economy pioneering a M&E system and being recognized worldwide as one of the strongest system. From which, the Mexican government adopted the evaluation model further explained in the methodology chapter.

Bergoeing et al. (2002) also note that the relevance is the recovery paths of the economic crisis during the 1980s, which differed markedly. On the one hand, the standard monetarist and real-wage story for Chile, help this country to recover more rapidly, while on the other hand, the debt overhang and structural reforms story of Mexico (e.g. trade policy, fiscal policy, privatization, the banking system and bankruptcy laws) were the principal causes for its slow recovery and stagnation while

---

<sup>5</sup> Also called transfer-learning, lesson-drawing, cross-national policy or benchmarking. Some of the policies adopted come from those devised by developed countries such as the United States of America, the United Kingdom, Australia, France, Canada, etc. The citizen charters, New Public Management, results-based budgeting and management are some examples of these influential models to Latin America.

Chile grew. Moreover, regarding administrative features, Dussauge (2013) adds low levels of administrative corruption, professionalization, the constitutional and legal authorities that Congress possesses, and the size of public administration, which are positive attributes in the Chilean government and contrasting in the Mexican.

Despite the potential bias when *small-n* cases are considered, these commonalities and variations in these features and comparing data in the outcomes from these countries produce a focused study research. By taking such an approach the study provides an original contribution by contrasting how institutional factors influence the effectiveness of evaluation in countries at different stages in the development of their evaluation systems (the relatively mature Chilean system and the still-evolving Mexican system). It will also offer a better understanding about the contribution of policy evaluation to good governance and administrative reform in emerging countries. The study needs to contrast institutional factors to show how these are connected with the effectiveness of evaluation.

### **Chapter outline**

The first chapter is a literature review, which provides an overview of the state of the art in evaluation studies, focusing on current debates regarding the effectiveness of evaluation as a part of the strategic planning and public policy process. The literature review also explores how evaluation has been pursued in developing countries, in the context of debates regarding public sector reforms and the doctrine of the New Public Management, with particular reference to their contribution to a 'rethinking of governance' in emerging economies. Building on related literatures this section reconceptualises the theory of evaluation use and the relevant institutional factors contributing to such theory.

The second chapter presents the methodology in detail to explain the research question, the dependent and independent variables, and the research strategy for data collection and analysis. Regarding the concept of effectiveness, this chapter explains the operationalisation of this variable, indicating the different approaches taken in the literature to its conceptualisation and measurement. Moreover, it also explains how various outcomes of evaluations are classified and related to effectiveness. It outlines the qualitative approach used to answer the research questions and to conduct the comparative study of the evaluation of programmes of the health sector. The chapter introduces the institutional or political factors identified in the academic literature, which will be tested in the following chapters.

The third and fourth chapters set out a documentary analysis of the historical evolution of evaluation activity in the Mexican and Chilean government. The chapters focus on identifying the major innovations in this field as well as the weaknesses of the evaluation systems, which might interfere in their effectiveness. The fifth and sixth chapters are concerned with the case study of the effectiveness of policy evaluation in the health sector in Mexico and Chile. In the Mexican case, the fifth chapter presents the evaluations conducted through the PbR/SED as the core of the Mexican system of policy evaluation and provides a comprehensive analysis of each factor in order to test the hypotheses regarding effectiveness, drawing upon documentary and interview material.

In the Chilean case, the analysis examines the evaluations performed under the results-based budgeting system at the core of the Chilean system of policy evaluation.

This is followed by a discussion of the level of influence of each factor to shed light on current practices of decision-making regarding its effectiveness.

The seventh chapter is centred on a comparative analysis between the institutional or political factors reviewed separately in previous chapters. It discusses whether the Chilean policy performs better than the Mexican based on these factors. The chapter then provides an in-depth analysis of the taking-up of recommendations per programme in order to examine whether the achievement of these recommendations is linked to the complexity of the task or the costs of implementing changes. Moreover, it also explores the factors shaping the effectiveness of the evaluation process through the comparative analysis of these factors within and across countries.

For instance, the differences found amongst these countries were the indicators regarding the variable IGR such as the reliability and quality of data; as well as the indicators of the variable of policy evaluation framework measuring perception of operationalization and methodological rigor. All these conditions impacting positively on the effectiveness of evaluation present in the Chilean but absent in the Mexican policy evaluation. At contrary, the availability of data and distribution of public resources of the variable IGR are outcomes present in Mexico, whereas in Chile were not identified in the course of time.

The final chapter summarises the findings of the dissertation and explores the theoretical and policy implications of the effectiveness of evaluation. In doing so, it set out some proposals for improving the effectiveness of evaluation in both countries and highlights the importance of adopting policy evaluation as an institutional process to improve public management and create public value.

## **CHAPTER 1. Broadening understandings of policy evaluation and effectiveness**

### **Introduction**

The aim of this chapter is to introduce the reader to the debate on policy evaluation and its effectiveness, identifying the significant literature and current debates regarding evaluation and its utilisation/use. Furthermore, it aims to identify the factors (institutional/political), which influence whether or not the recommendations made in evaluations are adopted. However, it cannot jump straight into these issues before introducing the contrasting approaches towards the evaluation process and the underlying theoretical frameworks regarding policy evaluation. The first section discusses some central concepts of this discipline and provides an overview of the state of the art in evaluation studies. Also, it reviews the historical evolution of evaluation and its emergence as a mechanism for improving management.

The second section examines how evaluation has developed in emerging countries as part of a wider trend towards public sector reform, transformation of the budgeting process and results based management, developments which are associated with the paradigm shift towards “the New Public Management”.<sup>6</sup> This section also emphasises the gap in terms of the State’s capability to understand and adopt policy evaluation. The third section will look at the debate on the effectiveness of evaluation as part of a broader re-conceptualisation of the theory of evaluation use. Furthermore, this section explains the importance of political and institutional factors in shaping the use of evaluation. Finally, the concluding remarks will review the importance of the effectiveness of evaluation in influencing decision-making, recognising the role of evaluation in improving public management and creating public value.

### **The roots of evaluation**

Evaluation began in early 2000 B.C. in China with the examination of officials in the civil service (Fitzpatrick et al., 2004: 31 and Guba & Lincoln, 1981: 1). Later on, during the 1800s, evaluation was directed towards education, social and health sector programmes in the UK as well as in the USA. These efforts continued in the 1900s with rapid advances thanks to technology, techniques and new methods of researching and assessment. Evaluation has come a long way since its beginnings in the last century. In the early years, it emerged as a tool to examine programmes, and eventually, evaluation approaches<sup>7</sup> became more formal, professional, systematic, rigorous and multidisciplinary cutting across several fields and disciplines.

The development of evaluation has been motivated by a number of rationales including to provide information and gain feedback of public policies; to increase knowledge of how programmes either contribute to the outcomes achieved or are implemented in complex organisations, and; to take better and effective decisions in the future.<sup>8</sup>

---

<sup>6</sup> This is confirmed by Power (1995, 1997, 2000 cited in Rose et al., 2006: 86) saying that technologies of budgets, audits, standards and benchmarks [implicit on policy evaluation] are forms of new public management developed under advanced liberalism (neo-liberalism).

<sup>7</sup> In evaluation field it would be more appropriate to use the term approaches or models (Alkin, 2004: 4-5).

<sup>8</sup> Mejía (2005) affirms that the most interested on evaluation systems are: i) citizenry seeking for information of how their taxes are used or the effectiveness of government in following electorate’s wishes; ii) politicians interested on learning about the quantity and quality of goods and services produced, and the extent that social problems have received attention and resources have been used properly; iii) officials interested on weakness and risks of management systems, programmes’ relevance, and whether public resources are sufficient; iv) policy makers seeking for the impact of their efforts and performance.

Since the 1960s and 1970s evaluation has been a commonplace in many developed countries, particularly in the USA, UK and some countries in Europe (most of the policy evaluation literature has dealt with these countries' experiences). Compared to developing countries in Asia, Africa or Latin America, the former properly developed theoretical knowledge and expanded the scope of evaluation (Alkin & Christie, 2004 and Rossi et al., 2004: 28). Moreover, recently, the topic has potentially increased its development mainly in the social field (e.g. education, health, housing, employment, agriculture) largely due to the large budget invested in these areas, where evaluation produces findings about programme's outcomes with the aim of improving them.

Another explanatory reason for expansion assumes that evaluation goes far beyond the immediate impact on society and focuses more on the control of finances, to provide accountability about how funds have been allocated given spending constraints. Indeed, it could be argued that evaluation emerges or expands more effectively in contexts where financial crises dominate, serving as a crucial tool to address inquiries about the efficiency of policies adopted and to control limited resources.<sup>9</sup> The development of evaluation in the 1960s and 1970s, drew upon methods such as Planning, Programming and Budgeting Systems (PPBS). First used in the Ford Motor Company, these approaches were extended to the public sector to improve their efficiency and effectiveness and to inform budget allocation decisions (Fitzpatrick et al., 2004: 36).

Clearly, performance-based and results-based budgeting become quite significant in terms of being a good incentive for management improvement.<sup>10</sup> In this context, evaluation has received notable attention by governments and international organisations such as the Organisation for Economic Co-operation and Development (OECD), World Bank (WB) and Inter-American Development Bank (IDB) and through paradigms such as New Public Management, which will be later explained.

Moreover, the literature suggests that evaluation can be conducted in numerous ways and with equally numerous approaches (e.g. methodologies and techniques), depending upon the purpose, resources and time.<sup>11</sup> Inasmuch as a government wants to know about the real impact of programmes, more complex and costly methods are required. International organisations such as the World Bank<sup>12</sup> propose a number of methodologies, which are widely accepted due to the influence they exert over governments. The principal methodology adopted – the 'logical framework

---

<sup>9</sup> This is visible by the UK HM Treasury's statement (2011: 9), which says that evaluation is more related to public resources '... to underpin practical resourcing and policy making exercises, and to provide accountability about how funds (public resources) have been spent'.

<sup>10</sup> Undoubtedly performance measurement has some implicit effectiveness or efficiency criteria (Knill & Tosun, 2012: 184).

<sup>11</sup> Models are developed under several visions and take different forms such as: cost-free evaluation; goal-free evaluation (Scriven, 1991); functional, tailored, comprehensive, and theory-driven evaluation (Chen 2004, Chen & Rossi, 1981, Rossi et al., 2004 and Weiss, 1997, 2007). The classification continues with naturalistic evaluation (Guba & Lincoln, 1981), which attempts to present the structure of reality based on informal reasoning of people's concerns, beliefs, perceptions and understandings (House, 1976: 37 and Wolf & Tymitz 1976-77: 6 cited in Guba & Lincoln, 1981: 78). The utilization-focused evaluation (Patton, 2004, Stufflebeam, 2004 and Weiss, 2004) for making decisions in collaboration with primary users focusing on intended uses; the CIPP model (Stufflebeam, 1983, 2004); and the administrative, judicial and political evaluation (Howlett et al., 2009 and Wu et al., 2010). Other scholars refer to the relation of costs to benefits and utility, performance and experiments (Parsons, 1995); prescriptive and descriptive model developed by Alkin (2004); objectives, management, consumer, expertise and participant-oriented approaches (Fitzpatrick et al., 2004); and the multiplist, design and naturalistic approaches (Parsons, 1995).

<sup>12</sup> The World Bank. Monitoring & Evaluation: Some Tools, Methods & Approaches.



approach' – has been promoted by these organisations to get homogeneity in the method,<sup>13</sup> followed by those countries utilising the Results based Budgeting approach.

Other authors such as Dye (2011: 326, 328) argue that many governments' approach to evaluation emphasises effectiveness reviews. For example, in developing countries involved in budgetary reform, administrative evaluation is widely accepted as the main mechanism given its emphasis on the efficiency and effectiveness of public services carried out within government bodies. Despite these developments, for most developing countries evaluation has emerged slowly compared to other stages of the strategic planning process (tailoring/design, implementation, control and monitoring), and seems to be the least developed discipline in terms of the State's capacity building.

Moreover, in practice, informal practices have prevailed over formal or scientific evaluations. Such outcomes might be due to the potential harm of evaluation to the manager's reputation [when results are not positive], the technical challenges presented in terms of expertise and data (Wu et al., 2010: 2), or the threats to the powerful when the programme evaluated is contentious (Scriven, 1991: 41).

The study of the use of evaluation has gained currency (Weiss, 1998) to monitoring such evaluation practices. Research has highlighted the importance of reviewing the effectiveness of evaluation as a method for closing the policy cycle and to acknowledge creation of public value. However, as Stiglitz (1998: 287 cited in Wiesner, 2011: 27) indicates 'Evaluation itself is an institution...' this perception might be due to all circumstances both internal and external to be considered in order to make evaluation effective. After discussing the evolution of evaluation and the utilisation of evaluation in multidisciplinary fields such as engineering, business, spatial, health, agriculture, army, politics and so on, this diversity might be the reason for the large number of definitions of evaluation according to theorists, academics and scholars in different contexts.

The more comprehensive concept of evaluation as assessing the performance of government and covering the entire public administration relates to the final and essential phase in the policy process, providing an explanation of the outputs, a learning process to find out about the consequences of public policy, an understanding of how and why it works or does not work, and an assessment which involves value judgments of the overall effectiveness of policies and public programmes in terms of the objectives, targets and goals that should be achieved (Cochran & Malone, 2010, Dye, 2011: 323, Hill, 2009: 8, Howlett & Ramesh, 1995: 168, Stufflebeam & Shinkfield, 2007: 9 and Weiss, 1997: 51, 2007: 77). Scholars also point out how evaluation should be linked to the government's commitment to the creation of public value.

The latter concept refers to the equal satisfaction of the human needs e.g. poverty reduction, improvement in education or health (Hintze, 2001: 40; 2003: 3), embracing a wider concept of people's living conditions. The worth of evaluation depends on the extent to which they can provide better quality in policy adoption and decision-making regarding the allocation of public resources oriented to produce such public value. Moreover, from a holistic perspective of policy, in order not to be

---

<sup>13</sup> In Mexico and Chile have been introduced the logical framework approach as a planning tool to interconnect through indicators, institutional inputs with outputs, outcomes, institutional, sectorial and national objectives. This methodology is compulsory by law and operates through their evaluation systems.

judge and jury the evaluation should be operated by external organisations independent from the executive or through the legislative branch – whose role is to represent constituency demands – or the agency that implement those public policies. Thus, the state of art in the conduct of evaluation is where the operation of evaluation is located at a superior level and is autonomous from the State structure.

### **Evaluation in emerging countries and the New Public Management**

The so-called third wave of evaluation<sup>14</sup> emerged during the late 1980s and 1990s (Wollman, 2003: 14), and recently, the growth of evaluation of public management has been closely associated to a wider process of public sector reform. The well-known cases of such reforms took place in New Zealand, the United Kingdom and the United States of America (Rhodes, 1991). In Latin America, efforts to develop reforms have been in response to historical problems in the region, such as the quality of public management and its capacity to implement policies and allocate resources efficiently. These efforts belong to the second generation of reforms<sup>15</sup> following the macro economical adjustments in the 1980s (Naím, 1994).

In the study “El gobierno del gobierno” (the governing of government), Aguilar (2013: 27) claims that there were two motivations for boosting public administration reforms in these countries. Firstly, the debt crisis, which forced them to implement a fiscal adjustment policy to achieve fiscal equilibrium, which in turn required the improvement of programmes, policies, public goods and services’ efficiency and of government performance. Secondly, the political crisis of these regimes characterized as corrupt and inefficient, forced them to democratize and introduce a more participative way of governing, one which was transparent, legal, and accountable. These issues converge into the New Public Management ‘NPM’, which responds to the State’s desire to be competitive, efficient, and to improve the macroeconomic performance of the national economy.

The NPM focused on ensuring citizen consumer satisfaction through greater use of market and contractual mechanisms, government transparency (Jones, 2002: 84 and Jones & Kettl, 2003: 10), and performance-related pay in public organisations (Hood & Peters, 2004: 278). These principles (Barzelay, 2001: 3, Hill, 2009: 291, Hood, 1995: 96, Osborne & Gaebler, 1992 and Peters, 1996: 13) belonged to a wider movement of a ‘rethink of governance’ in many developing countries following the example of developed economies. For Arellano & Ramírez (2000: 1-2) proposals to reconfigure the public sector were implemented to create governments based on results and real impacts.<sup>16</sup> In this context, the NPM (Hill, 2009: 291) responded to dilemmas about public bureaucracy and accountability, emphasizing the need to adopt an ‘evidence-based learning’ approach (Hood & Peters, 2004: 278).<sup>17</sup> A key factor in the reforms of

<sup>14</sup>Wollman (2003: 13) says that public-sector reform and evaluation are interlinked and in sequence over time: “the first wave of evaluation during the 1960s and 1970s; the second wave beginning in the mid-1970s; and a third wave related to the New Public Management (NPM) movement”.

<sup>15</sup>The Washington Consensus dates back to 1989 and is a set of guidelines of policy and economic reforms imposed by the Washington-based international financial institutions to developing countries in Latin America that should be adopted in order to reform their economies. The first-generation of reforms include macroeconomic stabilization, reduction of inflation, budget cuts and privatization, whereas second-generation focus on reforms of the State, civil service, better regulation, tax collection, delivery of public services and so on. For further information see Naím, 2000, 2000b, Navia & Velasco, 2003 and Williamson, 2002, 2003.

<sup>16</sup>This new way of government started with a management strategy, which was focused on performance achievements, through key performance indicators as well as effective monitoring and an evaluation system (Rhodes, 1991; Schedler, 2003).

<sup>17</sup>Hood (1995: 96) identifies seven ‘doctrines’ i) Hands-on professional management in the public sector; ii) Explicit standards and measures of performance; iii) Greater emphasis on output controls; iv) ...disaggregation of units in the

public administration proposed by the NPM is the transformation of the budgeting process, which means moving from a system of control of expenditures towards a system oriented to results.<sup>18</sup>

However, based on some arguments regarding the experience of and evidence from those countries where reforms originated (Jones, 2002: 84, Jones & Kettl, 2003: 2 and Manning, 2001: 297), it could be asserted that the NPM as a global trend is more a paradigm inspiring other countries than a guideline to promote public sector reforms. It appears to be 'somewhat mystical in essence' (Hood & Peters, 2004: 268), but without a consensus about its features and evidence of the real impact. For Manning (2001: 297) "the victory of NPM was very partial". For some developing countries the results may have been slightly better in terms of improving public sector responsiveness and efficiency to the political principals. However, in other respects it appears that hierarchical bureaucracies have not been downsized or reformed by market mechanisms (Cejudo, 2008 and Manning, 2001: 300).<sup>19</sup>

While emerging countries – at least in Latin America – arrived late to such reforms, compared to developed economies, the important factor to take into account is that such waves of reform were important in fostering the use of evaluation in policy-making. However, these countries have developed less systematic evaluation processes than was predicted and they have been more focused on government outputs than outcomes for citizens (Jones, 2002: 84). Hill (2009: 299) also identifies the potential contradiction within the NPM, which stresses accountability on the one hand, but on the other hand, advocate the traditional forms of top-down control.<sup>20</sup>

A common mistake in the NPM reforms is that they frequently adopted a 'one-size-fits-all' approach to bureaucracy (Hood & Peters, 2004: 278); as Ormond & Löffle (1999: 2) suggest, this kind of ideal form of 'Weberian' public administration never survives in reality. For Arellano (2002: summary), the solutions offered by the NPM do not fit in public administration with dominant constraints,

In other words, no administrative reform consider itself as an exclusively technical transformation and based on concepts such as 'quality', 'innovation', 'client', 'results-based evaluation' succeed in the short term. In administrative "habitats" based on loyalty with extensive clientelism, less accountable, who are used to lack of transparency and over regulation.

While the NPM's doctrines appear to have been implemented, and have followers in developing countries, the extent of change has been only superficial.<sup>21</sup> Consolidation of the State's democracy is attractive for such countries in 'transition', in which the NPM makes promises for greater public service responsiveness, accountability, and entrepreneurial public sector to grow their economies. Indeed, such waves of reform linked the adoption of policy evaluation to a process of developing the state's

---

public sector; v) ...greater competition in the public sector (to this may be added actual privatization); vi) ...private sector styles of management; vii) ...greater discipline and parsimony in resource use.

<sup>18</sup> Some of those characteristics are the basis of a *new managerial management* and *organisational cultural model* (SHCP, SED, 2007: 13) in order to create public value inside the Mexican public administration.

<sup>19</sup> The article by Cejudo (2008) suggests that not all reforms were inspired by the New Public Management, these were a consequence of political democratization and economic liberalization in the Mexican public sector.

<sup>20</sup> Aguilar (2013: 28) adds that the NPM and New Governance agree on something, a narrow interest to discern the decisional process within the governing of government, '...how the government decide about objectives, priorities, regulations, organisation, rules, actions, actors, resources, auditing and control systems and the performance of evaluation of the agencies inside public administration...'

<sup>21</sup> As those suggested by Jones & Kettl (2003: 2), a smaller, less interventionist and more decentralized government, and the visible improvements on governments' efficiency and effectiveness, despite such reforms are often 'evidence free' (Hood & Peters, 2004: 278).

institutional capabilities.<sup>22</sup> These capabilities are built upon technical-administrative and political capabilities (Repetto, 2003: 1,3), where interaction among individuals (officials) and groups operate within the framework of rules (formal and informal), organisational routines and practice.

Increasing administrative capacity through institutional development has been an important component for ensuring the State modernization, facilitating democracy and sustainable development (Prats, 1997).<sup>23</sup> In this sense, the adoption of evaluation can be seen as closely linked to these reforms efforts, signaling the modernization of the State, especially for those developing countries seeking to change their authoritarian and bureaucratic image in the world. Beyond everything, a major achievement of the NPM was bringing up the Results-based Budgeting as the epicentre of policy evaluation. The experience observed in developed countries (USA, UK, Australia, Canada, New Zealand), which positioned budgeting reforms as the core of evaluation, indicates that incentives should be the basis to influence budget decision-making.<sup>24</sup> Indeed, the value of such reforms given by political actors within the ministries or departments of finance was due to their power, leadership and experience.

Exploring whether and how evaluation 'demand' arises, scholars agree that the growing interest in appraising government performance in all dimensions (individual, programme, policy, organisational) is primarily linked to the scarcity of resources and interest in being efficient.<sup>25</sup> For Bangura & Larbi (2006: 83) it was more linked to the second-generation of state reforms during the 1990s, which emphasized the effectiveness of the public sector among others.<sup>26</sup> In general, the priority of governments is to seek through public expenditure the provision of services and deal with the demands of the public. However, public resources are insufficient to meet such demand, at least in developing countries.

Therefore, in principle the more evaluation is pushed forwards by government the less economic uncertainty and scarcity of resources the government have to deal with (see Bamberger, 1991 and Wiesner, 2011). Given this, evaluation should have become a powerful tool to determine whether programmes are effective and where to allocate these resources. For others, evaluation has been promoted by international

---

<sup>22</sup> To explain the utility of institutional capability, it is important to note that everyday public institutions and their stakeholders make decisions to solve problems distressing society under a fundamental premise (Barzelay, 2001: 3) that policy decisions amount to a substantial shift in the governance and management. The way to respond effectively to these public issues is closely linked to the rules, routines, practices, as well as other key factors pointed out by Repetto (2003: 6-8) such as the legitimacy and accountability of priorities and the particular interests of those involved in public-decision making, amongst others. Nonetheless, there is a gap between these countries and developed nations in terms of capability building. The reasons could be attributed to their position and control under donor agencies, or the consequences (transparency and accountability) that evaluation could bring to their governments' stability.

<sup>23</sup> To consolidate institutional capability it is not enough to have meritocratic and encouraged bureaucracies, with the ability for management under uncertain contexts (Repetto, 2003: 35). It is more related to institutional strengthening, which means the relation between the State, politics, market, administration, constituency and international context.

<sup>24</sup> Incentives align the process to efficient and sustainable solutions. Whether politics achieve linking results with the more budget allocation, evaluation will act as an incentive and the evaluation market will develop (Wiesner, 2002: 139). However, a study by the OECD (2009) argues that most countries have few incentives for bureaucrats, ministers or legislators to encourage learning or to base their decisions on performance data.

<sup>25</sup> Wiesner (2002: 140, 142) notes three sources: i) countries have to assess whether to do budgetary cuts or remain it; ii) governments cannot be passive if they do not want to pay high political costs; and iii) civil society has become more aware of its capacity to claim accountability from governments.

<sup>26</sup> This is known as the 'four Es': the effectiveness of public sector intervention in terms of coverage and quality of service; the economic efficiency of service delivery; the improved equity of service delivery; and the creation of an enabling environment for private sector development.

organisations or donor agencies<sup>27</sup> (Bamberger, 1991: 329, Dahler-Larsen, 2005: 621, Horton & Mackay, 1999: 11 and Wiesner, 2002: 132), whose aim is to support the improvement of evaluation capacity at the central government level.

What factors help to explain the attempts to develop evaluation systems in Latin American countries? The analysis of the circumstances under which evaluation emerged in developing countries shows numerous possible reasons: pressure by international organisations to develop best practice models, measure performance and create data systems (Andrews, 2008: 5); the influence of wider NPM reforms<sup>28</sup> (Ormond & Löffle, 1999: 1); the result of the economic liberalization and political democratization processes (Cejudo, 2008: 120-3), or; a state's interest in improving results due to domestic pressures.

Moreover, this seems to be the tendency of policy evaluation and one of the open questions might be elucidated in Rossi et al.'s (2004: 14-5) work. They suggest that the 21<sup>st</sup> century is dominated by conservatism on fiscal policy, the devolution of responsibility to the states and skepticism about programmes and their social impacts. Thus, the question since the 1990s is whether they (policies) are poorly or properly conceived, appropriately or inadequately implemented, and effectively or ineffectively administered and how effective the policy evaluation implemented is.

### **A re-conceptualisation of the theory of evaluation use**

According to the dictionary, effectiveness is the degree to which something is successful in producing a desired result; success (Oxford, Merriam-Webster), and the quality of being successful in achieving the results of what is wanted (Cambridge). Much of the literature refers to effectiveness as use or utilisation of evaluation (evidence-based policy). Within the evaluation literature, Leviton & Hughes (1981: 526) confined the concept of utilisation to the evaluation results for programmes and policies only; whereas Teirlinck et al. (2013: 369) refer to usefulness as a standard of evaluation "measured in terms of policy consequences related to the conduct of the evaluation".

For Dahler-Larsen (2005: 623), the utilisation dimension is based on the premise that evaluations are produced to play a role in future 'practical action situations' although they are not always used. A literature pioneered by Weiss (1998: 31) describes a connection between use and change based on findings "Any theory of evaluation use has to be a theory of change". Moreover, the theory of utilization-focused evaluation developed by Patton (2004: 278) is widely accepted because it attempts to identify a group of intended users within the decision-making process to improve utilisation.

In this sense, scholars of evaluation have identified five ways in which evaluations are used: i) instrumental, when findings modify the programme evaluated (Rich, 1977 cited in Leviton & Hughes, 1981: 528 and Stufflebeam & Shinkfield, 2007); ii) conceptual, when findings help staff to understand the programme in a new way (Fleischer & Christie, 2009, Rich, 1977 cited in Leviton & Hughes, 1981: 528, Sager & Ledermann, 2008 cited in Ledermann, 2011 and Weiss, 1998); iii) knowledge-generating (Patton, 2012) or enlightenment (Stufflebeam & Shinkfield, 2007), when

---

<sup>27</sup> The World Bank, the Organisation for Economic Co-operation and Development, the United Nations, the United States Agency for International Development, the European Union and the Inter-American Development Bank.

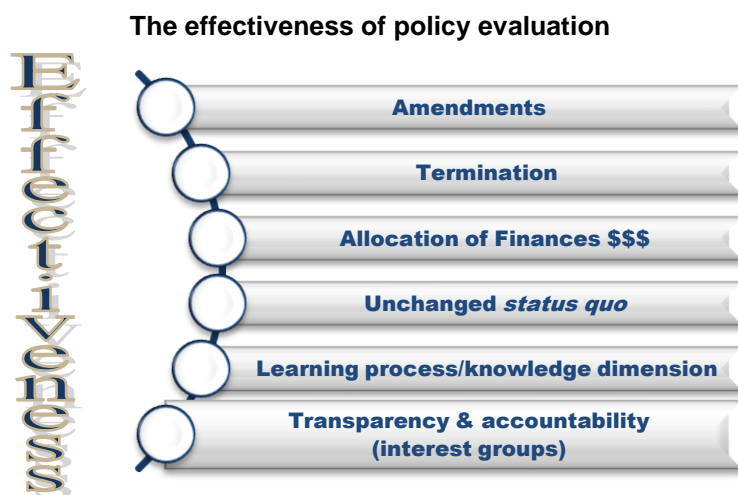
<sup>28</sup> Albeit those developing countries often lack the evidence and results of developed countries, they still follow the NPM doctrines and reforming governments as a guide to construct their future.

findings add knowledge to the field; iv) process, when findings produce cognitive, behavioural, program and organisational changes (Weiss, 1979); and v) persuasive or symbolic, when findings persuade stakeholders about programme or organisation's accountability (Florio et al., 1979 cited in Leviton & Hughes, 1981: 528, Patton, 2012, Stufflebeam & Shinkfield, 2007 and Weiss, 1998). However, Alkin & Taut (2003 cited on Herbert, 2014: 391) describe this process as “conducting evaluation as a symbolic act without intending to use the findings”.

In this research, the concept of “effectiveness” is employed instead of “utilisation” or “use” in order to illustrate how it can shed new light on the understanding of the evaluation process. The effectiveness of evaluation is defined in terms of the *impact of evaluation on the policy itself*; that is, the extent to which the evaluation of a policy or programme leads to changes in that policy. In other words, effectiveness is measured by assessing how many of the recommendations established in evaluation reports are adopted by policy makers and programme operators in response to those evaluations.

Assuming public policy aims to solve specific problems, the effectiveness of evaluation can be assessed as the extent to which programme improvement in a particular policy area can be attributed to the evaluation performed.<sup>29</sup> It is to find out how the key government outcomes compared to others are provided, considering contextual factors (Andrews, 2008: 36), and the criterion of evaluation about how an outcome is achieved (Dunn, 2012: 196), as a result of resource allocation and of sustainability of the desired results (Hintze, 2003: 4). The concept explained here is similar to that offered by Weiss<sup>30</sup> in a broad way. Thus, effectiveness is the essence of evaluation results both in public and in private sectors (Hintze, 2001: 14), the core criterion to demonstrate the value of evaluation and the purpose of this research. In this context, the criteria for judging the effectiveness of evaluation could take simpler forms compared with the taxonomy already mentioned, as shown in the figure below.

Figure 1



Source: author's own elaboration.

<sup>29</sup> As Essama-Nssah (2013: 1) affirms “Effective evaluation can produce reliable information on what works, what does not, and why”, and presumably answer questions such as: i) whether they are doing the right things in the right way; ii) if what they are doing is working and worth the cost, and iii) how the observed outcomes can be explained, which policy makers need to know.

<sup>30</sup> For Weiss (1972: 18) is to decide whether to continue or discontinue the programme, improve practices and procedures, add or drop specific programme strategies and techniques, institute similar programmes elsewhere, allocate resources among competing programmes, and accept or reject a programme approach or theory.

In the first place, the evaluations might recommend some “amendments” or modifications to policy in order to improve it. Secondly, the evidence from evaluations might highlight risks, high costs or the need for improvements in current policy/programmes itself, thereby leading to their “termination”.<sup>31</sup> In the third place, another indicator of change in the programmes is the “allocation of financial resources” during the budgetary process,<sup>32</sup> as enacted in law either by the executive or legislature at Congress. The latter categories could be linked to the instrumental use mentioned earlier, as Herbert (2014: 390) notes, “This type of use depends on evaluation results being the basis of a decision”; as well as to the process use of evaluation involves either some amendments, allocation of resources or cancellation of the programme or policy evaluated.

In the fourth place, the operators placed a programme in the “unchanged” category, which may indicate the maintenance of the *status quo*, with no consequences for programmes. Concerning this, an element gaining relevance in the literature on evaluation is the learning and knowledge dimension. Despite few indications of empirical links between learning and usefulness (Teirlinck et al., 2013: 369)<sup>33</sup> many programmes are going through an internal “learning process” due to evaluation, which is not reflected in specific changes of the programmes, and which should be properly acknowledged. In the fifth place, this learning category includes the role of learning inputs, outputs, channels, the dimensions enabling learning processes in policy-making, the conditions of failure or success (Teirlinck et al., 2013: 367), and the nature of problems and solutions [active learning] (Howlett & Ramesh, 1995: 175).

It also comprises the perspective of involving practitioners and what can be learnt from their experience (Coote et al., 2004: 9) as a way of achieving effectiveness. Similar to the conceptual use noted by Weiss & Bucuvalas (1980 cited on Herbert, 2014: 390) “...where the use of an evaluation is not direct, but rather the information is absorbed into the common knowledge...” As Howlett & Ramesh (1995: 170) claim “the greatest benefit of policy evaluation is not the direct results it generates but the process of policy learning that accompanies it”. For Weiss’ (1979: 430) this is the enlightenment model, which claims,

Research sensitizes decision makers to new issues and helps turn what were non-problems into policy problems... In the long run, along with other influences, it often redefines the policy agenda.

This is similar to Sundell’s (2014: 34) view of the indirect longer-term influence of evidence and evaluation studies on the organisational cultures of decision-makers. However, such “learning” processes (see knowledge dimension on Dahler-Larsen, 2005) generally prove to be difficult to measure and classify unless those implementers or coordinators of policy evaluation refer to it. This leads to a perspective, which might seem contradictory to the definitions already mentioned, recognizing the validity of Picciottos’ view (2005: 348 cited in Wiesner, 2011: 30) that,

---

<sup>31</sup> An example of disruption on the effectiveness of policy evaluation and main concern of this study is when termination of programmes is needed. In this sense, although the chances of programme termination increased with external consultants (Dahler-Larsen, 2000), according to his study in Danish municipalities shows they participation made no difference to the continuation, adjustment, and termination of programmes. This lack of uptake by policy-makers is attributed to institutional and ideological factors as mentioned above.

<sup>32</sup> At this point, it is important for the reader to keep in mind that results based budgeting is only an outcome of the policy process, and should be distinguished as part of a structure of incentives and not as a purpose in itself, because it only brings a partial perspective on policy evaluation.

<sup>33</sup> The three dimensions noted by Teirlinck et al. (2013: 369) of the learning effect of evaluation and factors to its utilization are: i) the evaluation content or the usable knowledge of recommendations and the evidence base of the study; ii) the process dimension from the angle of stakeholders’ involvement; and iii) the evaluation design and the methodological setting of evaluation studies.

If evaluation is only about learning, it does not make authority responsible. If it must churn out lessons to justify itself, it will generate an oversupply of simplistic and pious exhortations and platitudes. Worse: evaluators that do not encourage accountability for results fail to provide incentives for learning.

Lastly, another source of use comes by the citizenry and the NGOs, mainly as a form of accountability of government's performance and similar to the symbolic use of evaluations with the aim to legitimize its use. However, it is not the aim of this research to follow-up the use of evaluations by these actors outside the government due to limitations of time and length of the research. In any case, the evaluation should be used to improve the programmes in different forms such as those already mentioned either by the policy makers and officials, by the executive when tailoring the budget, by the legislative when approved the budget and/or by the citizenry.

### **What explains the effectiveness of policy evaluation?**

In addressing this question, a primary observation is that potential recommendations become actions when policymakers assume their responsibility of decision-making to make real improvements in programmes. However, the way in which policy makers, politicians, officials and bureaucracies value results usually has some ethics and politics implicit in decisions, which could influence an effective conclusion of the policy process. Thus, the potential to impact the decision-making process through evaluation only exists to the extent that programme effectiveness is valuable for those actors involved (Weiss, 1972: 5). In this process, Wiesner (2011: 34) argues that the degree of the demand-driven approach is more effective than supply-side, because the former provides incentives that encourage and reward results.

In other words, the demand for evaluation should emerge from those levels in which the policy is needed such as recipients and clients whose interest claim for results (effectiveness) from the government. It seems like a bottom-up approach of evaluation, yielding policy-makers, stakeholders and users' opinions to formulate the policy rather than a decision only made by the top-level officials and politicians. Arguments about effectiveness go in both ways. On the one hand, Hill (2009: 169) and Dahler-Larsen (2005: 619) assert that there is rarely a connection between evaluation and policy improvement, and consequently, the conclusions of evaluation are rarely applied,<sup>34</sup> as Dye (2011: 335) affirms, even when there is evidence of negative findings. Teirlinck et al. (2013: 368, 370) refer to the lack of uptake of recommendations as a disruption to the policy cycle between the evaluation process and the agenda of policy design.

Indeed, they offer this argument based on their case study saying "unexpected weakness with respect to the lack of broad stakeholder involvement in evaluation exercises" (Teirlinck et al., 2013: 368, 370). Seeking for the reasons why decision-makers avoid or ignore findings, the UK Cabinet Office (2008 cited in Sundell, 2014: 25) presents a realistic explanation of the process. For instance, they mention that the research itself was not always valued or well communicated within their own organisations; internally conducted research or commissioned research from consultants was more likely to be regarded as more relevant than academic research; external academic research was not seen as sufficiently timely, or as not sufficiently relevant to users' current needs; and research was much less likely to be used when findings were controversial or when findings upset the status quo.

---

<sup>34</sup> Reasons why it is so difficult to eliminate failed programmes and policies by governments are: i) concentrated benefits and dispersed costs; ii) legislative and bureaucratic interests; iii) incrementalism at work (governments seldom undertake to consider any programme as a whole in any given year).



On the other hand, from the perspective of other academics such as Askim (2007: 454) and Pollitt (2006a: 38; 2006b: 5) there is empirical evidence about utilisation, especially by politicians, although they made little use of it. Indeed, Askim's study (2007) proves that performance data and evaluations are used in the decision-making process. However, utilisation appears to vary according to the circumstances in different policy sectors<sup>35</sup> and countries. His analysis extends to particular patterns of use e.g. periods of time or interest groups. In surveying Norwegian councillors, his study sets the decision-making process in a timeline known as pre-decisional, decisional and post-decisional stage to explain how they respond differently over the policy cycle. He cited other scholars to show contradictory results around the topic. For example, Johansson (1995) claims there is more use of data when actors are dealing with 'hard core' tasks (e.g. technical services) than with 'soft' (e.g. social services).

By contrast, for Greeve (2003), decision increases when the complexity of the task decreases, or as Askim (2007), Bogt (2003), Macintosh (1985, 1994) and Wilson (2000) observed, where observability over programmes is high. Therefore, as Askim argues (2007: 464) the difficulty here is to distinguish between a soft and hard task, high or less observable, and complex or easy policies. Patton (2012: 136) adds to the complexity of studies of evaluation use by highlighting that "Evaluation findings typically have technical and methodological weaknesses; data must be interpreted; [and] other contextual factors must be taken into consideration". In another study, Melkers & Willoughby (2005) prove that data was most useful during budget preparation.

A different perspective on exploring the conditions where evaluation leads to policy changes is developed by Ledermann (2011). Through a case study of 11 programmes' evaluations commissioned by the Swiss Agency for Development and Cooperation and carried out by external evaluators, factors were connected such as the level of conflict (low, high) among the stakeholders and the amount of pressure for change (low, high) acting as a mechanism (awakener, trigger, referee and conciliator) to measure how much these influence the use of evaluations. The study also adds the novelty value and evaluation quality as conditions *sine qua non* for changing decision-making. Ledermann concludes that in contexts of low pressure and conflict, evaluation contributes to decision-making to the extent that evaluations are of good quality and novelty. However, when high pressure and conflict prevails, evaluation is used as a referee to decide changes neither considering novelty or quality.

Conversely, the findings of the study of Jennings & Hall (2012: 16) suggest that in high-conflict environments, evidence-based decision-making is less likely to occur, whereas when a political conflict over the agency declines the evidence-based practices increase; although this author noted that these depend on who controls the agency. However, this is a partial perspective based only in two conditions, leaving aside the inherent political factor. This means that despite factors are *ceteris paribus* or remain the same and keep controlled by the agency, the decision made by actors after evaluation is uncertain and unpredictable due to other conditions examined here.

In reflecting on the insightful arguments of these academics, this research further discusses under what circumstances evaluation is more likely to be used. Although

---

<sup>35</sup> Askim (2007: 464) concludes that "utilization is higher among councillors working with elderly care, administrative and educational affairs" and constant across policy sectors increasing during decisional stage (later on explained) working with administrative affairs and technical services.

evaluations are government-ordered, in general policy-makers, officials and policy analysts are rarely or unsystematically involved in closing the policy cycle. Appropriation of evaluation results by political actors remains a difficult task and varies from country to country. Moreover, in large administrative systems (e.g. the Mexican government) it is complex to isolate the causal effects of evaluation (Dahler-Larsen, 2000: 77). To identify these causal factors, institutional or political, requires a case study to provide an explanation of the effectiveness of policy evaluation.

### **Factors influencing the effectiveness of evaluation**

To determine which factors are likely to impact the use of evaluation results, a review of the literature regarding this topic is applied here highlighting a myriad of factors identified by many academics that might affect effectiveness. The discussion will focus on the extent to which this literature engages with the variables noted as the most relevant. In doing so, Leviton & Hughes (1981) agreed there are five variables on the utilisation between users: the relevance of evaluation addressing clients' needs and timeliness of information and the importance of good communication between evaluators and potential users.

Another issue is related to the information processing, as these authors claim "The way evaluation is presented to users affects their comprehension and thus the extent of use". Thus, evaluators should be aware of the way reports are communicated to users. The credibility of the evaluation producer and the quality and objectivity of evaluation are also likely to enhance utilisation. Utilisation might also be influenced by user involvement and advocacy and the commitment of decision-makers.

Other authors (Fleischer & Christie, 2009) identify four factors mainly related to the role of evaluator: organisational outcomes associated with evaluation capacity building (e.g. organisational support structures, capacity to do evaluation, evaluative inquiry); the role of evaluator, which can be influenced by the purpose and approach of the evaluation, stakeholders' needs and evaluator's epistemological preferences; stakeholder's involvement; and misuse of evaluation findings based on intentionality (intentional, unintentional and nonuse). Cousins & Leithwood (1986: 347-8) grouped in two those factors affecting effectiveness. One set of factors related to the conduct of an evaluation such as relevance, credibility, quality, communication, findings and timeliness while the other set was concerned with decision or policy setting such as information needs, decision characteristics, political climate, competing information, personal characteristics and commitment/receptiveness to evaluation.

However, the impact of the latter factors regarding decision or policy-making seems to be mixed. This group of factors not only include organisational characteristics but they extended to information, commitment and decisions, which might belong to another set of characteristics related to actors and policy makers' decision who operates within a political context. More recently, the debate in the field also emphasizes different stances influencing effectiveness. Contandriopoulos et al. (2010) and Contandriopoulos & Brousselle (2012) identify the social structuring, level of polarization (problematization, prioritization and salience of the issue, and criteria to assess solutions) and cost-sharing dimension (knowledge exchange imply time, money, attention). However, as these authors affirm, these stances belong to the 'collective-

level knowledge use<sup>36</sup> to attain knowledge from cumulative experience, in which evaluation results could be used to the 'learning process' as explained here.

In order to provide a comprehensive overview of factors facilitating the use of evaluation findings, these will be analysed from a different perspective and in order of importance as a narrative in terms of whether and how they are implicit in the effectiveness of policy evaluation. It is important to bear in mind that institutional or political factors studied here embrace those relatively under-explored by scholars and academics within a comprehensive approach. It is therefore of interest of this research to study those factors which are at the foundation of policy evaluation. This is the conceptual basis for the development of these M&E models in Chile and Mexico.

With respect to the role and attitudes of individuals toward evaluation a key factor which might affect the effectiveness of evaluation and has not given much attention in the literature, is the nature of *Inter/Intra Governmental Relations* 'IGR'. Despite the difficulty to identify the most relevant factors promoting the use of evaluation, the findings of a recent analysis of 41 empirical studies of evaluation use conducted between 1986 and 2005 (Johnson et al., 2009: 388) highlight the importance of engagement, interaction and communications between stakeholders, clients and evaluators as the "key to maximizing the use of the evaluation in the long run". These are characteristics of the concept IGR, as explained in the methodology chapter, this structure of relationships embraces diverse forms of coordination such as negotiation, exchange and legal authority (Franz, 1985 and Kaufmann, 1985).

It also endorses the idea that the availability, timing, reliability and quality of data, as well as economic resources represent a key input to make the evaluation effective. How these elements (forms of coordination, information and economic resources) are incorporated in the policy process will have an impact on the level of agreement or conflict in the IGR. As Franz (1985: 484) argues 'the more intergovernmental linkages exist the more the single organisation gains some degree of autonomy as to its interorganisational behaviour'. A general agreement is that such forms of coordination often determine the decision or nondecision-making regarding effectiveness though they are difficult to validate. This may be because the decision-making process occurs in the political arenas where the human relations, ethical issues and interests could bias any effort,<sup>37</sup> along with the difficulty of proving these intangible and subjective issues.

For example, at Congress for the budgetary process where politicians need to pay more attention to programme performance when preparing budgets instead of discretionary decision-making. The importance of the coordination between these actors and the key role of Congress is also noted in Bundi's study (2016) about the demand for evaluation in the parliament, who is "an important demander of evaluations" to make decisions and fulfil their oversight function. From a comprehensive standpoint, these are the subtle nuances of politics and ethics implicit on the decision-making of utilisation (effectiveness) of evaluation. However, some of

---

<sup>36</sup> Contandriopoulos et al. (2010) and Contandriopoulos & Brousselle (2012) refer to it as "the process by which users incorporate specific information into action proposals to influence others' thought, practices and collective action rules".

<sup>37</sup> For instance, Fleischer & Christie (2009: 172) in their comparative study of the Preskill and Caracelli's and the U.S. American Evaluation Association (AEA), as well as in Ledermann's survey (2011: 15) pointed out to the hurdle for utilising evaluation data when most of decision makers reject conclusions based on beliefs and values rather than data. This belongs to the political arenas, which go further than the user or evaluator characteristics of the human factor identified above and will explained later on.

these characteristics are not referred to in detail by authors yet and considered here as a key matter.

As was mentioned earlier, political factors shape every public decision (Mayntz, 1994: 80-4) including the decision-making process around policy evaluation. Thus, the intentions of those demanding the evaluation (Teirlinck et al., 2013) along with the decisions of actors responsible to make effective the evaluation are influenced by such political circumstances, which could either modify or not the programmes according to the evaluation results expected. Indeed, it might not be wrong to argue that the possibility of success in expanding the use of evidence-based policy will depend on the ability of political actors to move away from the *status quo* to integrate a more results-focused approach based on the effectiveness of evaluation into the strategic planning process.

Another perspective recently emerged from the debate is the theory of influence by Kirkhart (2000), in which dimensions (source of influence, intention and time frame) exert greater impact on the use/utilisation of evaluation. For this author, time dimension occurs in three stages: immediately, at the end-of-cycle, and in the long term of evaluation. However, the time dimension refers only to the influence of evaluations over time, this is in the short, medium and long term. To make effective evaluations, a main concern of this study is the need to provide data (evaluations) timely to actors shaping the decision-making and continuity of the process e.g. during the budgetary process. Indeed, this element is closely linked to the framework factor to strengthen the structure of the M&E system.

Regarding the source of influence (agent of change), this identifies two elements: the results and process-based influence to generate change in the policy or programme. Whereas the intention (intended or unintended) dimension is explicitly described in the purpose of the evaluation, theory employed and evaluator-client contract. This dimension constrains the intentionality to the tailoring process of evaluation and does not identify an intentional/unintentional intervention of those responsible of decision-making once they have the evaluations results. These are the subtle nuances of politics implicit in the study of the IGR as here is examined.

Another discussion in debate is how mis-utilisation or 'non-change' decisions represent a risk to close the policy cycle (Teirlinck et al., 2013: 368). In situations where all conditions are settled such as the accepted rules of the game, instruments of force, values and power relations but there seems not to play for decisions effectively, presumably a *non decision-making* exists (Bachrach & Baratz, 1963: 641) and should be analysed on detail. Indeed, the mere existence of evaluations is not sufficient to call for decisions by actors and prevent a latent misuse of them despite these dominant values. For example, when findings suggest major changes or cancellation of the programme/policy but this represent a political risk in terms of votes, might be a nondecision-making prevails due to political consequences to the party in power. Here, power emerges as the realisation of preferences in the decision-making process (Hay, 1997: 46).

For Lukes (2005: 22), this is the *two-dimensional* view of *power* operating as a "qualified critique" to prevent potential issues and conflict interests over decisions. The latter example indicates either a conflict within the structures of power<sup>38</sup> relations

---

<sup>38</sup>For Bachrach & Baratz (1962: 947) a power structure is "an integral part and the mirror image of the organisation's stratification".

or how the nondecision-making is also exercised excluding relevant issues from the process without any conflict and under apparent consensus, as Bernstein (2000: 5) asserts to say, power relations create, legitimize (relations of order) and reproduce boundaries between different groups, discourses and agents. Thus, the subtle nuances and meanings emerging from this process are closely linked to the relationship between actors and branches as well as the circumstances around the decisions made by actors are embraced into the IGR and these will be discussed here. Then, maybe the questions are *Under what circumstances decision-makers are likely to use evaluation results?* and *Do some actors make more use of evaluations than others?*

Regarding the question that *the political factor in the IGR is decisive for the effectiveness of policy evaluation* it is assumed in the theory that the more intra-governmental linkages are present, the more co-governance exists between branches of government and the more effective the evaluation is. If evaluation of public policies is closely linked to coordination, the latter enhances the efficiency and effectiveness of government programmes (Peters, 2006: 133, 135). Therefore, an effective coordination between actors seeking the effectiveness of evaluation enhances the performance of government programmes. The impact of this independent variable on the effectiveness of evaluation is a result of the combination of the above factors as well as the decisions made by different actors inside agencies.

To ensure effective decision-making, the *quality of the bureaucracy* of those actors and stakeholders<sup>39</sup> involved in the process should be taken into account. This is a core element of the State's capacity to adopt policy evaluation as part of an institutional process. Alkin pointed out the experience, skills (know-how) and abilities as decisive for improving public sector performance. The so-called human factor e.g. user or evaluator characteristics<sup>40</sup> (Alkin, 1985) and the educational level to develop effectiveness (Askim, 2007) indicate that the role of the evaluator, evaluation capacity building (Fleischer & Christie, 2009) and the involvement of administrators (Dahler-Larsen, 2000, Ledermann, 2011 and Teirlinck et al., 2013) should become a professionalized area of work across the government and non-government sectors (Sundell, 2014: 31) introducing skilled and experienced cadres to strengthen capacity development.

In essence, the quality of bureaucracy points toward the fact that policy-makers, politicians, officials and stakeholders involved in the evaluation process and operation of programme/policy should be skilled with a basic level of know-how, technical support to fulfil the needs of evaluation and monitoring systems. These capabilities help to perform efficient data-gathering, data-entry, analysis – these will be more sophisticated depending on the amount of data provided by programmes and government institutions – and most importantly to improve decision-making based on evaluations. The more they are involved from the beginning of evaluation, the more likely that the evaluation will impact upon the redesign and expansion of the policy (Teirlinck et al., 2013: 371). For Dahler-Larsen (2000: 85) and Ledermann (2011: 16), the involvement of administrators leads to a better understanding of the evaluation process influencing it.

---

<sup>39</sup> Stakeholders are those policy-makers, politicians, ministers, officials and evaluation practitioners directly responsible of success or drawback in evaluations.

<sup>40</sup> The human factor concerns to characteristics such as knowledge, skills, and experienced cadres, which affect the quality of the bureaucracy necessary to lead changes on programmes.

Similarly, Teirlinck et al. (2013: 368) draw attention to their [bureaucracy] role and how far the lack of involvement, along with the design (methods and approaches) and content (formulation of recommendations and use of evidence base) of evaluation are linked to the failure of evaluations to influence policy. These factors can be broadly categorized into a widely acknowledged concept such as “civil service”. The literature on administrative reforms, new public management and governance agrees that the permanent civil service is central to improving public sector performance and to maximizing the chance of adequate policy implementation. The policymaking process could perform more efficiently if government gives an active role to the civil service because that is where experience can be found (Peters, 1996: 91-107).

Overall the quality of the bureaucracy will have a decisive impact on policy making and the effectiveness of evaluation. On the one hand, public administrations characterised by personal loyalties, patronage networks and a lack of expertise will adversely affect governments’ ability to manage under uncertain contexts. Khan (1988: 322) argues ‘...a lack of training and lack of loyalty to the organisation put independence in reporting at risk’. On the other hand, a merit-based career civil service will ensure the officials’ commitment and experience to improve public sector performance and respond to citizen’s demands. Such a bureaucracy looks for coordination between departments and agencies to get the most from evaluations for the benefit of the programmes. As Peters (2006: 126) argues, the civil service provides a better possibility of adequate policy coordination, while its absence may be an impediment to effective coordination within government.

In this context, it is assumed that the more professional officials working inside the areas operating the policy evaluation the more committed they will be to achieving the effectiveness of evaluation. Moreover, the variable is relevant in all systems but particularly important for those emerging countries, where the reform of the civil service has not been fully implemented. That is, the scale of complexity increases for developing countries due to their traditional bureaucracy, which has less regard for their value of being democracies ‘in transition’ (Lijphart, 1968 cited in Cabrero, 2000: 197 and Linz, 1973).

To that extent, it seems that all factors highlighted by academics are focused on the decisions directly made by the actors involved in the process but utilisation does not depend only upon the abilities of evaluators as Jennings & Hall (2012: 16) refer,

The use of scientific evidence and EBP [evidence-based practice] does not simply depend on its availability, relevance, credibility, and staff capacity. It also depends on the mission and mandates of the agency, its political environment, and its internal characteristics.

There are some other circumstances and/or contextual factors denoted here, which offer alternative explanations of the decision-making process as Hojlund (2014: 38) concludes in his article, “...the literature on evaluation use has been very good at describing the evaluation, its conditioning factors, etc. while neglecting the organisational context in which the evaluation organisation operates...” For instance, the democratic context and autonomy, in which operate those coordinating the policy evaluation, which are not constantly referred. This might due to the characteristics of the context in which use of evaluation has been analysed by many scholars i.e. most of the research has been conducted in democratic countries: Denmark (Hojlund; Dahler-Larsen), Switzerland (Ledermann; Bundi), Australia (Herbert) and the USA (Jennings; Hall) amongst others.

Given that evaluation has been promoted by developed countries such as these earlier mentioned, it is assumed that democracy is an implicit characteristic not usually measured. However, in emerging economies, conditions such as the flawed nature of democracy and the autonomy (control) of public agencies are issues that have not been solved yet. These conditions may offer alternative explanations for the effectiveness of evaluation and the foundation of policy evaluation frameworks. Alkin (1985 cited in Fleischer & Christie, 2009: 171) highlights organisational and political background, while Askim notes the political climate (Askim, 2007: 467), as potentially influential in terms of transparency and accountability systems.

These factors are related to the impact of the *level of democracy* on the effectiveness of evaluation. This variable refers not only to free and fair elections, the right to vote, and freedom of expression but also the capacity to effectively implement and evaluate policies within a system of checks and balances (Dahl, 1971: 3 cited in Lijphart, 2012: 46; EIU democracy index, 2011) such as monitoring system and policy evaluation. In such a democratic context, evaluation becomes a powerful tool to ascertain whether the government's actions are in accordance with those planned, thereby enhancing transparency and accountability.

This is confirmed by Jennings & Hall's (2012) survey of the use of information by US agencies, which indicates that the level of conflict in the political environment affects the degree to which decision makers are interested in utilising evidence "Differences in the degree to which decisions are driven by a variety of political considerations..." Moreover, after reviewing the experience of M&E systems in developed countries such as the USA, UK, Australia, Canada, New Zealand (Andrews, 2008, Bamberger, 1991, Bangura & Larbi, 2006, Cabrero, 2000, Horton & Mackay, 1999, Ospina et al., 2004, Pollitt, 2006, Sundell, 2014 and Wollmann, 2003), this variable seems especially relevant to economies 'in transition', seeking to change their authoritarian and bureaucratic image, in which evaluation represents an indication of modernization.

This takes the discussion to the *autonomy* of coordinators and evaluators of the policy evaluation. This factor embraces both the organisational and political context (Teirlinck et al., 2013: 374), and should provide an external, objective and unbiased way to view the programmes. Scholars define the concept of *autonomy* as the extent to which agency [either the evaluation agency or the public agency coordinating the evaluation] can decide by itself what matters it considers important (Verhoest et al., 2004: 14).

It concerns the degree of discretion and responsibility to take control from government and execute decisions by itself (Rhodes, 1996: 659) insulated from full political supervision by the central head authority (Christensen, 1999). Whereas for Rhodes (1996: 660) to some extent, the degree of autonomy is a characteristic of governance, emphasizing a bureaucratic approach,

Autonomous agencies possess externally differentiated and internally coherent preferences (primarily over policy outcomes) that they are able to achieve either directly, by setting policy, or indirectly, through the political process. An agency's capacity to achieve desired outcomes depends on its organizational resources and on the extent to which other actors can veto, reverse, or punish the agency. Caughey et al. (2009: 17)

Verhoest et al. (2004: 104) carefully provide a distinction between autonomy as: i) the level of decision-making competencies of the agency (concerning management and agency policy); and ii) the exemption of constraints on the actual use of decision-

making competencies of the agency (structural, financial, legal and interventional constraints). Whereas Caughey et al. (2009: 2) also envision two key sub-concepts, the independent goal formation and the capacity to achieve desired outcomes but these distinctions do not necessarily examine other indirect factors influencing decision-making. Moving forward to a more comprehensive concept, the literature agrees that autonomy enjoyed by those responsible for evaluation entails the delegation of competences, power and functions from the hierarchical supervisory body, and enables the responsible agency to operate independently on the basis of its formal/legal status.

Building on these ideas, for instance, the relative independence of the evaluator will affect the objectivity, transparency and accountability of the results obtained, as well as the credibility of the evaluation amongst scholars, citizens and even inside the public sector. To exemplify how autonomy affect effectiveness, Dahler-Larsen (2000: 73) noted that external evaluators are less burdened by internal organisational culture and concerns over sanctions, and are more likely to recommend programme termination than internal evaluators, who prefer to make adjustment in programmes as a result of evaluation. In this context, the current view from studies conducted by Caughey et al. (2009: 18-24) include a set of characteristics concern the autonomy of any public agencies. These refer to matters which involve sources of information, investigative powers, internal veto points in making decisions about agency goals, politicization, and independent leadership, amongst others.<sup>41</sup>

An additional factor to consider is the degree of centralization of the monitoring and evaluation system. On the one hand, centralization attempts to control the implementation of a homogenous policy but, on the other hand, could also jeopardize the flexibility and freedom to adjust policy in the light of particular needs. In this sense, control is the inverse of autonomy (Verhoest et al., 2004: 106), where the government could take control of agency behaviour to influence its decisions. As these authors indicate in their study, the positive effect of autonomy on one dimension can be offset by negative effects on the other, with consequent tensions.

While we have so far focused on how evaluation might be affected by institutional and contextual factors, another influence may be the nature of the evaluation itself. Alkin (1985 cited in Fleischer & Christie, 2009: 171) and Teirlinck et al. (2013) both stress the design of evaluation as a factor shaping effectiveness, whereas Askim (2007: 467) notes that the infrastructure and demand side of performance data operate as barriers to evaluation use. These categories relate to matters of the *policy evaluation framework* adding another variable to the research. Indeed, a review of a variety of evaluation systems from Australia, Colombia, United Kingdom, Japan, Spain, United States, Peru, France (Lopez-Acevedo et al., 2012 and Rist et al., 2011) confirms its importance to produce an effect. This variable is related to the formulation stage of the policy evaluation process and the impact of the framework on the effectiveness of evaluation and classifying the types of factors has been the subject of continuing discussion.

This factor should consider evaluation within a comprehensive system and focuses on the importance to embrace two elements: a coherent design and a feasible system

---

<sup>41</sup> Others are the autonomy to get qualified staff, money for investigation, report quality, organisational identity career stability, legal mandate, financial, human and time resources, case studies of "clash", interest group alliances, reputation, public goods provided by agency, existence of multiple principals, numerical count of policy invalidation, legal factors, competing policies, relationships with interest groups (networks), constituency and salience of issue. See Caughey et al. (2009)



operationalization of evaluation (see methodology chapter). Ideally, *coherence in evaluation design* is needed and should strike a balance between gathering and obtaining relevant information on the one hand and enabling the evaluated to carry out their tasks on the other. At one extreme, the imposition of highly complex systems, accompanied by excessively legalistic obligations for regulation, monitoring and evaluation may impose high burdens and set ambiguous requirements on the evaluation framework and reduce the stakeholders' ability to improve and innovate. At the other extreme, underspecified and weak systems may give too much discretion to actors such as officials, ministers, legislators and especially policy implementers and permit evasion.

In reviewing the debate on the design of policy to date, Ospina et al. (2004: 232) argue that a broad model should be consistent at the 'vertical-horizontal' levels of evaluation as outlined by the Latin American Centre for the Administration of Development (CLAD) and explained in the methodology chapter. In addition, evaluation requires a strong legal framework and participation of interest groups including citizenry, NGOs and target groups as key users, whose actions and behaviour of the demand side could alter government action (Wu et al., 2010: 69) gaining legitimacy and autonomy. Focusing merely on the coherence of design does not capture broader influences such as the importance of infrastructure of the system (Askim, 2007) to incorporate all performance information timeliness (Leviton & Hughes, 1981) and evaluation reports in a data system that should operate as a comprehensive model.

In this sense, the *structure of monitoring and evaluation system* means the high technical level of sophistication needed due to the large amount of data to gather, interpret and use by agencies timeliness. An example of the issues linked to the structure and lack of coherence at the 'vertical-horizontal' levels of evaluation is when several monitoring and evaluation systems operate within organisations but without coordination between them. This is based on the assumption that all data have a purpose in providing results to different actors, which should be feasible to operate and interconnected between them. Thus, a complex scenario could jeopardize the decision-making process because they (actors) are puzzled over a "mountain" of data from different sources, which makes it difficult to process. Hence, the evaluation process could be hindered from providing objective and timely results which could lead to policy change.

The policy evaluation framework also needs methodological rigor, for which many authors consider it as a characteristic of the quality factor (Cousins & Leithwood, 1986 and Leviton & Hughes, 1981). The rigor concerns the order in which evaluations should be executed in order to be coherent with the strategic planning process. For example, in the Mexican case there are many types of evaluations e.g. design, consistency and results, indicator, process, impact, strategic, complementary and specific (see chapter 'Policy evaluation in Mexico'). However, the rigor of the framework suggests that the coordinator of evaluations should not recommend to an agency to develop an evaluation of impact or process before doing an evaluation of design or consistency. Similarly, the Chilean programmes, in which the evaluations of public programmes and of new programmes should be executed before others such as impact, comprehensive spending review and/or ex-ante.

### **Concluding remarks**

The discussion so far has sought to provide a basis for understanding how the most influential factors affect the effectiveness of evaluation. This chapter also reviewed the origins of evaluation and the evolution of the concept until recent decades when it became a multidisciplinary method and acquired visibility. Having reviewed the main literature on evaluation, some conclusions can be drawn on the strengths and weaknesses of the field. A first conclusion is that effectiveness of evaluation is valuable in government intervention, regardless of the approach, evaluation is a key analytical process, drawing upon a variety of disciplines to understand whether government interventions are effective. The chapter also noted how the growth of evaluation was due to the public sector reforms associated with the New Public Management (NPM). Indeed, the NPM brought the *results-based budgeting* to the core of policy evaluation practices in most emerging economies.

At the present time there is a question mark over how these developing countries are dealing with the utilisation of evaluation and closing the policy cycle, especially given that governments have invested a large amount of money to improve the way in which evaluation operates. Thus, this research is trying to go further to the next step in the policy cycle, to analyse the effectiveness of such policy. However, significant research gaps remain in our understanding about what happens inside the process of the effectiveness of evaluation. Indeed, there is a fair level of agreement about the most influential factors, and how nearly all of these conditions explained above are necessary but not necessarily sufficient to achieve effectiveness. In some cases, critics ignore factors whereas others stressed particular issues, which cannot be used as generalizations and perceived as a limitation of such studies.

However, having established the current debates around effectiveness and after looking across into different stances and factors implicit, and the discussion of the overall factors, which undoubtedly have a particular influence on effectiveness, the Intra Governmental Relations, quality of bureaucracy, democracy, autonomy, and policy evaluation framework will be tested in the following chapters to validate generalizations. As was mentioned, the institutional and political factors studied here have been relatively under-explored by scholars and academics. This research looks at these factors in greater depth through a comparative study of the Mexican and the Chilean experiences of policy evaluation to address the core questions of effectiveness.

This research aims to contribute to filling the gap addressing the question *Whether these are the most influential institutional (political) factors shaping the effectiveness of evaluation?* The point of discussion is whether these factors influence the extent to which evaluations bring about policy change. The following chapter attempts to explore in detail these institutional and political factors and summarize the criteria for assessing their impact on effectiveness and presents the research strategy for the comparative case study.

## **CHAPTER 2. Methodological approach**

### **Introduction**

In his article “Where is political science going?” Sartori (2004: 351) argues that politics is the interaction between public organisations and behaviour, and there is a risk of creating a useless science when there is no linkage between theory and practice. In this context, the research strategy adopted here will be of a flexible design, using qualitative methods to determine those linkages. This research attempts to explain whether the policy evaluation is effective and what are the factors affecting such effectiveness. This chapter offers a detailed description of the qualitative method used to answer the research questions. After an overview of the research strategy, the chapter focuses on the dependent and independent variables. The dependent variable at the centre of the research is “the effectiveness of policy evaluation”.

Effectiveness will be considered in terms of the extent to which the findings and recommendations from evaluations are associated with changes to the evaluated policies and programmes. This section elaborates on the possible effects that evaluation may or may not have including amendments to the policy, changes in allocation of resources, cancelation of programmes, maintenance of existing programmes and more general “learning processes” by the programmes’ operators following evaluations. The next section outlines the independent variables, the institutional and political factors that might influence effectiveness. These factors comprise the Intra Governmental Relations, quality of bureaucracy, level of democracy, policy evaluation framework and degree of autonomy.

This dissertation aims to provide an in-depth analysis of the causal relationship between the dependent and independent variables, following the approach offered by Gerring (2008: 650). Using the case study will allow us to explore causal mechanisms and to confirm, disconfirm or reframe a hypothesis. The following section therefore, focuses on the case study as a research approach and explains the advantages and shortcomings of this method, for the analysis of the Chilean and Mexican health sectors. The analysis also outlines some of the characteristics of the health sector in the two countries.

The final section outlines the empirical basis for the case study research and explains how the data was collected. The empirical resources comprise a mixture of documentary analysis, involving close examination of different sources across countries, complemented by elite interviews with officials, politicians and experts in academia and non-governmental organisations (NGOs). These sources underpin the research strategy taken to reach the conclusions of the dissertation.

### **Research strategy or how the research is to be carried out**

To answer the question raised by Pawson & Tilley (1994: 305) “What more do we know at the end of the evaluation about the patterns of outcome effectiveness of the initiative, and what else do we need to know?” it is necessary to investigate and test empirically the relationship between independent variables. In doing so, this dissertation is built upon a bottom-up or inductive approach of reasoning (Bernard, 2013, Burnham et al. 2004, Hantrais & Mangen, 1996 and Pennings et al., 2006) to explore evaluation to establish broader assumptions. The research relies on a qualitative method, specifically a comparative study between the Mexican and the Chilean experiences of policy evaluation.

This study is based on a *small-n* sample of evaluation models e.g. health sector and country settings, which will be contrasted on the basis of economic, legal and political factors. In taking this approach the research draws upon what Pennings et al. (2006: 5) called the crucial triad of *what, when* and *how to compare*.

To examine the trends of decision-making over time as result of evaluation, i.e. the effectiveness of evaluation, it is important to identify whether the programme improvement could be attributed to evaluation. In order to identify this, the analysis is divided into two components. One is the review of programmes of the health sector and whether recommendations of evaluations performed from 2007 to 2014 led to changes classified as amendments, status quo, learning process, cancelation or no data available. The other examines the budgetary process at Congress to determine whether those recommendations of evaluations were used for programmes' changes in the policy sector. It also examines the allocation of financial resources to identify whether they increased, decreased or were unchanged after evaluations took place. In both components is considered the role of political and institutional factors.

These factors are potential influences in terms of policy (recommendations) take-up. Here, the most important factors in the decision-making process for effectiveness to measure are the following factors known as independent variables: IntraGovernmental/Interagency Relations, quality of bureaucracy, level of democracy, policy evaluation framework and degree of autonomy. These will be tested against each other across countries based on the assumption that some variables may differ in some way to explain differences in the effectiveness and to explain which one should be expected to be the most important and why this prevails over the others.

Based on Pawson & Tilley's (1994: 305) list of enquiries to address the problem of evaluation research, this thesis investigates the following questions as the most important interrelated to the object of study:

*What explains the differences in the effectiveness of policy evaluation across these countries?*

*What are the most influential institutional (political) factors shaping the effectiveness of evaluation?*

*Whether policy evaluation matters for the public policy and budgetary process?*

To investigate these questions, the research adopts a historical analytical approach of the evolution of the evaluation process and country settings in Mexico and Chile. Moreover, to shed light on the practices between branches of government, the analysis also uses primary data from interviews with politicians and officials. Through comparison of evaluations of both countries, the research will attempt to address the central problem investigated here *whether policy evaluation is effective across these countries*.

### *1. Dependent variable and operationalisation of evaluation effectiveness*

According to Howlett, Ramesh & Perl (2009: 186), the task of evaluation is to find out if the programme is doing what it is supposed to be doing or whether it needs to be adjusted, based on its intended goals. With this in mind, the dependent variable is *the effectiveness of policy evaluation in the Mexican and Chilean public administration*.

The operationalisation of this variable captures whether and how policy changes in response to evaluation. For Dahler-Larsen (2005: 623), the "utilization dimension" is

based on the premise that although they are not always used, evaluations are produced to play a role in future “practical action situations”. Effectiveness is the essence of evaluation results both in public and private sectors (Hintze, 2001: 14). For the purpose of this research, this effectiveness is defined in terms of the impact of evaluation on policy itself, specifically the extent to which evaluation of a policy/programme leads to changes in that policy or programme; in other words, how many of the recommendations noted in evaluation reports are adopted by programme operators in response to those evaluations.

To examine the trend in decision-making over time, this research looks at those changes in programmes, which are made as a result of evaluations. Looking for such changes this research will use the formal evaluation and monitoring systems issued by government in both countries.

In Mexico, the research uses the database developed by the Ministry of Finance (SHCP) called “Presupuesto basado en Resultados and Sistema de Evaluación del Desempeño” (Results-based Budgeting and Performance Evaluation System ‘PbR/SED’) and by the National Evaluation Council of the Social Development Policy (CONEVAL),<sup>42</sup> comprising the recommendations of evaluations issued by external evaluators. In Chile, the so-called “Sistema de Evaluación y Control de Gestión” (System of Evaluation and Management Control ‘SECG’) is the database developed by the Ministry of Finance (MoF) to track the recommendations.<sup>43</sup> In both cases, the research examines only the evaluations of the health sector as a sample to infer or generalize findings to other sectors (later on explained).

In both countries, as a result of evaluation reports, programme operators then draw up the working report and the so-called “agency statement”. The former is the response by such operators to recommendations listed in the evaluations, indicating actions to be taken and the timeframe for implementation. The latter is the report establishing the position to be adopted (agreement or disagreement) by programme operators on behalf of the programme with regard to recommendations and evaluation. These reports can highlight problems faced by operators in addressing recommendations or indicate when these cannot be implemented. On that basis, the analysis in this dissertation uses these reports as the principal resource for operationalising the concept of evaluation effectiveness since they provide evidence of the extent to which agents have responded to the recommendations of evaluators.

Concerning the speed of adopting changes, it depends on [institutional] agencies’ capacity and their commitment to addressing the recommendations, which could complicate the process of judging effectiveness; for example, when implementation takes longer than the timeframe established in reports. However, according to regulations regarding evaluation and implementation of recommendations in both countries, there is no criteria and limited timeframe of adopting changes (recommendations) neither sanctions for delays. Indeed, it is considered a participatory, gradual and progressive process, in which, decision-makers and programme operators indicate the deadline for implementation.

---

<sup>42</sup> The “Mecanismo para el seguimiento a los aspectos susceptibles de mejora derivados de informes y evaluaciones a los programas presupuestarios de la Administración Pública Federal” (Mechanism to follow-up evaluation results) is the guideline to operate the database developed by government in which public agencies upload their evaluation results. This database is available at: <http://www.transparenciapresupuestaria.gob.mx/> and <http://www.coneval.org.mx/Paginas/principal.aspx>

<sup>43</sup> The database is available at: <http://www.dipres.gob.cl/594/w3-channel.html>

To address such cases, where programme operators are working on recommendations by the deadline, changes are considered within the category “in progress”, which means a positive effect regarding the outcome of effectiveness. Conversely, when programme operators do not indicate current situation of adopting changes by the deadline, these are classified as *status quo*.

It is also important to note that identifying the precise drivers of any adopted changes is difficult if these are not specified explicitly in reports. It could be that decision-makers are improving their programmes, although, when the reasons for the actions to be taken are not explicitly expressed in such reports and data is not publicly available, it is difficult to clarify the motivations for such changes within the agencies. Nevertheless, it is reasonable to conclude that changes made in line with recommendations are at least in part a response to those recommendations. Moreover, every aspect (recommendation) identified by the external evaluator during the assessment of the programme or policy should be listed in the working report, in order to avoid that evaluation might lead to an outcome that cannot be observed.

However, when the analysis of each programme evaluated leads to an outcome not observed or expressed in the evaluation and working reports, this is noted in the case study chapters of both countries with no effect regarding the outcome of effectiveness. This is because the main purpose of this analysis is to prove that recommendations noted in evaluation reports are adopted by programme operators in response to those evaluations, this is the effectiveness of policy evaluation.

Both reports (working and agencies' statements) are open to the public in the SHCP, CONEVAL and DIPRES websites, which are regularly updated. However, some of the problems related to operationalising “effectiveness” were those reports, which were not easily accessible. Indeed, these documents were uploaded in pdf format in a data repository not in a database system as specified in the regulations. One of the consequences resulting from this pdf format is that these documents are easily interchangeable with one another with the risk that recommendations from those reported by evaluators and deadlines previously committed could be changed jeopardizing the following-up process (see case study chapters for relevant findings on this). Thus, the review of recommendations took longer than was expected in order to verify congruence between recommendations observed in evaluations and actions committed in working reports and agencies' statements.

Another challenge in operationalising the concept of effectiveness is related to acceptance of evaluation results by programme operators who could either accept or not such results due to many reasons usually explained in the agency statement. For example, a recommendation which requires additional financial resources will be difficult to implement and will depend upon the action of external actors such as Congress people, MoF or local authorities. This situation has to be considered by coordinators of policy evaluation but without any sanctions for non utilisation of recommendations. In dealing with such cases, the research considers the response made by programme operators in such agencies' statements to classify each recommendation according to the criteria for judging the effectiveness explained further below.

It should be stressed that it is not the purpose of this study to judge whether these recommendations are the right changes needed to the programme or any variation

in the precision with which recommendations are made. The rightness of these changes (recommendations) has already been determined by the external evaluator based on their judgement and responsibility to provide an impartial and objective assessment of the programme. However, where some data is different or missing e.g. actions taken are different from those recommended by the evaluator, this is assessed according to the criteria for judging the effectiveness of evaluation as explained below and classified in the *status quo* category. Moreover, this is pointed out in every programme reviewed in the case study chapters and table 3 and 6 annexed.

Returning to the criteria for judging the effectiveness of evaluation, this refers to the relationship between evaluations and the response of programme operators to recommendations. In the review of the literature, the concepts regarding use or utilisation identified by many scholars refer to different categories such as conceptual, process or persuasive, amongst others. However, in practice, these are difficult to measure to prove utilisation and the reason to employ the concept of effectiveness as re-conceptualisation of the theory of use or utilisation.

To operationalise this concept of effectiveness, it is relevant to define all possible outcomes once evaluation took place to determine whether this (evaluation) is effective or not. Thus, on the basis of the empirical research, it was found that effectiveness of evaluation could take various and simpler forms such as *amendments*, *status quo*, *allocation of resources*, *cancellation* or *learning process* in the programmes or policies, all discussed in detail in the review of the literature. This approach is the most effective and feasible way of determining whether an evaluation has accomplished its objective and whether decision-makers are utilising recommendations to lead changes on programmes. However, it is worth highlighting some relevant points regarding the effects.

Concerning the category of *status quo*, to classify a programme evaluated in this category is important to keep in mind that such classification occurs when the implemented recommendations of the evaluation represent less than 30% of the total improved. In other words, the policy makers have only adopted 30% or less from the total of recommendations suggested by external evaluators in the evaluation's report. The main reason to establish this percentage is based on the low rate of accomplishment of recommendations. Therefore, on the assumption that most implementers are contributing to the broader improvement of their programmes in some way, the study sets 30% as the threshold for classifying whether significant amendments are or are not made. The tables 3 and 6 in the annexes illustrate the low rate of recommendations achieved per evaluation and programmes.

Another reason for policies remaining unchanged is when those responsible of the programme(s) do not consider any recommendations for further improvements or changes or when some data is missing and actions taken are different from those observed by external evaluators. This is, despite the recommendation was issued by the external evaluator, policy makers do not mention any action taken to solve the problem or they make different amendments from those recommended by the evaluators.

Regarding the programmes of the category of *learning process*, despite the importance to acknowledge them, it is not possible to provide evidence in the later chapters to assume that some programmes or policies are going through an internal "learning process". That is, it will not address any learning and knowledge processes

from evaluations due to the lack of information available to track either any recommendation or any action taken by policy-makers, which should be properly acknowledged to this category as analysed in the case study chapters. As mentioned in the review of the literature, another category refers to termination of programmes, when data from evaluations highlight risks or the need for improvements in such programmes.

Another source of use comes during the budgetary process, when the executive (MoF) and Congress people use evaluations in order to reallocate finances. Finally, after the revision of programmes, it is also included the category “in progress”, which refers to these programmes in which operators are still working in the recommendations suggested. At the time this study was completed, if programme operators were still working on recommendations by the deadline as noted in the working reports, these were classified as “in progress” with no negative effect regarding the outcome of effectiveness.

## *2. Independent variables*

A number of institutional and political factors were examined in the review of the literature including the institutional design of policy evaluation to establish whether they influence the effectiveness of evaluation. After the analysis of multiple factors it was determined that the independent variables to be tested comprised the following factors: IntraGovernmental/Interagency Relations, the quality of bureaucracy, the level of democracy, the policy evaluation framework and the degree of evaluator autonomy. This section complements the analysis of variables made in the review of the literature, providing a descriptive account of their characteristics and explaining how they are to be operationalized and measured.

As noted below, for each variable there is a database from different sources, which complicate the task to code them into a single coding system due to different scales of measurement. For instance, regarding the variable of quality of the bureaucracy, some sources were identified such as the Governance Indicators Database (DataGov) by the Inter-American Development Bank or the Worldwide Governance Indicators (WGI) by the World Bank both measuring either the degree of professionalism or the perception of the quality of service with some differences on the scale utilised. However, in the end, the decision to use the latter source for measuring this variable instead the former would not have produced different results in the analysis.

With respect to the level of democracy of these developing countries, the following sources were considered: the democracy index by the Economist Intelligence Unit and the democracy ranking of the quality of democracy developed by the Democracy Ranking Association. However, the former index is the best measure for this variable because this provides a clearer scenario of four categories of democratic system.

The same pattern was followed concerning the degree of autonomy, for which, the dimensions identified by Verhoest et al. (2004) capture the different levels based on a distinction between autonomy as: i) the level of decision-making competencies of the agency (concerning management and agency policy); and ii) the exemption of constraints on the actual use of decision-making competencies of the agency (structural, financial, legal and interventional constraints). Thus, this measure was the best choice for this variable.



Given this fact, when data come from various sources, which essentially are different measures, it was difficult to design an elaborate scale and therefore basic binary scale was devised. Their measurement is based on binary selection of every variable, where the variable is coded “0” if the analysis indicated that the condition is absent or it is coded “1” if the outcome is present or high. (Annex 1 illustrates the dichotomization of each variable according to the outcome analysed). However, any nuances that are lost through binary coding are revealed and discussed more fully in case study chapters.

#### Intra Governmental and Interagency Relations ‘IGR’

Although this concept was identified in the first half of the last century by many scholars (Anderson, 1960, Bromage, 1943 and Wright, 1974), it remains important to explain the interaction between governments, branches and actors operating policy evaluation. While much of the literature refers to intergovernmental relations in terms of the “vertical” dynamics between the federal, state and local tiers of government, the focus of this analysis is on the “horizontal” dynamics between and within government branches and agencies at the federal level because the focus of this study is on the federal level of both countries, where most evaluations of the health sector have been developed.

Moreover, in Wright’s overlapping model of IGR, the flow of authority of governmental bodies among national, state and local units (Agranoff & Radin, 2015: 3) provides the conceptual basis for moving from the vertical to horizontal dynamics studied here. We refer to these horizontal dynamics as *Intra Governmental Relations* and *Interagency Relations* and considered as one category, hereinafter IGR. Indeed, the latter concepts are those examined along this research due to this is conducted at the federal level leaving aside vertical dynamics noted by InterGovernmental Relations. The focus is also a function of practical limitations to the conduct of the research, explained later.

Thus, this dissertation is exclusively focused on the horizontal dimension, capturing the relationship between different branches of governments and the relationship between different agencies involved in policy making and evaluation. Regarding inclusion of the legislative branch as part of the government, as noted by Wright (1974: 3), the role of members of Congress and public officials such as governors, state legislators, appointed administrators, etc. is a characteristic of IGR. Indeed, scholars commonly refer to the three branches of government – executive, legislature and judiciary –.

Although, the role of Congress might be an object of major criticism, for instance, when there is a minority government, the governing parties do not control the legislature and might not be considered as part of government or when a government has a strong majority in the legislature in the so-called dominant party systems, such as Mexico. In the end, the legislature is composed from different parties – including opposition – operating within government structure regardless minority or majority government in the legislature. Therefore, it is appropriate to treat it as part of government.

Continuing with this concept, the IGR are those patterns of interaction among actors (officials, ministers, legislators, policy makers and politicians), agencies (ministries or ministries-state enterprises) and branches (executive-legislative-judicial) of government. When these operate between ministries, state enterprises and/or

agencies inside the executive or inside other branch (judicial or legislative) are called Interagency Relationships. Whereas these that operate between the executive (ministries/agencies), legislative and judicial branches are named Intra Governmental Relations.

The importance of the IGR for the effectiveness of evaluation depends on the extent to which those evaluations results are acted upon and lead to changes in policies operated by actors within these branches. Incidentally, it is important to bear in mind that judicial and legislative branches are excluded from the evaluation process (e.g. object under evaluation) in both countries. Indeed, the legislative only operates as decision-maker regarding budgetary process and policy sector. However, later on it is explained the importance of these branches in a comprehensive view of policy evaluation.

To understand the role of many actors in the effectiveness of evaluation, it is important to point out to different dimensions of political relationships within the variable IGR. After the analysis of these dimensions at the federal (horizontal) level,<sup>44</sup> the following relationships are identified. In the first place, there is a pattern of interaction between those operators (bureaucracy) of programmes and policies and coordinators of policy evaluation (DIPRES in Chile and SHCP/CONEVAL in Mexico) mostly located in the MoF. The former receives evaluations and implement those recommendations suggested by external evaluators while the latter is responsible for the coordination of policy evaluation but most important for the planning of the federal Budget.

In the second place, there is a pattern of coordination amongst bureaucracy or programme operators of different agencies to solve problems related to recommendations of evaluations, for which two or more programmes are interlinked in some way. Lastly, it was identified the role of the president through the MoF and Congress. The former develops the budget through the MoF while the Congress authorizes it. Regarding evaluation, the MoF coordinates the policy evaluation while Congress oversees its effectiveness and use evaluations to lead changes on policy sector. These different dimensions of political relationship within this variable will be contrasted in the comparative chapter.

Another matter is the lack of an index or indicator to determine the interrelations amongst agencies or public organisations. However, in order to have an inclusive variable with the most important characteristics describing these relations, there will be integrated different approaches by scholars. Indeed, the IGR settings vary in many ways and need to incorporate factors such as *forms of coordination, information and economic resources*, explained below to be more cooperative and coordinated to make effective evaluation. Franz (1985: 485) affirms that treasure (finances), along with legal authority, information and organisations are closely linked due to the high degree of complexity in these relationships. Thus, the variable (IGR) is based on characteristics highlighted by many authors to determine its impact on the effectiveness of policy evaluation.

The first element is the “forms of coordination” which will be measured in terms of how of those mechanisms operate to improve relations between actors inside the government to empower policy evaluation. These forms are divided into:

---

<sup>44</sup>The focus of this study is on the federal level of both countries.

- i) Negotiation, common agreement/interest or partnership (performance), where objectives are worthwhile for the organisation's purposes. This could be on an *ad hoc* or permanent basis and refers to an ideally strong form of coordination.
- ii) Opportunistic incentives (also called exchange), which produce conflicts sub-optimal strategies as Kaufmann et al. (1985: 801) state. These occur where an asymmetry of roles prevails between principals and agents,<sup>45</sup> superiors and subordinates, the rulers and the ruled, and the few can easily collude to exploit the many. This represents a weak form of coordination.
- iii) Legal authority or obligation to enforce cooperation between those involved in the policy evaluation process; particularly relevant in countries where federalism is fairly achieved and power is not centralised, primarily by hierarchical coordination executed at the central level.<sup>46</sup> This third mode is classified as the minimum form of coordination.

The second component of this variable is the "information". This essentially refers to whether a wider performance information database under the control of policy implementers is used for feedback and effective policy-making. The availability of data to key stakeholders involved in evaluation is essential to provide the necessary information for the continuity of the process, but also for accountability, informing citizens of the government's results. There is also a need to provide data to the budgetary cycle and policy process in a timely manner to all users.

The reliability and quality of performance data is also decisive for better decision-making across the branches of government. In many countries the veracity of data in evaluation is taken for granted due to the use of external evaluators independent from the government, reducing the risk of bias and subjectivity. However, sometimes a source of power and political control emerge from the central authority when it has access to crucial information and has a monopoly of evaluation in some way, which jeopardizes such veracity. Thus, the aim is getting equilibrium between misutilisation and availability of data by operators given the control they have of the monitoring and evaluation system.

The last element of this variable is related to the "economic resources". This element produces a classification between even or uneven distribution of public resources and has an impact on the operation of policy evaluation. For Sharpe (1985: 166), resources are important for the central authority due to the control they can exert over other branches of government. Whereas for Peters (2006: 128) 'The one major control instrument that remains in place is the budget, so that ministries of finance in particular become crucial in the process of controlling administration'. Apart from these references, finances also determine the capacity of implementers to boost policy evaluation and, regarding the strategy of results-based budgeting, the way incentives are applied via reduction, elimination or funding of programmes.

Thus, the state of the art of these relationships is one in which policy and decision-makers, users and actors involved (stakeholders) are willing to use and exchange

---

<sup>45</sup> In the agency theory, there is an agency relationship between two or more parties, in which the principal delegates work to the agent, who is the performer. The aim is to resolve two problems that arise when the principal cannot verify whether the agent is doing or behaved appropriately, and when both actors have different attitudes toward risk and make different actions, which may conflict them. (See Eisenhardt, 1989, Mitnick, 1975 and Ross, 1973)

<sup>46</sup> Regarding this is characterized by the strong influence of the president and centralization of government. These relationships are often defined by the structure set by legal authority (Franz, 1985: 483) and could be a source of effectiveness or weakness. This premise is based on the definition offered by Ostrom & Ostrom (1965: 138) of the system of government.

data (evaluations) for further decisions. It also endorses the idea that the availability, timing, reliability and quality of data, as well as economic resources represent a key input to make the evaluation effective. Indeed, it might not be wrong to argue that the possibility of success in expanding the use of evidence-based policy will depend on the ability of political actors to move away from the *status quo* to embrace a more results-focused approach based on the effectiveness of evaluation into the strategic planning process. For this, these elements (forms of coordination, information and economic resources) will be measured according to binary selection to point out whether the element is present or not to prove the premises stated here as well as to address how these are incorporated in the policy process and their impact on the IGR to lead changes in policies and programmes.

The criterion to assign a code depends upon the outcome is present or not in every indicator and subindicator mentioned above (see annex 1 for a description of each variable and indicators). For instance, regarding forms of coordination the variable is coded “0” when the analysis of these relationships results in minimum or weak coordination, whereas when the outcome shows there is a strong coordination, it is coded “1”. The second element of information measures the availability of data, timing and reliability and quality, when these are present a code “1” is assigned and when these are absent a code “0” is given. Lastly, even distribution of public resources ensures a code “1”, whereas a code “0” is assigned when there is uneven distribution. To complement the analysis questions such as perception of the relationships between actors as well as availability, reliability and quality of data are included.

#### Quality of the bureaucracy

As noted in the review of the literature the *quality of bureaucracy* is a key variable of the State’s capacity to improving public sector performance, in which policy evaluation maximizes the government efforts to perform efficiently. Based on the assumption that *the more meritocratic the bureaucracy, the more effective the policy evaluation*, units responsible of policy evaluation should be staffed by skilled and experienced members to fulfil the needs of evaluation. In this context, the research attempts to measure the variable in terms of the characteristics of the ‘Permanent civil service’.

Lonti and Woods (2008: 27) define the Senior Civil Service as ‘a separate, structured and recognized system of personnel for the higher, non-political positions in government, in order to provide stability and professionalism at senior levels’ and include features such as competitive appointment based on merit, requirement for specific skills and being centrally managed.

Taking into account complexities observed in the review of the literature including developing a specific indicator to measure this variable, the Worldwide Governance Indicators (WGI) source developed by the World Bank will be used. The WGI<sup>47</sup> reports on six broad dimensions of governance (voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption) of 215 countries over the period 1996-2015. However, for the purpose of this research only the dimension of ‘government effectiveness’ will be considered, which includes among others perceptions of the quality of the civil service

---

<sup>47</sup> See Worldwide Governance Indicators <http://info.worldbank.org/governance/wgi/index.aspx#home>

(Kaufmann, Kraay & Mastruzzi, 2010: 4). These results are approximate due to this indicator also measure other elements.<sup>48</sup>

Moreover, the review is from 2007 to 2014 according to evaluations examined. Its score ranges from 0 for the lower degree of government effectiveness to 100 as the higher score indicating the opposite. For this, the variable is coded “1” if results of the country are above 50 in the scale from 0 to 100, whereas the variable is coded “0” when results are under 50. It is important to bear in mind that quality of bureaucracy should be implicit to all actors involved in the evaluation process. However, the quality investigated here is referred to those coordinating policy evaluation including Congress people and those being evaluated excluding external evaluators.

For this, data from the interviewees of both countries such as background, skills and experience will be included (see questionnaire in annex 2) to determine the relationship between the bureaucracy and their capabilities to respond to evaluation recommendations and the consequences of that relationship for effectiveness.

#### Level of democracy

There is a reciprocal relationship between evaluation and democracy where the former operates as a tool helping to boost transparency and accountability. According to Dahl (1971: 3 cited in Lijphart, 2012: 46), democracy is measured by the right to vote, the right to be elected, the right of political leaders to compete for support and votes, elections that are free and fair, freedom of association, freedom of expression, alternative sources of information and institutions which depend on votes and other expressions of preference for making public policies. In democratic countries, people use their vote to obtain public value from the government.

The public value of the government is the equal satisfaction of the human needs e.g. poverty reduction, improvement in education or health (Hintze, 2001: 40; 2003: 3) as well as providing the conditions for economic development, amongst others. In this sense, policy evaluation operates as a method used by governments to assess the creation of public value. In doing so, the Economist Intelligence Unit has developed the *democracy index*, which is based on the UNDP's report and Latin American Barometer, and on the review of 167 countries divided into the five categories: electoral process, pluralism civil liberties, functioning of government, political participation and political culture,<sup>49</sup> to determine how democratic the State is.

Based on this index, democracy ranges into the four categories of democratic system from authoritarian regimes to fully democratic countries.

- 1) In *authoritarian regimes* political pluralism is absent, many of them have dictators and media are state-owned or controlled, repression and no independent judiciary.
- 2) An *hybrid democracy* is due to irregularities on elections, pressure on opposition parties by government and serious weaknesses in functioning of government, political culture and participation.
- 3) *Flawed democracies* are characterized by a low level of democracy and also show various problems notably in terms of governance, political culture and participation.

---

<sup>48</sup> It also captures perceptions of the quality of public services, the quality of policy formulation and implementation, the credibility of the government's commitment to such policies, and the quality of the civil service and the degree of its independence from political pressures already mentioned (see WGI methodology).

<sup>49</sup> The Economist Intelligence Unit. Democracy index 2014.

- 4) *Full democracies* demonstrate good governance in terms of more capacity to implement policies effectively within a system of checks and balances.

This source will be used and to measure it, the indicator is coded “1” when the democracy index classifies either Chile or Mexico into the category of full democracy. Thus, the other categories will be coded “0”. The latter source and to demonstrate whether *the more developed the democracy, the more goals of policy evaluation can be achieved*, the historical evaluation context is referred to understand whether the democratic context shapes the policy evaluation. This offers a different perspective on whether policy evaluation is imposed upon or engages with stakeholders within these particular country settings.

#### Degree of autonomy

As mentioned in the review of the literature, the *level of autonomy* is the extent to which the agency [either the evaluation agency or the public agency coordinating the evaluation] can decide by itself<sup>50</sup> what matters considers important (Verhoest et al., 2004: 14) and exempted from full political supervision by the central head authority (Christensen, 1999). Based on the statement *the more autonomous those who are responsible for policy evaluation, the more credible and effective the evaluation*, an implicit assumption is that the effectiveness of evaluation will depend upon the extent the agency that coordinates evaluation would be operationally independent from the central authority.

The degree of autonomy will be measured following the dimensions offered by Verhoest et al., e.g. managerial, policy, structural, financial, legal and interventional autonomy, mainly due to these dimensions embrace most of the definitions reviewed in the literature. However, this study is primarily focused on the three aspects (policy, financial and legal) which are considered the most important. An analysis of them can shed light on the autonomy of evaluation agencies across countries, and will be used to determine how effective the autonomy granted to an agency is towards pursuing policy evaluation goals.

These dimensions will be used to rank agencies into one of the two levels of agency's autonomy as follows:

- 1) Policy autonomy: low and high.
- 2) Financial autonomy: low and high.
- 3) Legal autonomy: low and high.

*Policy autonomy* is the degree to which agency may or may not decide on individual applications of general rules or to issue policy instruments in general. *Financial autonomy* establishes the level of primarily financial resources, fully funded by central government, or partially funded by other governments, product sale or loans, or fully covered by agency including deficits. As can be noted from the description of the element of finances in the IGR factor, both are related on the extent to which these determine the capacity of implementers to boost policy evaluation and its effectiveness. Finally, *legal autonomy* covers status aspects such as if the agency is part of the central government, or has a separate one and does not own legal personality, or has a legal personality and is created by a parliamentary act or under private law.

---

<sup>50</sup> Due to complexity for measuring leadership as an indicator, autonomy will be considered as a proxy of leadership for this case study.

For this, a code “1” is assigned when the level of policy, financial and legal autonomy is high, whereas a code “0” is when the level of autonomy is low. To complement the analysis questions such as perception of level of autonomy of public agencies and actors are included. Moreover, while autonomy may ultimately be desirable at any level and at any agency, the autonomy investigated here is focused on those coordinating policy evaluation including Congress and external evaluators to prevent them from playing the roles of judge and jury.

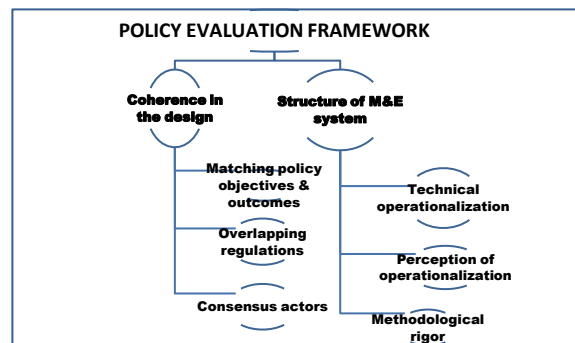
### Policy evaluation framework

Based upon the broad definition of this concept in the review of the literature to analyse the institutional design of policy evaluation, the variable is divided in two sub-variables (see figure 2 below). The first sub-variable is the *coherence in the design*, an essential instrument for subsequent implementation. This includes matching policy evaluation objectives and outcomes based on the theory of cause and effect,<sup>51</sup> i.e. that these problems and targets identified, as well as expected outputs and outcomes are addressed by the policy. For this, the design of every policy, programme and project of government should be correlated between them at the vertical and horizontal levels (macro, meso and micro)<sup>52</sup> as suggested by the CLAD (Ospina et al., 2004) in order to have a coherent policy design.

Other subjects to measure are whether this could be overlapped with similar regulations and whether an inclusive policy should attempts to reach a consensus between stakeholders. It means, the extent to which policy coordinators hold hearings to allow interest groups including citizenry, NGOs and target groups to present arguments, counterarguments and evidence regarding the proposed policy evaluation and its regulation in order to find consensus between them.

Figure 2

### Elements of the policy evaluation framework



Source: author's own elaboration.

The *structure of monitoring and evaluation system* is the second sub-variable entailing a technical mechanism to attain policy evaluation goals. The architecture of the system needs a high technical level of sophistication due to the large amount of data to gather, interpret and use by the government agencies. It also needs

<sup>51</sup> A causal theory is about what causes the problem and what intervention would alleviate that problem (Birkland, 2011: 241).

<sup>52</sup> The micro level refers to individuals, meso level to programmes and agencies, and the macro level or meta evaluation focus on sector, regional and national policies in the medium and/or long term goals. The macro level or meta evaluation identifies the impact of public policies on sector, regional and national objectives. For instance, in Mexico, those established in the National Development Plan (PND), whereas in Chile, those established in the Government Programme. Regarding the long term goals those established in the Sustainable Development Goals <http://www.un.org/sustainabledevelopment/>

methodological rigor to harmonize evaluations to specific needs of programmes or policies. The rigor concerns the order in which evaluations should be executed in order to be coherent to the strategic planning process and those evaluations performed. These issues along with the results of interviewees' perception will be analysed to find out the complexity (difficult or easy) to operate the system by their users. The results clearly provide arguments attributable to the evaluation system's operationalization to affirm that *the more ambiguous and complex the framework of policy evaluation, the less effective the evaluation*.

The criterion to assign a code "1" is when policy evaluation objectives and outcomes match between them or regulations are overlapped or policy coordinators hold hearings with interest groups, when these are absent a code "0" is given. Regarding the structure of monitoring and evaluation system, a code "1" is assigned when technical operationalization is easy and "0" when operationalization is difficult. The perception of interviewees will determine whether operationalization is easy "1" or difficult "0". Lastly, the subindicator of methodological rigor indicates whether evaluations are executed coherently to assign a code "1" or not "0". See annex 1 for a description of each variable and indicators along with the countries used for comparison.

### **The comparative dimension of policy evaluation**

A central task of the research is to study the effectiveness of policy evaluation from a comparative perspective.<sup>53</sup> This technique emerged as part of the post-war development of public administration as a discipline, which included an attempt to acknowledge best practices and benchmarks from other countries, and to draw general lessons from particular cases (George & Bennett, 2005 and Heady, 2000).<sup>54</sup> This practice also known as cross-national, cross-country, cross-systemic or cross-institutional study (Hantrais & Mangen, 1996: 1)<sup>55</sup> offers a different view of two or more objects of study for better comprehension of the diversity of political systems based on categories, characteristics, and interactions between political organisations and society (Einsestadt, 1962).

Moreover, to respond to the core question of *whether policy evaluation is effective across these countries*, the cross-case study contributes to analysing the problem within a broad context. Some advantages of the comparative method for political scientists are mentioned by Hague & Harrop (2004: 69-71), including learning about other governments and the potential of prediction.

To investigate the dependent variable *the effectiveness of policy evaluation in the Mexican and Chilean public administration* mentioned at the beginning of this chapter, the research design is based on Mill's method of agreement and indirect method of difference. According to Ragin (2014: 36-8) the latter method attempts to approximate experimental design with non experimental data, while the former proceeds by the elimination of explanatory variables. This author noted that to use the indirect method of difference it is necessary to identify negative cases to reinforce

<sup>53</sup>For Pennings et al. (2006: 3) the art of comparing is: "One of the most important cornerstones of the development of knowledge about society and politics and insights into what is going on, how things develop and, more often than not, the formulation of statements about why this is the case and what it may mean to all of us".

<sup>54</sup>For better understanding of the comparative studies in the public administration see Heady (2000).

<sup>55</sup>These authors add that this method also aims to point to possible directions of being followed, to sharpen the focus of analysis by suggesting new perspectives, identification of gaps in knowledge by suggesting useful avenues for future research, offer possible explanations in terms of national similarity and difference (Hantrais & Mangen, 1996: 3-5)



conclusions drawn from positive cases (Ragin, 2014: 41), e.g. countries in which there is an assumption that policy evaluation is not effective.

It also involves searching for differences between variables that display an outcome and those where is absent. These differences among positive cases can address patterns of multiple causation (Ragin, 2014: 44). Based upon this method of difference, the main features to be considered in this comparative study are the following.

In the first place, the decision to compare Chile and Mexico is related to the main advantage claimed by these authors, to contextualize and acquire knowledge in more formal political analysis, considering overseas events, which could help to interpret new developments and practical political relationships (Hague & Harrop, 2004: 69-71).<sup>56</sup> Although the countries have some differences in terms of the civil service, corruption, democracy and government effectiveness (see case study chapters), they also share some similarities regarding the political system (presidentialism) and administrative reforms e.g. the programme of administrative modernization, Managing for Results and Results based Budgeting. Indeed, various policies adopted by the Mexican government come from the Chilean context.

Thus, these countries are worth studying in parallel to introduce a discussion about how institutional factors in Mexico differ (or not) from countries considered as more 'mature' such as Chile in terms of their monitoring and evaluation system. In terms of Ragin's approach, this would be a negative case compared to the positive. The premise is that these systems have different reputations for effective policy evaluation, so, to the extent that the institutional factors can help to explain different outcomes and identify gaps, they will also be able to explain their effectiveness.

The experience of Chile will be included due to its importance as an emerging economy pioneering the development of a monitoring and evaluating system<sup>57</sup> (World Bank, 2005). The Chilean National Budget Office's introduction of performance indicators in the budget began in 1993 and was maintained with some ups and downs until 2001, when performance indicators were reintegrated into the budgetary process (Ministry of Finance, 2003). Since then the country has developed an ambitious evaluation agenda comprising performance indicators, evaluation of institutional programmes and comprehensive annual reviews (World Bank, 2005b: 11).

The Chilean system is one of the strongest and most recognized monitoring and evaluating systems in the world; as Mackay (World Bank, 2005b: 18) states: 'the single most important factor is probably the location of the M&E system... in the finance ministry...' These characteristics might explain its strength, power and capability to implement highly cost-effective policy evaluation and it is expected somewhat better and effective Chilean policy evaluation compared to the Mexican.

---

<sup>56</sup> Other advantages are the following: it has the potential to improve the classifications of politics; it enables hypothesis testing for explaining the particular "So explaining the particular calls forth the general; only theories explain cases"; and to make predictions about politics once generalizations are validated.

<sup>57</sup> Indeed, "Chile's system is one of the strongest government M&E systems in the world. In achieving this success, the single most important success factor is probably the location of the M&E system where the system was designed, progressively developed, and is still managed in the Ministerio de Hacienda, specifically its Budget Directorate (DIPRES). The Hacienda has succeeded in creating an 'evaluation factory,' which includes a well-developed process for planning, commissioning, managing, reporting, and using a range of types of evaluation". (Keith Mackay, discussion at World Bank)

Therefore, studying asymmetries or similarities in the results of this comparative study could explain how these factors influence effectiveness, in terms of the success or failure of policy evaluation. Moreover, the differences between these cases help to identify whether these factors are sufficient conditions for the outcome as Mill (1843: 455) suggests,

If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former, the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon.

However, following the analysis presented in the case study chapters and comparing institutional factors, it appeared that the Chilean policy evaluation has not been more effective than the Mexican one; neither of the outcomes were constant over time. Therefore, the Mill's method of difference was not sufficient to explain effectiveness across and within these countries in the course of time. Considering this, the method of agreement could be used to analyse the evidence, according to Mill (1843: 456), this method "stands on the ground that whatever can be eliminated, is not connected with the phenomenon by any law."

However, the latter method neither was sufficient to explain differences. As Ragin (2014: 43) noted "The method of agreement may show no common cause or set of causes". At the end, the Mill's method of agreement and difference was used jointly to find factors in common amongst all cases where utilisation of evaluations appears or when the factor is absent and effectiveness does not occurs.

In the second place, the decision to select cases on the basis of a *small-n* comparative analysis essentially depends upon the availability of countries with similar characteristics for systematic and detailed analysis, and the reason for not being randomly selected. This is what Hantrais & Mangen (1996: 3-5) refer to as the three major problems in the management of research facing scholars: the availability of and access to comparable datasets, definition of research parameters and associated issues of equivalence of concepts. Although, on the one hand, the *small-n* scale studies might produce bias on selection, quality and comparability of data, on the other hand, bias is "a systematic error that is expected to occur in a given context of research..." (Collier & Mahoney, 1996: 59). Therefore, small studies should not be abandoned because they could provide insights into the most developed evaluation systems that any government attempts to achieve.<sup>58</sup>

Returning to the research design and the "crucial triad" of *what, when and how to compare* raised by Pennings et al. (2006: 5) to respond to the question of *how*, the research will focus upon the evaluations of the health sector programmes in both countries to track the recommendations issued by evaluators. These programmes are the main case study for the evaluation of the hypotheses and the third factor to be considered in this comparative study. For this process, some considerations need to be highlighted.

A main concern of this research is to isolate evaluations performed of the social sector (at least in the Mexican case) programmes, in order to get a wider perspective of use of evaluations across other sectors within public administration. For instance, in the

---

<sup>58</sup>The 'too many variables, not enough cases' problem of comparison arises because the political world that is the research environment of political science is too rich and varied (that is, it consists of too many variables) for the researcher to be able to find enough cases to control for all the effects of these variables... (Ragin, 1987: 23-6 cited in Burnham et al. 2004: 71).

Mexican case, evaluations of the social sector programmes are coordinated by the CONEVAL, which operates exclusively for monitoring and evaluating this sector. However, this exceptional case (agency) does not operate to the remaining public agencies and sectors of public administration, excluding few cases in the energy and environment sector, better explained in the case study chapter. Moreover, the number of budgetary [public] programmes of the social sector is non-representative of the total of these programmes across the federal [central] government.

Then, the next options were the energy and education sector, which are the largest sectors with public resources. However, during the period under review (2007 to 2014) only a few evaluations have been carried out. Therefore, the best option to investigate the effectiveness of evaluations performed was the health sector. The choice of the health sector is due to it being a key sector in the modern welfare state, in terms of budgetary importance, public interest and political focus (Bovens, 't Hart & Peters, 2001: 606).

The ranking developed by the international World Health Organisation 'WHO' (Tandon et al., n.d., 2-3, 18-21) measures the health system performance and overall efficiency<sup>59</sup> of 191 countries, where Chile is ranked 33<sup>rd</sup> and Mexico is ranked 61<sup>st</sup>. However, it is important to bear in mind that the aim of this study is to examine how the policies of the health sector change in response to evaluation rather than an exhaustive analysis of the health sector in these countries.

Moreover, in the Mexican case, evaluations of the health sector were coordinated by three agencies<sup>60</sup> but currently coordinated by two agencies (Ministry of Finance 'SHCP' and the National Evaluation Council of the Social Development Policy 'CONEVAL'), which can provide more elements to address the problem of effectiveness when this is compared to the Chilean policy evaluation. Regarding its importance, according to the PbR/SED strategy, which is the core of the Mexican policy evaluation, the health sector has 268 budgetary programmes representing the third largest sector with public resources,<sup>61</sup> and in 2010, accounted for 6.2% of Gross Domestic Product (GDP)<sup>62</sup>. The sector comprises the Ministry of Health (SSA), the Mexican Social Security Institute (IMSS) and the Institute for Social Security and Services for State Workers (ISSSTE), as well as the health services offered to minor population by the Ministry of National Defense (SEDENA), the Ministry of the Navy (SEMAR) and PEMEX (Oil State enterprise).<sup>63</sup>

By contrast, the Chilean policy evaluation highlights that in 2013, the health sector accounted for 7.5% of Gross Domestic Product (GDP).<sup>64</sup> It comprises two sectors, the public sector constitutes the National System of Health Services and the Health

<sup>59</sup>This could be considered as the effectiveness of policy evaluation in the health sector. For instance, an effective country-led coordination mechanism to address problems identified in evaluations is an option, but also regular assessments or report of progress and performance of recommendations implemented. Clear linkages of the evaluation of key interventions in specific programme areas to the overall evaluation of the national health strategy. Systematic linkages between health sector reviews, disease and programme specific reviews, and global reporting. Comprehensiveness of review mechanisms in central and subnational levels. Key institutions and stakeholders have clear roles and responsibilities to adopt evaluation results and have mechanisms to provide routine feedback to sub-national stakeholders. Finally, results from evaluations are incorporated into decision-making, including resource allocation and financial disbursement.

<sup>60</sup>Until 2015, the Mexican policy evaluation was coordinated by the Ministry of Public Administration (SFP) along with the Ministry of Finance (SHCP) and the Council (CONEVAL).

<sup>61</sup>Expenditure Budget of the Federation for fiscal year 2015 (PEF 2015).

<sup>62</sup>Health at a Glance 2013: OECD.

<sup>63</sup>See 'Health system in Mexico', Gómez-Dantés et al., 2011.

<sup>64</sup>Health at a Glance 2013: OECD.

Services for the Armed Forces, and covers 80% of the population and the private sector, which supplies medical services to the 17.5% of the population (mostly the upper middle-class and high income groups).<sup>65</sup>

The focus of this study is on the federal level of both countries, where most evaluations of the health sector have been developed. However, the decision to focus on this level is also due to practical limitations such as availability of evaluations at the state level and the time and the length of the research. At this point, it is important to remind that the judicial and legislative branches in both countries are excluded to perform any evaluation (e.g. object under evaluation), consequently to be accountable of their performance. These exclusions are analysed in the case study chapters, including some progress recently made by Congress noted in the conclusions chapter.

It should also be borne in mind that, although the quality of evaluations is constantly referred in this research, this characteristic is not analysed in detail neither have an impact on the outcome of this research. Finally, another factor to be considered is the period selected to review evaluations of the health sector. Following the triad of Pennings et al. (2006: 5) this period of analysis means the *when to compare*. The time under review is mainly determined by the Mexican case and the PbR/SED strategy, which initiated since 2007. Thus, despite the Chilean policy evaluation was developed earlier, the review is from 2007 to 2014 for both countries selected to have the same point of comparison.

To sum up, this sector is a good reference offering an in-depth analysis of a set of evaluations of government programmes and how they may or may not lead to changes in policy. It is looking for a pattern in the occurrence of the effectiveness of policy evaluation in emerging economies in similar conditions. Hence, the findings of this sector across countries will also contribute to the investigation process providing crucial data for developing evaluation systems with regards to the problems inherent of different economies.

### **Documentary and analysis sources**

The thesis carries out the fieldwork research through a deep documentary analysis of evaluation in both historical and contemporary contexts. This analysis consisted of a close examination of a range of governance and government reports including, National Development Plans, Account of Public Treasury, Congress reports, internal evaluation reports (administrative, accountability, government), internal organisation manual, expenditure budget, programme and legislative materials, daily debates in Congress, laws comprising initiatives, government plans, projects and programmes. The analysis also covered academic journal articles, PhD theses, statistical databases and international organisations reports.

This analysis provide the basis for both tracking the evolution of policy evaluation in the two countries and for testing whether evaluation is effective, tracing programmes from their initial formulation and implementation, through their subsequent evaluation to the post-evaluation phase. Inspired by the historical analytical perspective the analysis also explores the development of evaluation in both countries highlighting important developments such as the actions taken to strengthen policy evaluation in Chile by Eduardo Frei during his administration

---

<sup>65</sup>See 'The health system of Chile', Becerril-Montekio, Reyes and Manuel, 2011.

between 1994-2000 and in Mexico, during the presidency of Vicente Fox between 2000-06. One of the main issues faced by this kind of study within different contexts is the difficult to find empirical evidence to assign values of every variable in each country, which makes analysis and operationalization particularly difficult.

Hantrais & Mangen (1996: 5) also add that difficulties arise due to differences in research traditions and administrative structures. In this regard, the elements of the study are confined to description, due to constraints in terms of data sources, time and funding. The sections on the effectiveness of the policy evaluation of both countries will be more descriptive through a deductive method, starting from hypotheses and moving towards their analysis and verification deducting conclusions from these premises.

### *Elite interviews*

Finally, the analysis is complemented by *elite interviews* with the main actors of the selected countries who basically are well-informed and connected people. This concept refers to an interview with an interviewee who is given special and non-standardized treatment (Riesman cited in Dexter, 2006: 18) due to hierarchical or political position. For this purpose, a semi-structured questionnaire with open-ended questions is applied by the interviewer; who lets the interviewee explain the problem, question or situation investigated (Dexter, 2006: 19). However, the balance is usually in favour of the respondent due to high levels of knowledge of the subject matter under discussion (Burnham et al., 2004: 205). The interview is a tool to tap political constructs – involving political actors' beliefs – that may otherwise be difficult to examine (Beamer, 2002: 87).

In this context, the aim of interviewing is to understand the process of evaluation in these countries to identify patterns and circumstances, which explain shortcomings in evaluation's strategy. It is important to note that some independent variables will be measured by indicators provided by organisations, whereas others that cannot be measured so easily such as the perception of operationalization of the monitoring & evaluation system depend upon interviewees' opinion to determine whether the outcome is present or not. Moreover, interviews are helpful to complement the analysis of possible causal mechanism linking independent to dependent variable(s). For this, questions regarding the quality of bureaucracy such as background, skills and experience, perception of the relationships between actors (Congress, executive, policy-makers), availability, reliability and quality of data and level of autonomy of public agencies are included (see annex 2).

Interviews are conducted on the basis of a *small-n* sample with officials in ministries and agencies, academic experts, specialists in NGOs, and politicians from legislatures, to understand the evaluation process from the point of view of those involved in it. These are structured in order to contribute to the discussion of the way in which evaluation is managed within the executive and between the executive and the legislature. As Beamer (2002: 87) suggests, this (elite interview) approach is the first stage to define concepts and the research question to be answered, and to examine important parameters guiding the elite's definitions of problems and responses to them (Aberbach & Rockman, 2002: 673).

Defining the sample entails compiling a complete list of key informants based on documentary analysis of administrative, government and accountability reports to identify representative people who have a say in this topic. As Burnham et al. (2004:

205) affirm “... one of the defining characteristics of elite interviewing is that some respondents may count more than others in terms of their influence on the decision-making process”. For this purpose, the sample of people interviewed is as follows:

Table 2

| Position/Country                   | MEXICO                 |                                | CHILE                  |                                |            |
|------------------------------------|------------------------|--------------------------------|------------------------|--------------------------------|------------|
|                                    | Congress (legislative) | Ministry or agency (executive) | Congress (legislative) | Ministry or agency (executive) | University |
| Deputy                             | 6                      |                                | 2                      |                                |            |
| Official (middle-level and senior) | 2                      | 11                             |                        | 1                              |            |
| Academic                           |                        |                                |                        |                                | 1          |
| Total                              | 8                      | 11                             | 2                      | 1                              | 1          |

Source: authors's own elaboration.

The sample is based on gathering a balance of interviewees from within the different branches of government and from outside. Firstly, to contrast views and perspectives regarding the evaluation sector, representatives from the legislative branch (deputies, officials and advisors) were selected. Secondly, the selection of interviewees within the executive embraced both those senior officials with a deep understanding of the latest reforms regarding evaluation and the most experienced middle-level officials' view in their role as coordinators of the evaluation and policy makers and executors of programmes of the health sector. The final category of interviewee included respondents from academic institutions who have been involved in the policy evaluation process and could therefore provide a particular perspective on the issue from a broader perspective.

The next step was to apply a standardized and semi-structured questionnaire with open-ended questions, starting with some closed enquiries to ensure that the interviewee knows about the research topic. To ensure confidentiality of personal information, and in order to prevent any inconvenience to those people interviewed, citations from interviews are anonymous. The first section comprises a set of closed questions seeking of general information of respondents such as gender, age, job position, length or years in public service, and formal education. Later on, open-ended questions seek not to impose a rigid framework, allowing respondents to express themselves freely to engage in a wide-ranging discussion. This is especially useful for those contentious questions, which require more explanation by the interviewee.

An example of this arises when referring to questions targeted at Congressional interviewees about the use of evaluation results for better decision-making within the budgetary process. Questions which provide important insights into the political factor or capacity building implicit on the adoption of evaluation findings within the policy process are not easily answered by the interviewees due to the risk of political retaliation. Hence, a follow-up question of whether Congress and the executive are coordinated to enhance effectiveness helps in providing a different approach to the subject.

Another factor to guarantee the reliability of the questionnaire is to construct convergent questions, which aim to investigate whether a respondent has a consistent orientation toward the point of interest (research objective). For example, questions

on matters such as the origins of evaluation, and the role of those who are or were responsible for it, could offer another perspective in terms of the different contexts when policy evaluation was developed (i.e. the 1980s). In this way the interviews provide a comparison between the time when policy evaluation was launched and the point of view of key informants about policy evaluation currently operating. The questionnaire includes particular conditions of the policy evaluation and health sector such as perception of the system operationalization, forms of coordination, exchange agreement, incentives, information and economic resources as factors to respond to the variables under analysis (see annex 2).

Lastly, as table 2 shows, there are considerably more interviewees from Mexico than Chile. This is a limitation as a result of the circumstances of this research and practical constraints for travelling to Chile. Thus, most of interviews conducted in Mexico were face to face and they kindly accepted being recorded. In other cases, they preferred to respond the questionnaire by e-mail. All the interviews conducted in Chile were coordinated and responded by e-mail.

This chapter has outlined the underlying research design and methodology which has informed the dissertation. It has highlighted the factors underpinning the definition and operationalization of the dependent and independent variables as well as the rationale for the comparative case study approach which will be elaborated over the following chapters.

## CHAPTER 3. Policy evaluation in Mexico

### Introduction

This chapter places the study of the Mexican policy evaluation in an historical perspective. The main reason for doing so is because there is a general perception and concern that in every new administration new policies, programmes and projects are created without looking back to the past to identify facts and success but also failures in public policy making. Therefore, it is important for the reader to know about the evolution of evaluation and the relevant institutional and political factors which have influenced the effectiveness of policy evaluation across time. This chapter aims at presenting evaluation activity in the Mexican government and the major initiatives in the field, which were generated (mostly) by the executive. In considering the above, the information is presented for the period between the 1930s and 1970s, and thereafter in six-year presidential terms up to the present. The evaluation performed by the Mexican government emerged as a way of controlling and monitoring practices.

Moreover, evaluation has been considered relevant to the extent that it is closely related to the budgetary process. As Mejía (2005:4) claims, the government interest was more linked to finance and public resources than to a review of the programmes and agencies' results. In this sense, the power exerted by the Ministry of Finance (MoF) has been decisive in driving evaluation. Thus, after a long period of ups and downs since the 1930s, in 1983 the planning process was integrated into the regulatory framework as the Planning Law and National Democratic Planning System (SNPD). The chapter also confirms that administrative reforms and doctrines of the New Public Management (NPM) boosted evaluation and promoted its 'virtues' in terms of transparency, accountability and democracy. However, although political conditions favoured the consolidation of evaluation, its operation remains unstable due to multiple factors. The chapter explores whether or not the relevant institutional and political factors have worked in favour of effective evaluations over time.

### Historical perspective

For much of its history, the Mexican political context has been characterized by an authoritarian regime and strong presidential system of government where the executive branch has an essential role of leadership and power. This power has even been apparent in the planning process and in the monitoring and evaluation practices. Meyer (2000: 50) claims that since Porfirio Díaz ruled the country from 1876-1911, 'el presidencialismo' (presidentialism) has prevailed.<sup>66</sup> This characteristic was emphasized in the figure of Lázaro Cardenas (president from 1934-40), who achieved the greatest level of control over the Mexican political system. Furthermore, Carpizo (1983 cited in Meyer, 2000: 50) defined this power as 'metaconstitutional' to explain the magnitude of power achieved by the presidency, especially in the second half of the 20<sup>th</sup> century. Moreover, the State was under the control of the political party PRI (Institutional Revolutionary Party).

Writing during this authoritarian period, Grimes & Simmons (1969: 74) suggest that 'The special relationship between the PRI, the bureaucracy, and the functional *interest groups* has, up to now, prevented the formation of other significant political

---

<sup>66</sup> Díaz recognised as a dictator ruled this era for more than 30 years (1876-1911). This period of rule is so-called presidentialism, in which the leader of the executive exerts authority over Congress through the majority of the party, and the latter cannot achieve independence to act as intermediary between citizenry and government (Duverger, 1992: 213).



structures, and thus, at least in one sense, it has inhibited political development'. Another analysis refers to the lack of appropriate systems of checks and balances for a democratic State (Meyer, 2000: 53). These factors might help explain the attempt to establish a proper evaluation system in the earlier years of the past century. Indeed, there is no consensus regarding the beginning of evaluation. Arguably, the 'Special efficiency Commissions' in 1926 were the earliest examples, which were developed to organise public services and to streamline government agencies. Oropeza (2005: 82) and Sánchez (1993: 943; 1996: 708) agree that there are precedents before 1930, when the First National Congress of Planning and the National Economic Congress were held.

Other scholars argue that after the depression in 1929, a commitment was made to look for an effective method for allocating public resources (Grimes & Simmons, 1969: 75). Indeed, in 1930 the General Planning of the Republic Law<sup>67</sup> was enacted as one of the first government efforts, establishing a National Planning Commission under the supervision of the Ministry of Communications and Public Works. However, for these authors, the fact of belonging to this ministry explains why the planning process was more focused on geography, urban planning, natural resources, communication and infrastructure issues, rather than on modern planning as it is currently understood.<sup>68</sup> For other authors as Mejía (2005:4), the government's interest in evaluation began in the 1940s and was more linked to finance and public resources than to a review of the programmes and agencies' results.

Some monitoring practices emerged in those years such as in 1947, when the Control Law of the Decentralized Agencies and State Enterprises was enacted to monitor agencies' performance but it was abolished under the next administration. Between 1958 and 1976, the absence of inter-agency plans, policy guidance of planning, methodologies, evaluation criteria and incentives to collaborate (Bailey, 1980: 16, 20) resulted in planning fragmentation and relegated evaluation to a secondary role whilst the budgeting processes were prioritised (see table 1 annexed). Economic planning became predominant due to the importance and size of the State's activity. Bailey argues this emphasis was in line with the economic challenges and the 'statist development' model or the so-called 'Mexican miracle' for the achievement of economic growth. In the 1970s, the model of 'desarrollo compartido' (shared development) implemented at the national level by government, required the establishment of 'internal commissions of administration' by public agencies to improve their efficiency, control and evaluation.

The Ministry of the Presidency became the implementing agency responsible for coordinating the reform of the public administration (Oropeza, 2005: 87). For this author, this was the first time (1970s) that the political discourse referred to the strengthening of the democratic state and the need for public sector reforms. Apparently, political factors interfered with these efforts from its inception. On paper, planning seemed to be an effective and ambitious policy instrument which aimed to control public resources and organize the expanding government. However, those

<sup>67</sup> Official Journal (Diario Oficial de la Federación, 12/07/1930).

<sup>68</sup> At a first glance, Pichardo (1972: 16) summarizes all of them in a wide spread of agencies. The most representative before the 1940s was the Programme Commission of Revolutionary National Party [Comisión del Programa del Partido Nacional Revolucionario] created in 1933 and the precedent of the 'Partido Revolucionario Institucional' (PRI) party. It was integrated by government officials and party members to carry out the first six-year plan of social and economic reforms (1934-40). The Special Advisory Committee [Comité Asesor Especial] created in 1935 by the President Lázaro Cárdenas was in charge of coordinating national planning policies.

with political power had a remarkable ability to side-line the planning process and there was a sort of passivity and resistance to act on its recommendations.

### 1976-1982

In *José López Portillo's* administration, and during his prior electoral campaign, he had repeatedly announced his intention to carry out a substantial government reform to increase economic and social development with “planning and programming as essential tools” (Lanz quoting Carrillo and García, 1981). Planning acquired significance within a wider vision of promoting federalism and regional balance (Chapoy, 2003:10-11). This strategy sought to tackle the economic crisis through government initiatives such as the political reform and the strategy ‘Alianza para la Producción’. In this context, the most important evaluation-related initiative was the streamlining of the federal bureaucracy through an administrative reform agenda (Bailey, 1980: 11). By attempting to link performance and management, the administrative reform was focused on the reorganisation of state enterprises and ministries, as well as on programme budgeting.<sup>69</sup>

In addition, the executive created the Ministry of Programming and Budget [Secretaría de Programación y Presupuesto ‘SPP’] responsible for the planning, programming, budgeting, monitoring and evaluation of public expenditure.<sup>70</sup> The agency was also responsible for policy implementation which is the basis for the development of a policy evaluation framework. As Lanz (1987) indicates, ‘for the first time, the assessment role was institutionalized and assigned to this office; in the past, these functions were usually performed by the Ministry of the Presidency and the Ministry of National Heritage’. Although the aim was to centralize these processes (planning, budgeting and evaluation) under one ministry, the President was interested in promoting the autonomy of sector coordinators in setting their objectives, public resources and evaluation, providing the managerial and policy autonomy referred by Verhoest et al. (2004: 105).

The autonomy given to agencies would help to release some pressure over the president – who controlled all public agencies –, rewarding top officials based on results and tackling problems of overlapping among agencies (Bailey, 1980: 23). However, the SPP was not very effective, at first glance, some reasons to modify the systems previously mentioned became evident in the next presidential-term. The *obese* State was the justification for amending the State legislation and implementing an administrative reorganisation to create the new Ministry of the General Controller of the Federation ‘SECOGEF’ amongst others. Although the SPP had an Undersecretary of Evaluation, it only had an office (General Office of Analysis for Evaluation) without a formal evaluation system.

From its inception, the SPP was modified several times to articulate the processes, did not have a clear evaluation process and was enforced to develop *ad-hoc* studies ordered by the President. Moreover, in 1978, a technical body called Accounting Office<sup>71</sup> was established, whose objectives were closely linked to the SPP’s goals. At the end of 1977, evaluation functions were transferred from the SPP to the Ministry of the

<sup>69</sup> For this purpose, the Legal Base and Guidelines for Organisational and Sector Performance in the Federal Public Administration was issued by the Presidency (22<sup>nd</sup> July, 1977). Regarding reorganisation, around 900 state enterprises were grouped under the supervision of ministries as sector coordinators, while the budgeting programme represents a transition between the traditional way of budgeting based on resources to budgeting based on objectives/goals and the medium-term (Bailey, 1980: 22; SPP, 1988).

<sup>70</sup> The Federal Government Organic Law (LOAPF, art. 32). Official Journal, 29/12/1976.

<sup>71</sup> Replaced after 1999 by the Federal Superior Audit.

Presidency while its structure (bureaucracy) remained in the SPP (Moctezuma, 1989: 88; Bailey, 1980: 27) and explained in the following period.

### 1982-1988

This presidential term is one of the most critical in terms of the economic crisis, which Mexico experienced throughout the 1980s and 1990s. According to the president *Miguel de la Madrid Hurtado* 'The *obese* State has been a critical component of the fiscal crisis to the extent that public resources are used for its administration instead of for solving social problems...' (Aguilar, 1996: 188). In those days a powerful new political grouping of 'technocrats' who belonged to the bureaucratic elite with a common background based on a neoliberal economic ideology (Llerenas & Huertas, 2004: 230) emerged. Their ideas were reflected in the growing influence of new ideas on the management of public administration. Mexico like other Latin American countries adopted some management methods and techniques from the private sector, such as public services delivery and performance evaluations to tackle and transform the bureaucratic paradoxes of public administration.<sup>72</sup>

The regime's so-called "Moral Restoration of Society" strategy and the slogan of *Moral Restoration* [Renovación Moral],<sup>73</sup> which referred to "an effort to return to sobriety and austerity related to the republican system..." was the driving force behind the amendments to the Federal Government Organic Law 'LOAPF' [Ley Orgánica de la Administración Pública Federal]<sup>74</sup> and the creation of the Ministry of the General Controller of the Federation 'SECOGEF' [Secretaría de la Contraloría General de la Federación]. The reforms of the LOAPF<sup>75</sup> (article 32 BIS) conferred powers on the ministry to regulate monitoring and evaluation activities:

...to carry out the assessments of ministries and state enterprises of the Federal Public Administration (APF) in order to promote efficiency and verify the achievement of programmes objectives.

This presidential term is well-known as the one that formalized the evaluation process in official documents. In 1983, a notable aspect was the creation of the *National Democratic Planning System* (SNPD) consolidated through the first National Development Plan (PND) 1983-1988 and the 'Planning Law' framed as part of the Constitution's reform.<sup>76</sup> This period marked the consolidation of the planning process for the first time since the 1930s, making it mandatory by law to have a planning document with detailed goals and objectives, to continue its implementation, monitoring and evaluation of government performance in the medium term. Although evaluation functions including planning, programming and

---

<sup>72</sup> Indeed, in those years the major elimination of state enterprises framed in the organic law began, continuing the administrative reforms undertaken by the previous administration; the argument to dissolve them was the high growth of this sector and their negative cost for the budget (Guerrero, 1989: 787). In 1983, there were approximately 1150 state enterprises in the Mexican government. By 1988 this number was reduced to 595. At the end of the process in 1991, the number of enterprises downsized was 969, which represented 16.75% in terms of Gross Domestic Product (Aguilar, 1996: 193; Rebollo, 1993: 157; Sixth Governance Report, 1988: 407 and Tamayo, 1992: 115).

<sup>73</sup> 1st. Governance Report (Sep. 1, 1983).

<sup>74</sup> This law represented a challenge by the government to organize federal public administration and state enterprises. LOAPF, art. 32. Official Journal, 29/12/1976.

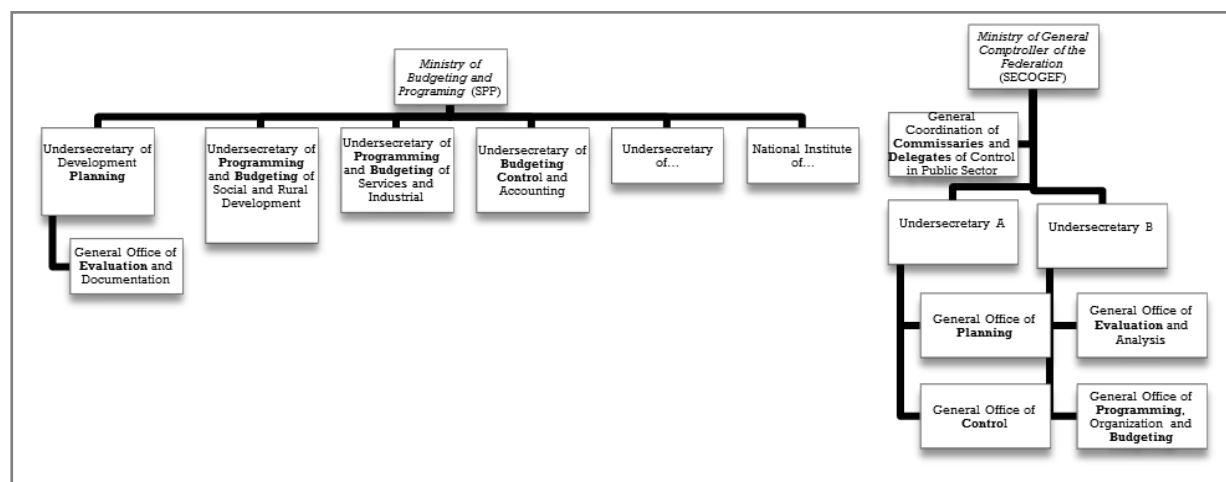
<sup>75</sup> Official Journal 29/12/1982.

<sup>76</sup> Regarding planning and conducting national development, the Constitution's reforms include the articles 25, 27, 28 and 115. The SNPD operates at three levels: global, sector and institutional. The law (art. 115) granted autonomy to the state and municipalities on the extent that the federal and state's executives agreed to coordinate actions. The result of this coordination was the State Planning System. In terms of time, the tools of the SNPD are classified in: i) medium-term programmes such as the PND, sector, special, regional and institutional; and ii) short-term programmes such as the annual programmes which include operational activities, goals, policies, resources and priorities of medium-term programmes. Secretaría de Programación y Presupuesto (1988) *Planeación democrática*.

control were shifted from the SPP to the SECOGEF, in practice, the operation was confusing even for these ministries, which were overlapping regulations and roles.

Figure 3

### Planning and evaluation functions in the SPP and SECOGEF



Source: Official Journal, Internal organisation manuals of the SPP and SECOGEF in 1983.

Especially during 1983, the rules for internal organisation manual issued by both ministries were similar in terms of responsibilities (see figure 3). On paper, both ministries had an office of evaluation, of planning and of programming and budgeting but in practice, evaluation functions inside the SECOGEF were distinguished from the SPP, because the former was more focused on overseeing of public resources, audits and review systems while the latter focused on monitoring plans and programmes. However, despite some relevant results<sup>77</sup> and the president's interest in establishing a system for monitoring, controlling and evaluating government performance, evaluation had not impacted on such performance. Administrative reforms did not use evaluation as the main tool for changing government management. These reforms seem more a response to the economic crisis than a process where evaluation helped decision-makers to determine the agencies' efficiency and dismantling "bureaucratic" structures.

By 1985, there was no evidence of any evaluation unit and the General Office of Evaluation was eliminated from the structure of the SECOGEF due to economic restrictions. Furthermore, the government did not link the main evaluation system called the System for Government Control and Evaluation (SICEG) operated by the SECOGEF with the SNPD, so they took separate paths. The former aimed to review public resources and infrastructure, while the SNPD organized a complex government structure through programmes, plans and agreements between the state and executive and with regional authorities.

### 1988-1994

The presidential term of *Carlos Salinas de Gortari* achieved strong political power and influence over public administration. As Arellano & Guerrero (2003: 3) point out '...the administrative structure became the arena in which contenders competed for the presidency and became the main channel through which conflict among social

<sup>77</sup> According to the Government (Sept. 1984-Aug. 1985) and Work (1988) Reports, more than 800 evaluations were performed by Commissioners to improve the administration of state enterprises, such reports also refer to the SICEG's implementation in states and municipalities, and to the promotion of self-evaluation reports.

sectors was managed'. There are some advantages and disadvantages of such power achieved. On the one hand, the disadvantage of a large structure of loyalties and oversight to control the government but on the other hand, such power was an advantage to the government in terms of the permanence and stability of policies, programmes and projects. In this context, the General Programme for the Simplification of the Federal Public Administration [Programa General de Simplificación de la Administración Pública Federal]<sup>78</sup> was the result of government reorganisation and the State reforms (see Aguilar, 1996). It aimed to consolidate an administrative culture fostering efficiency and productivity in government management, promoting evaluation of public services and tackling corruption.

The administrative reorganisation was also a response to pressures for economic and political reforms. However, for Arellano & Guerrero (2003: 10), this administrative reform is more difficult to understand than the other two major reforms undertaken by this government (economic and political). As these authors point out, substantial government transformation might have jeopardized the political control and power over resources, agencies and their bureaucracy. For example, the essence of the SPP – including evaluation functions – was transferred to other ministries until it was eliminated in 1992.<sup>79</sup> This action took place despite the president's power and interest to keep the SPP alive; the administrative reform had several allies but also opponents (Aguilar, 1996: 189), and the president did not secure for political support.

To achieve any major reform needs consensus among political actors (inter/intra governmental relations), who should strongly believe in it towards the creation of public value. Hence, in regards of implementing evaluation based on the results achieved<sup>80</sup> and failures of management of the economic policy during this presidency, the incentive to implement evaluation was probably ineffective due to the lack of political consensus between those assessed, and the consequences of bad or poor performance as a result of evaluations. The evidence of evaluation in this presidency is the evaluation of social programmes, which started in 1992 operated by the Ministry of Social Development [Secretaría de Desarrollo Social 'SEDESOL'] and eventually transferred to other sectors.

### 1994-2000

The early years of president *Ernesto Zedillo Ponce de León's* administration were constrained due to the 'error de Diciembre', when the economy collapsed at the end of the year 1994 and during 1995.<sup>81</sup> The period was also characterised by democratization (even some scholars Navarrete, 2008, Reyna, 2006 and Sáenz et al., 2007, argue that the process is incomplete), which was beginning to change the old

<sup>78</sup> Agreement for implementation, coordination and evaluation of the General Programme for the Simplification of the Federal Public Administration. Official Journal, 09/02/1989.

<sup>79</sup> Reforms issued by the LOAPF. Official Journal, 21/02/1992.

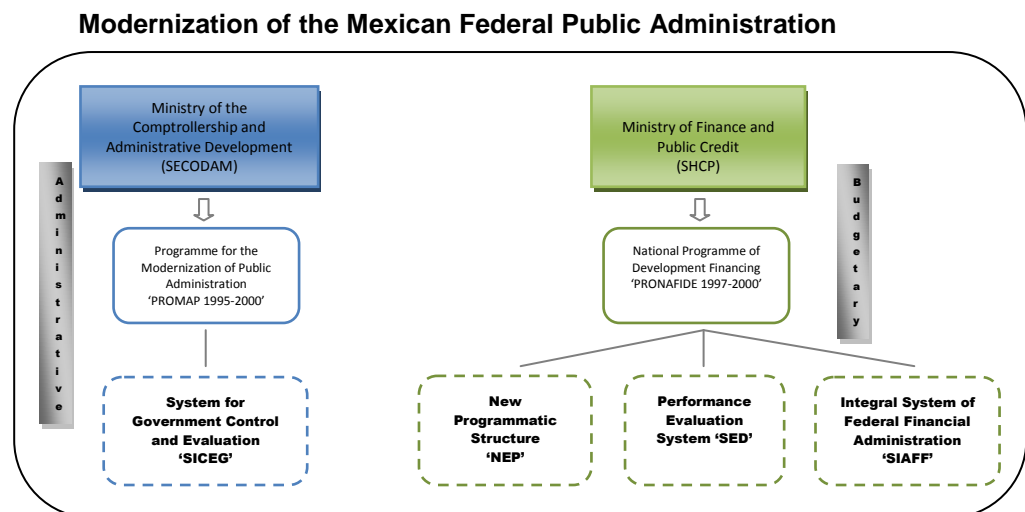
<sup>80</sup> The report of activities 1992-1993 [Informe de labores] of the SECOGEF in the section 'control, inspection and evaluation of public management' refers to evaluation results as a mechanism for the oversight of public resources, training courses of control and evaluation of public management for bureaucracy and Internal Controllers. At the regional and local level, activities were focused on the mechanism of Social Controller in the National Programme of Solidarity and the effectiveness of resources. Finally, a monitoring and evaluation of the president's campaign promises show that 5000 works and actions were executed. However, the report did not provide further information of these actions.

<sup>81</sup> The SHCP affirms that several factors occurred at the same time, the decrease in domestic saving, a rise in private consumption, deficit on public accounts, and external factors, such as an increase in interest rates on international markets which pushed investment capital towards other markets offering better benefits and less risk. See Decree of the PRONAFIDE, Official Journal 7/11/1997.

political system.<sup>82</sup> Although these political and external factors had some effect in government, the hegemony of the 'PRI' prevailed in the form of presidential power over all branches of government (Legislative, Judicial and Executive), and even civil society including trade unions, confederations, groups representing peasants and teachers (Arellano & Guerrero, 2003: 2-3).

Moreover, the democratic transition did not alter the immovable nature of the president as supreme leader. Indeed he was able to continue the administrative reform launched in the previous two administrations. From this moment until 2012, evaluation activities were conducted by the Ministry of the Comptrollership and Administrative Development<sup>83</sup> [Secretaría de Contraloría y Desarrollo Administrativo 'SECODAM'] and the Ministry of Finance and Public Credit [Secretaría de Hacienda y Crédito Público 'SHCP']. The former has the responsibility for organising and coordinating the SICEG, while the SHCP<sup>84</sup> was responsible for the assessment of the banking system, state enterprises' investment programmes, federal public expenditure and expenditure budget, and for the oversight of the compliance of national planning, programming, budgeting, accounting and evaluation. Figure 4 presents the connection between these ministries and systems.

Figure 4



Source: The Modernization of the Federal Public Administration and Budgetary System Programme, 1999. The SECODAM and SHCP.

The Programme for the Modernization of Public Administration 'PROMAP' 1995-2000 [Programa para la Modernización de la Administración Pública]<sup>85</sup> was the first initiative of the SECODAM.<sup>86</sup> The Programme comprised a scheme for the evaluation and measurement of public management, a comprehensive information system, and the Internal Control Bodies [Órganos Internos de Control 'OICs'] for better control of

<sup>82</sup> The first historic event was the elections of 1989, when Baja California became the first state dominated by the opposition party 'PAN' (National Action Party) and Mexico city also was led by the opposition party 'PRD' (Party of the Democratic Revolution). Afterwards, federal elections in 1997 transformed the power structure and the PRI no longer had a majority in Congress. The incorporation of Mexico into the Organisation for Economic Co-operation and Development (OECD) in 1994 also represented a major challenge to this administration leading to the need to transform the government.

<sup>83</sup> Afterwards the SFP (see table 1).

<sup>84</sup> Article 31, LOAPF. Official Journal 21/02/1992.

<sup>85</sup> See Decree of the PROMAP, Official Journal 28/05/1996.

<sup>86</sup> Through the Unit for Administrative Development 'UDA' [Unidad de Desarrollo Administrativo] responsible of the modernization project based on reforms' experiences in the OECD countries and state governments in Mexico. See Cejudo (2008: 116) and table 1.

public resources. The National Programme of Development Financing 'PRONAFIDE 1997-2000' [Programa Nacional de Financiamiento del Desarrollo] was the first special mid-term programme of the fiscal and financial sector issued by the SHCP.<sup>87</sup> Arellano & Guerrero (2003: 13) affirm 'The SIAFF (Integral System of Federal Financial Administration), NEP (New Programmatic Structure) and SED (Performance Evaluation System) have also been ambitious efforts to transform the traditional budgetary system into a performance-driven budget system'.

As Mejía (2005: 17) identifies, 'these systems were in line with international trends in budgetary matters, seeking to achieve a results-based approach'. The NEP aimed, for the first time in public history, to establish a comprehensive system of strategic performance indicators for the monitoring and evaluation of programmes and government activities, and to classify such programmes – at federal, state and local level – into 'budgetary programmes' categories, organising public resources into the so-called 'estructura programática' as it is currently known. In this context, evaluation became part of the large-scale of reforms inside the public sector, and it is remarkable that for first time the SHCP and SECODAM worked by common interest and agreement – at least in theory – to promote an evaluation strategy along with other administrative reforms.

However, these efforts were not enough to deal with the problem of control and power exercised by those who would have been evaluated. The implementation of the SICEG, NEP and SED could reduce discretionary control by political actors via evaluation results. In the end, the ambitious efforts to establish a large-scale of reforms could not secure support from such political actors. As Arellano & Guerrero (2003: 14) argue this might be a reason to reform public administration with a "neutral" technique instead of major reforms because the former is "painless" for the public sector and easier to implement. In other words, the government led the reform called 'the managerialist strategy',<sup>88</sup> alongside the economic and political reforms (Arellano & Guerrero, 2003: 10-11).

This strategy refers to the introduction of minor changes at the micro level (bureaucracy) through administrative techniques (re-engineering, total quality management, corporate planning, budgetary planning and service-oriented bureaucracy), which could transform the macro level and improve government's efficiency without jeopardizing the government's traditional way of functioning,

The hope of this managerial proposal is that freeing bureaucrats from micro-management (or over-involvement of Congress or controller agencies in specific management of the agency), pushing them towards evaluation of performance and competition, the administrative apparatus would be more efficient and responsive to society (Arellano & Guerrero, 2003: 10-11).

In those years, the achievements of government were limited to some 'Performance Agreements' signed by State enterprises. In addition, Congress modified the draft budget in 1999 to reallocate public resources to strategic social programmes, though it is unclear whether or not these amendments were due to evaluations. Although the SICEG produced some relevant results (see sixth Governance Report, Sept. 1,

<sup>87</sup> According to the SNPD, the PROMAP and PRONAFIDE are classified into the category of special programmes. The aim is to provide a stable macroeconomic framework, financial system and fiscal policy that enhance external and internal saving and investment, and stimulate growth in the long-term. Official Journal, 7/11/1997. It also establishes a quantitative baseline scenario for main macroeconomic variables, and stresses the importance of domestic saving to finance investment in order to reduce the economy's vulnerability (OECD, 1998:63).

<sup>88</sup> An OECD's strategy inspired and supported by global financial institutions and NPM practices (OECD, 1997 and Osborne & Gaebler, 1992 cited in Llerenas & Huerta, 2004: 232).

2000), these were more related to audits and oversights than evaluation of public management. As regards evaluation, a first assumption of limited results is that the financial, fiscal and administrative reform required technical capacity and political consensus (Arellano & Guerrero, 2003: 13) between actors involved in the process. For instance, a civil service with experience and knowledge is certainly needed and important as a strategic factor to influence any administrative reform, public sector modernization, efficiency and political transformation (Arellano & Guerrero, 2003: 2-3, 15 and Llerenas & Huerta, 2004: 233).

However, according to these authors, public administration was characterized by personal loyalties, patronage networks by those people who were looking for a job, relations, political support or recommendation. This fact reduced any possibility to establish a career civil service to ensure officials' skills and experience. Moreover, while an evaluation system needs time and proper data system for consolidation and homogenous implementation. This reform and evaluation system did not have sufficient time to achieve results. Additionally, indicators required by the system did not apply to some activities and ministries. Mejía (2005: 17) notes the cases of the Ministries of Internal Affairs, External Affairs, Commerce and the SHCP, and the difficulty of measuring their impact, their performance and other activities such as negotiations or promotions.

Although it was expected that the NEP's regulatory framework would be independent to avoid micro-management from controller agencies such as the SHCP and SECODAM (Arellano & Guerrero, 2003: 13), these ministries had to control it and deal with bureaucratic, inflexible procedures and regulations in order to implement the NEP homogeneously. In addition, there was no development of institutional mechanisms which would facilitate the use of evaluation results by Congress, to enable the reallocation of resources, nor were instruments developed between government and the public to diminish the gap in public information.

In the last year of this administration, the political system under the control of the PRI encountered a loss of confidence from the society, due to corruption scandals, mismanagement of public resources and inefficient public services (Llerenas & Huerta, 2004: 234). This along with other factors,<sup>89</sup> led the PRI to lose – for first time in seven decades – the presidency at the election in 2000. For Arellano & Guerrero (2003: 4), the change in the presidency marked the culmination of a long process of democratization.

## **2000-2006**

During *Vicente Fox Quesada's* administration, the PRI opposed to the majority of the executive reforms launched by the first president of the opposition party PAN (National Action Party), reinforcing the impression of the collapse of 'presidencialismo' (presidentialism). The assumption is that the previous presidency was the last stand of 'presidencialismo'. As Sáenz et al. (2007: 19) suggest '...not only was (government) alternated, currently we have in the presidency a leader who has to negotiate to enable him to govern, even with his own party – PAN – at Congress'. As noted by Reyna (2006: 131), 'the Mexican government has not dismantled its authoritarian structure although it has undergone a political change, the public debate is more open and citizens tend to be more participatory'. In this context, the new image of

---

<sup>89</sup> These include the citizenry politicization and organisation, loss of confidence in political parties combined with greater trust in electoral institutions, the continuing fall in the voting of PRI, media campaigns, political openness due to the negative image of the PRI, and the need to legitimize itself to the public (Sáenz et. al. 2007: 19).



the president still operates as key decision-maker (Arellano & Guerrero, 2003: 4), which indeed will help for further improvements in public policies.

The emerging democracy saw the development of evaluation regarding social policy. Firstly, despite only a few cases of evaluation of social programmes in 1992 inside the Ministry of Social Development [Secretaría de Desarrollo Social 'SEDESOL'], evaluation spread across the public administration,<sup>90</sup> thanks to the innovative evaluation model of impact of the programme Progresá (rebranded as "Prospera" in 2014) implemented in this administration. Indeed, during 2000-01, Congress broadened the obligation to evaluate social programmes with 'Reglas de Operación' (operating rules) (Cardozo, 2006: 122 and Medina, 2007: 20).

In these cases, Congress led the evaluations and was responsible for hiring external evaluators to develop them. By 2004, this proposal was formally enacted by the General Law of Social Development<sup>91</sup> [Ley General de Desarrollo Social 'LGDS'], along with the creation of the National Evaluation Council of the Social Development Policy [Consejo Nacional de Evaluación de la Política de Desarrollo Social 'CONEVAL']. In 2005, the CONEVAL emerged as a state enterprise<sup>92</sup> of the social sector to coordinate policy evaluation and to measure poverty, establishing criteria and methodologies for both purposes. In the medium term, the council worked in coordination with the SFP and SHCP to perform evaluation functions.

Secondly, in 2000 Congress gave further capacity to the Superior Audit of the Federation [Auditoría Superior de la Federación 'ASF'] also known as Congressional oversight or Audit Office for the oversight and control of public resources but also for programme evaluation,<sup>93</sup> which aimed '(t)o assess goals and objectives set in federal programmes on the basis of strategic indicators approved by the budget to verify their performance and legality of public resources'. The ASF adopted new evaluation functions and units such as the Programme Evaluation General Office. As Guerrero (2002: 190) points out, 'this was the result of a large transformation process of this agency inside the legislature'. In effect, this agency and the SECODAM performed the same activities concerning programme evaluation, albeit from the respective positions of the legislature and executive.

Thirdly, in 2001 the Good Government Agenda [Agenda de Buen Gobierno] became the most important strategy by the executive, which incorporated amongst other priorities<sup>94</sup> the Presidential Goals System [Sistema de Metas Presidenciales 'SIMEP'] for performance measurement. The SIMEP operated in 2002 as a top management

<sup>90</sup> Since 1992 (Official journal 25/05/1992) the organic law created the SEDESOL with the responsibility to conduct and evaluate the social development policy. In 1998, the budget reform embraced the obligation to evaluate social programmes, e.g. in 1999, subsidy and cash-transfer programmes should establish indicators to assess the effective budget execution. Moreover, in 2001 the Internal organisation manual (official journal 13/09/2001) ordered to put more emphasis on external evaluation by academic or specialized organisations. This operated through the Undersecretary of Prospective, Planning and Evaluation and the Social Programmes Evaluation General Office inside the ministry.

<sup>91</sup> 'To establish evaluation and monitoring mechanisms of actions and activities of the National Policy of Social Development' (article 1). Official Journal, 20/01/2004.

<sup>92</sup> Official Journal, 24/08/2005.

<sup>93</sup> Article 16 of the Superior Audit Law of the Federation [Ley de Fiscalización Superior de la Federación], Official Journal, 29/12/2000. Currently is the Audit and Accountability Law of the Federation [Ley de Fiscalización y Rendición de Cuentas de la Federación], Official Journal 18/06/2010.

<sup>94</sup> This agenda operated through six strategic areas (see table 1). Eventually, in 2002 (Official journal, 22/04/2002) the SIMEP along with the National Programme for Combating Corruption, Promoting Transparency and Administrative Development 2001-2006 [Programa Nacional de Combate a la Corrupción, Fomento a la Transparencia y Desarrollo Administrativo 'PNCCFTDA'] were integrated into the 'Agenda' <http://innova.fox.presidencia.gob.mx/archivos/8/8/3/files/archivos/sip-2944.pdf>

evaluation system, which identified high impact, processes and projects linked to the PND and set goals and deadlines for its sector programmes. Its main characteristic was the agreement between the head of the ministry or state enterprise and the president, under supervision by the SECODAM (Accountability Report 2000-2006: 157, Diener, 2007: 85, Medina, 2007: 50-1, Mejía, 2005: 21-2 and Pardo, 2007: 899).

Fourthly, the National Institute for Evaluation of Education<sup>95</sup> [Instituto Nacional para la Evaluación de la Educación 'INEE'] was established in 2002 as a public, decentralized agency with financial, technical, and operational autonomy for the assessment of the educational sector. Its roles were linked to the creation and development of educational indicators, statistics, tools, research and knowledge, as well as the measurement and evaluation of education issues and it was able to produce relevant findings within a relatively short period. Fifthly, in 2003 the Planning Law was modified.<sup>96</sup> Article 9 points out the importance of a system that could measure government performance,

The Executive shall establish an Evaluation and Compensation System for Performance to measure progress of ministries of the Federal Public Administration on achieving goals and objectives of the Plan –PND– and its sector programmes...

Another element in the long list of existing evaluation practices and systems was the reform to the organic law<sup>97</sup> (LOAPF) to transform the SECODAM into the Ministry of Public Administration [Secretaría de la Función Pública 'SFP']. Sosa (2011: 112-3) adds that the approach of this new ministry is based on a modified public management policy version of the Public Management Programme 'PUMA' of the OECD and by the Ministry of Public Administrations 'MAP' of the Government of Spain during the 1990s. The SFP developed a number of evaluation-related initiatives. In 2003, after several failed attempts, the SFP issued the Professional Career Civil Service Law<sup>98</sup> [Ley del Servicio Profesional de Carrera en la Administración Pública Federal], a key element of which was a system for the performance evaluation of officials (bureaucracy). This law is particularly relevant in emerging countries to improve the operation of the public sector and evaluation system, and to provide stability at the senior levels.

The following year, the SFP implemented the Integral Performance Model of Control and Oversight Bodies<sup>99</sup> [Modelo Integral de Desempeño de Órganos de Vigilancia y Control 'MIDO']. The MIDO was a comprehensive performance evaluation and management model<sup>100</sup> of the OICs designed to support public agencies in achieving their objectives and goals successfully (Franco, 2006: 69 and SFP, 2004).<sup>101</sup> Finally, in 2006 the Federal Law of Budget and Fiscal Responsibility<sup>102</sup> [Ley Federal de Presupuesto y Responsabilidad Hacendaria 'LFPRH'] was enacted. This law

---

<sup>95</sup> Official journal 8/08/2002.

<sup>96</sup> Official journal 10/04/2003.

<sup>97</sup> Idem.

<sup>98</sup> Idem.

<sup>99</sup> The model is based on a "modern" internal control approach of the COSO methodology (Committee of Sponsors Organisations) report published in 1992 in the USA, which emerged due to the failure of traditional control systems, which could not avoid and prevent financial and accounting frauds. Nowadays it is an international standard for public management (Medina, 2007: 57).

<sup>100</sup> The MIDO also was linked with the reform wave of the New Public Management, according to the CIDE, an academic institution that studied this model (SFP, 2004: 27).

<sup>101</sup> Guidelines for MIDO's design, integration, operation, evaluation and control. Official journal, 7/10/2004.

<sup>102</sup> Official journal, 30/03/2006.

embraced the budgetary process and the evaluation results of the PND's objectives and sector programmes, based on the 'SED'.<sup>103</sup>

This was related to the so-called Results-based Budgeting strategy 'PbR', seeking to link governmental planning, design, and implementation of public programmes to the budgetary process through evaluation:

The SHCP and the SFP verify every two months the implementation of programmes and budgets in ministries and state enterprises. To identify the efficiency, economy, efficacy and quality of the Federal Public Administration, as well as social impact of public expenditure... These indicators (SED) emphasize the quality of goods and public services, as well as citizens' satisfaction... (art. 111)

Clearly, these strategies gave a boost to evaluation, which previously had been only a rhetorical commitment. According to the Report of execution of the PND 2006 [Informe de Ejecución del Plan Nacional de Desarrollo], the 6<sup>th</sup> Government Report [Sexto Informe de Gobierno] and the Account of Public Treasury, the results were as follows. In the first place, the *Evaluation and Compensation System for Performance* was linked to the Professional Career Civil Service Law linking results of programmes and this system through bonuses.<sup>104</sup> However, the reports did not provide evidence about this process and evaluation results were not used as an input and incentive for the improvement of a career civil service. Besides that, in those years the Austerity Decree arose as a measure for reducing wages, salaries and downsizing the bureaucracy, among other subjects, being this, presumably, the reason for these poor results. In other words, the decree was an opposite action by government against the compensation system to improve the quality of bureaucracy.

The SIMEP and the Good Government Agenda had a similar ending. The main reason for their cancellation was due to the overlap of coordination between the SFP (the lead department) and the Office of the Presidency. Pardo (2007: 899) notes that this generated confusion amongst government agencies in terms of priority of actions and dual supervision. Regarding the SIMEP, she claims 'This managerial (private sector) and short-term mechanism, in practice, nullified the accomplishment and assessment of the long-term goals...' Another reason was the difficulty to achieve policy objectives. Initially, public agencies established ambitious goals but eventually they had to make amendments to deadlines and even to the goals in order to facilitate their achievement.

The SIMEP was adopted based on the obligation to meet these goals. This obligation is a pattern of coordination referred in the Intra Governmental and Interagency Relations 'IGR' variable. As Medina (2007: 51) suggests based on interviews of SFP's officials 'No head of ministry or state enterprise wants to be questioned by the president about the red or yellow goals, if it was, they reacted immediately; in this sense it was effective'.<sup>105</sup> Moreover, in practice, the consequences of the evaluation process were not as envisaged in the SIMEP's framework. The sanctions promoted by the OICs were informal, moral in nature, and had limited potential to change decision-makers' behaviour. Thus, the system worked only to the extent that the president was involved in following up its findings.

---

<sup>103</sup> This system operates with indicators for measurement of coverage, efficiency, social and economic impact, quality and equality, based on the logical framework approach methodology [Matriz de Marco Lógico 'MML'].

<sup>104</sup> According to articles 2 and 68 of the LFPRH.

<sup>105</sup> The SIMEP was a system for performance measurement based in a traffic light system, in which the president monitored the level of achievement of specific goals set by agencies.

As for the evaluations led by Congress, the CEF (2008: 36) asserts that the results were not as expected. Firstly, not all ministries submitted their indicators and evaluations to Congress and secondly, most government institutions had unsatisfactory marks or scores. This might be the reason for the creation of the CONEVAL. Finally, the executive agenda was focused on the SIMEP rather than the PbR/SED and the System for Government Control and Evaluation (SICEG). Although for Diener (2007: 85), the SIMEP was established on the basis of the budgetary system reform (explained in the following administration), the NEP, SED and SIAFF were the only established to do so 'for real results-based budgeting'. According to the CEF (2008: 33), the introduction of the SIMEP undermined the consolidation of the SED despite the importance it acquired in the Federation's Draft of Expenditure Budget 2001.<sup>106</sup>

The PbR/SED emerged in the transition period between presidents Fox and Calderon but did not formally operate till the latter's presidential term. The delay in its operation was mainly because the structure and evaluation staff required for its implementation were not finalised until the next administration was in power, and when there was a renewed interest in promoting results-based budgeting.

## 2006-2012

At the beginning of President *Felipe Calderon Hinojosa's* administration, the most relevant components regarding evaluation were the SICEG along with the PbR/SED. Certainly, in recent years, the issue of evaluation in emerging countries has become more prominent. Administrative reforms and doctrines of the New Public Management (NPM) undertaken during the past decades by most developing countries seem to be the core of rethinking the policy process, in which evaluation plays an essential role.<sup>107</sup> The new Presidency's commitment to evaluation was arguably reinforced by the financial crisis, which had a dramatic impact in Latin American countries.

For instance, in Mexico the annual GDP<sup>108</sup> growth rates fell from 3.3% in 2007 to -6% in 2009. Wiesner (2011: 32) argues that the world crisis in 2008 and 2009 increased interest in evaluation as an accountability and learning exercise, in order to determine its causes, implications and policy changes. Moreover, the need to address the budget problems caused by the crisis encouraged the implementation of systems for the monitoring, control and evaluation of public resources. In this context, the Mexican policy evaluation embraces various evaluation systems (see annex 4) but here is explained the two major systems.

The first of these systems was the Results-based Budgeting/Performance Evaluation System 'PbR/SED'. Although it was designed in the previous administration, this strategy was not pursued in a systematic way until the reforms of the LFPRH in 2006. For Vásquez (2006: 111), this budgetary reform was different from those implemented in the 1970s and especially from that in 1998, because it incorporated some elements of private sector management into the public budgetary process, such as the strategic planning, performance indicators and administration by objectives. There is

<sup>106</sup> Decree of Federation's Expenditure Budget 2001. Official Journal, 31/12/2000. See (ASF, 2007: 505)

<sup>107</sup> As mentioned in the review of the literature, political and public administration studies have underlined that evaluation must increase transparency and accountability, essential elements of a democratic state and efficient administration. (Arnold, 2004: 4, Cochran & Malone, 2010: 54-55, Hill, 2009: 278, 282, Jones, 2002: 84, Jones & Kettl, 2003: 2, Knill & Tosun, 2012: 176, Ospina, Cunill & Zaltsman, 2004: 232, Palumbo, 1987: 40 cited in Parsons, 1995: 545 and Wiesner, 2011: 30).

<sup>108</sup> Gross Domestic Product. International Monetary Fund [www.imf.org](http://www.imf.org)

no doubt that its implementation acquired relevance during this administration. In 2007, the SHCP's initiative<sup>109</sup> (Law) proposed the creation of the National Evaluation Council for Public Policies to replace the CONEVAL, with the aim of extending the scope of the PbR/SED across the whole federal programmes, not only the social sector, and providing it with greater authority and power. This initiative triggered friction in intragovernmental relations between the agencies (SHCP-SFP) competing to lead policy evaluation.

Separately, a member of Congress proposed a Law of Government Management Evaluation, which envisaged the “institutionalisation” of evaluation and the creation of a National Evaluation Council of Government Management to replace the CONEVAL.<sup>110</sup> However, the initiative was not approved by the legislative. The PbR/SED was spearheaded into both ministries due to the obligation to do so, whilst the CONEVAL continued leading the evaluation of the social programmes. This represents a major achievement for the PbR/SED, even more than for the SICEG, because the SHCP was given a further boost due to its power (public resources) and control over public administration, even though the SFP was supposed to have a wider responsibility for evaluate at federal and local levels than the SHCP.

To understand the structure of the Mexican evaluation system, it is important to refer to Ospina et al. (2004: 232) who suggest that a comprehensive evaluation model should considers the vertical and horizontal integration of the three levels of evaluation – *micro, meso and macro* – explained in the methodology chapter. In this context, the PbR/SED<sup>111</sup> is located at the *meso* level, aiming to provide information about the performance of programmes through evaluation. The first element in the figure below, is the quality of public expenditure through the Medium-Term Programme<sup>112</sup> (PMP), which was linked in 2008 to the Management Improvement Programme [Programa de Mejora de la Gestión ‘PMG’] run by the SFP, where even objectives seem to overlap.

The former (PMP) foresees “the results of evaluation emphasizing the quality of goods and services as well as citizen satisfaction”, while the PMG’s objective is to “maximize the quality of goods and services supplied by the Federal Public Administration”.<sup>113</sup> These similarities were used to create synergies between these ministries to avoid overlaps, but this occurred more as a result of improvement measures promoted by the SFP rather than a planning process for establishing a common programme.<sup>114</sup>

<sup>109</sup> The initiative also included the operational mechanism of the performance evaluation system and the medium-term programme for promoting the efficiency and effectiveness of public management. Eventually, they evolved into the SED and the PMP (later the PMG), respectively. Also, the evaluation of public resources will expand to state and local levels through the ‘Ramo General 33’. The planning, evaluation and mechanisms of transparency and accountability of public infrastructure projects are run by the SHCP. See: Iniciativa con proyecto de decreto que reforma, adiciona y deroga diversas disposiciones de las leyes Federal de Presupuesto y Responsabilidad Hacendaria; Orgánica de la Administración Pública Federal; de Coordinación Fiscal; General de Desarrollo Social; de Adquisiciones, Arrendamientos y Servicios del Sector Público, y de Obras Públicas y Servicios Relacionados con las Mismas. Official journal 1/10/2007. Available at:

[http://www.diputados.gob.mx/LeyesBiblio/proceso/lx/052\\_DOF\\_01oct07.pdf](http://www.diputados.gob.mx/LeyesBiblio/proceso/lx/052_DOF_01oct07.pdf)

<sup>110</sup> See initiative of Law of Government Management Evaluation by Xavier López, deputy at Congress.

<sup>111</sup> The strategy design includes a world-wide concept ‘Managing for Development Results’ and ‘Results-Based Management’ of the United Nations Development Group, the Organisation for Economic Co-operation and Development and the World Bank. See: <http://www.mfdr.org/>

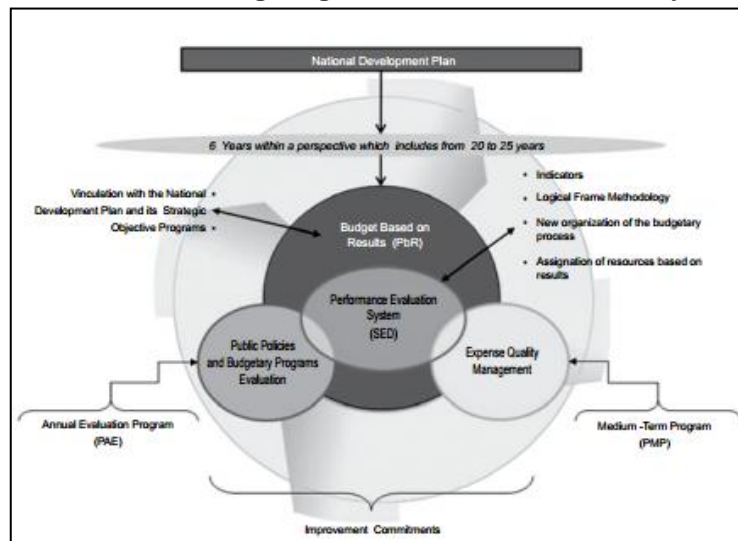
<sup>112</sup> Official Journal, 05/02/2009.

<sup>113</sup> Idem.

<sup>114</sup> Indeed, in the new special programme ‘Programa para un Gobierno Cercano y Moderno 2013-18’ mentioned there was deficient coordination inside public administration which reduced their efficiency and efficacy. This was caused by a disconnection of both programmes (PMP, PMG). Official journal, 30/08/2013.

Figure 5

### Results-based Budgeting/Performance Evaluation System



Source: SED, Ministry of Finance and Public Credit.

A key mechanism is the Annual Evaluation Programme (PAE), jointly issued by the SHCP, SFP and CONEVAL. This programme establishes the obligation of all government agencies to design the Matrix of Indicators for Results [Matriz de Indicadores para Resultados 'MIR'] of each budgetary programme, the type of evaluations (design, consistency and results, indicator, process, impact, strategic, complementary and specific), time of execution, and mechanism to follow-up evaluation result<sup>115</sup> (recommendations) on an annual basis. The diagram below shows the process.

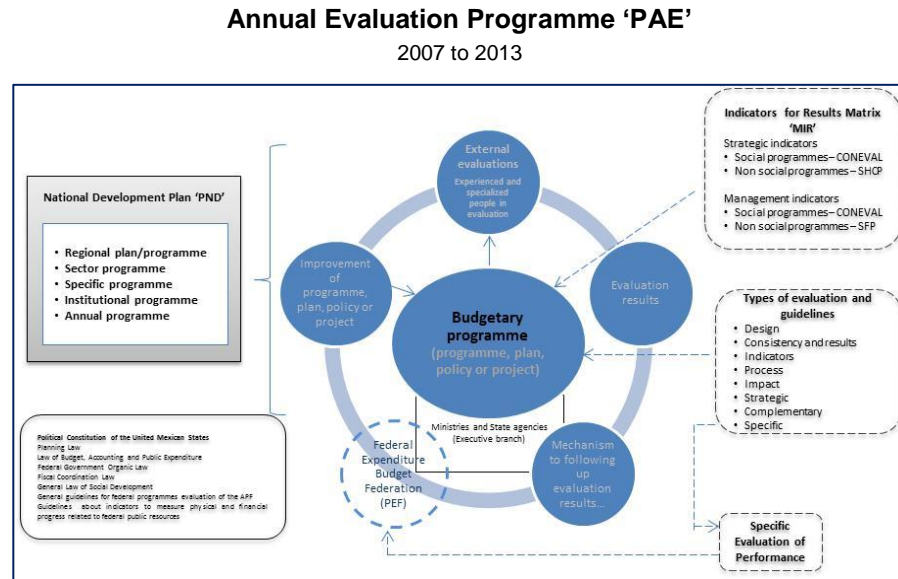
The second relevant component of the Mexican evaluation system is the SICEG located at the *meso* level similar to the PbR/SED. Since the reforms<sup>116</sup> of 2009, it changed to its initial name 'System for Government Control and Evaluation' [Sistema de Control y Evaluación Gubernamental], under the control of the new Evaluation Management and Government Performance Unit (UEGDG) in the SFP. This was created as a response to the demands for improvement in government performance and Congress recommendations. It also introduced a distinction between evaluation and control functions probably as an attempt to boost the former. The system should embrace other systems such as the career civil service, budgetary-programmatic evaluations, performance agreements and the MIDO as figure in the annex 4 shows. However the SICEG was not interconnected with them.

<sup>115</sup> Mecanismo para el seguimiento a los aspectos susceptibles de mejora derivados de informes y evaluaciones a programas federales de la Administración Pública Federal.

<sup>116</sup> Article 37 of the Federal Government Organic Law, official journal, 17/06/2009 and Internal organisation manual of the SPF (articles 24, 25, 76), official journal, 15/04/2009.



Figure 6



Source: author's own elaboration based on the SED and PAE.

**The Ministry of Public Administration. An expected ending or an effective political decision?**

Although the SFP performed a key role in policy evaluation for a number of years, its leading role has diminished over time,<sup>117</sup> A number of factors may have contributed to its collapse and dissolution at the end of 2012 and it was officially abolished<sup>118</sup> in 2013 and its functions should had been distributed amongst the SHCP, the new National Anti-Corruption Commission (in a stage of development) and the Presidency Office.<sup>119</sup>

In fact, the decision was taken in 2009, when the executive PAN-led executive called on Congress to dissolve the SFP,<sup>120</sup> along with the Ministry of Tourism and of Agrarian Reform. However, this initiative was rejected by Congress due to the opposition of deputies from the PRI party. Eventually, and perhaps ironically, when the PRI took over control in 2012, the new president made the same proposal and this time it was approved. Currently, it seems a regret on the decision made by the executive, because in the same presidential term, the president<sup>121</sup> promoted the new head responsible of the ministry (SFP), which it had already eliminated on the paper (Federal Government Organic Law) but currently still operating.

By the middle of year 2016, the executive took a political U-turn and abandoned the initiative, which indeed was reversed through the new reforms of the Federal Government Organic Law.<sup>122</sup> The latter reforms restored power conferred to the SFP (article 37). However, regarding evaluation, a year after, it was enacted the new

<sup>117</sup> See Sosa (2011) for implications of the SPF in the budgetary process.

<sup>118</sup> Reforms to the Federal Government Organic Law (article 37). Official journal, 02/01/2013.

<sup>119</sup> As for evaluation functions, there are some evaluation systems and monitoring practices awaiting to be reassigned, such as the evaluation performance subsystem of the career civil service, evaluation system at state and local level, the MIDO, bimonthly evaluations (expenditure quality and performance evaluation), performance agreements, evaluation of actions to accomplish the PND, the OICs evaluation activities, evaluation of the institutional plan and evaluation reports of government institutions to the Presidency Office.

<sup>120</sup> The law (initiative) was sent in 15/09/2009 by the Ministry of Internal Affairs to the Senate.

<sup>121</sup> See "executive actions to tackle corruption and interest conflict" by the SFP. 03/02/2015

<sup>122</sup> Federal Government Organic Law (article 37). Official journal, 18/07/2016.

Internal Organisation Manual of the SFP,<sup>123</sup> in which, evaluation operates separately from the PbR/SED focused on the public management. Moreover, the UEGDG formally disappeared and the roles and responsibilities were transferred to another unit within the same ministry.

### **The current state of affairs of policy evaluation**

As this chapter has shown, while there have been numerous attempts to establish evaluation since the 1930s, they have been constantly undermined by a variety of political, economic, legal and external factors (see table 1 annexed for better reference of the historical evaluation context in the Mexican government). Despite efforts in recent years to promote “policy evaluation” in the form of the SICEG, two law initiatives and even new evaluation agencies, the Mexican government has been more focused upon reforming the budget process than on developing an integral evaluation policy. Many factors are imputed to the SPF’s role in this process, which results had a severe impact on this ministry. Another argument pointed out by Wiesner (2002: 134-6) is related to the downsizing of the public sector, where evaluation has played an essential role to reduce public expenses efficiently. The conclusion here is that although many evaluations have been performed over a long time they have not had any impact on such expenditure.

For Sosa (2011: 117) ‘the main factor which has limited the SFP’s ability to respond effectively is what analysts called a return to political and discretionary management of technical institutions’, where discretionary decision-making trumps audit and control mechanisms and the professional civil service. Moreover, this confirms that the political control of the Ministry of Finance (SHCP) through budget has prevailed, despite the executive’s attempts to divide planning and evaluation functions into the SHCP and the SFP respectively.

There are additional issues implicit as well. Firstly, although the career civil service was established in 2003, with the aim of enhancing officials’ commitment and experience to improve public sector performance,<sup>124</sup> the ministry (SFP) lacks competent and experienced officials with the specific skills required for evaluation purposes. Secondly, there has been no continuity in policy evaluation, even when the same party was leading government, for example, the PRI in 1988-1994 and 1994-2000, and PAN in 2000-06 and 2006-12. Indeed, evaluation has not been considered as a formal element of the planning process while responsibility for it has shifted from the SFP, to the Presidency Office, SFP and SHCP. The SFP as coordinator of the PbR/SED has performed a secondary role.

Lastly, the SHCP’s internal organisation manual<sup>125</sup> created the new Performance Evaluation Unit to perform evaluation functions regarding the PbR/SED, which impacted to eliminate the SFP and the coordination between them to implement a comprehensive evaluation not only focused on the budget as the SHCP currently does. Bailey’s argument (1980: 15) that ‘conflict seems foreseeable’ is valid in a context where two ministries are responsible for policy evaluation, and the SHCP keeps hold of power through the control of the budget restricting financial autonomy of agencies. Finally, there was no political consensus on the role of evaluation. This source of weakness in intragovernmental relations (Franz, 1985: 483) amongst agencies and

<sup>123</sup> Internal organisation manual. Official journal, 19/07/2017.

<sup>124</sup> See Arellano & Guerrero (2003) for a better understanding of obstacles and problems in the implementation of the career civil service.

<sup>125</sup> Official journal, 10/10/2012.



stakeholders increased the complexity of joint action to carry out evaluations. In later chapter it will be analysed how this set of institutional and political factors influence the effectiveness of policy evaluation.

## **CHAPTER 4. Policy evaluation in Chile**

### **Introduction**

This chapter introduces the reader to the evolution of evaluation in Chile from a historical perspective. Like most other Latin American economies, it is classified as “developing”, “emergent” or “in transition”.<sup>126</sup> As in these countries, the Chilean approach to policy evaluation has been transferred from administrative reforms carried out elsewhere and is based on a Results-based Budgeting ‘PbR’ strategy. However, while Chile and Mexico have borrowed from other countries’ policy approaches, they have taken separate paths in order to develop a unique system congruent to local settings. It is also important to bear in mind that Chile is one of the most developed Latin American countries. The analysis is organised on the basis of the different presidential administrations and follows the development of policy since the end of the military regime and the subsequent commitment to a modernization of the State.<sup>127</sup>

The most important actions taken in every administration are included to explain how evaluation has developed to become more effective. Arguably, the golden age of evaluation occurred in recent years (2000s) under the responsibility of the DIPRES inside the Ministry of Finance (MoF). This unit developed a System of Evaluation and Management Control ‘SECG’ and a Management Improvement Programme ‘PMG’. The former operates as a mechanism for the achievement of agencies’ management goals established in the latter, which are linked to economic incentives.<sup>128</sup> The positive results achieved by these mechanisms are well-recognized as a worthy and successful experience for further dissemination abroad and have been adopted by the public sector across Latin America region.

The chapter explores the development of the Chilean approach over the last 25 years and examines some of the current changes in that approach which aim to improve its operation. However, regarding the data obtained, there is considerably more information of Mexico than Chile due to the circumstances of this research explained in the methodology chapter and practical constraints to gather additional data from Chilean policy evaluation and public management from other sources than those published online.

### **Historical perspective**

The Constitution of 1980 enacted the Chilean political system as a democratic republic (article 4) and the administration of the State as decentralized (article 3).<sup>129</sup> The government is under the responsibility of the President as the State’s chief who is elected for a four-year term (article 24) with no re-election for the next period (article 25). The broad territory is divided into regions (intendencias) and municipalities (provincias) with financial and political autonomy from the central government (articles 99, 110, 107). To start the analysis it is important to mention that across the world, the State’s role was reduced in the 80s, given the economic and political crisis.

---

<sup>126</sup> According to Heady (2000: 132), some relevant features shared by them are inequalities of economic development, high-level of poverty, agricultural under development and weak currency compared to developed countries.

<sup>127</sup> Most periods ruled by the Chilean presidents were of four years until 2000, where the Constitution was reformed to extend the period to six. However, since 2006 the period went back to the original presidential term of four years.

<sup>128</sup> To understand the operation of both mechanisms see: [http://www.dipres.gob.cl/594/articles-60578\\_doc\\_pdf.pdf](http://www.dipres.gob.cl/594/articles-60578_doc_pdf.pdf)

<sup>129</sup> The country was ruled under a military dictatorship at the time.

In contrast to much of the rest of Latin America, where economic conditions were very difficult due to financial crises and where reforms of the state – including the development of evaluation – were required to improve its performance, in Chile the economy performed much better and public finances were in good shape, even running a fiscal surplus (Armijo, 2002: 275).

In Chile, therefore, the State's modernization took place later than elsewhere in the region and adopted some management methods later on and started the modernization – based on downsizing the bureaucracy –, mainly, due to the need to improve the economy and to address issues regarding social justice in the post-authoritarian Chile. Oddly enough the State reform was implemented during the military regime of Augusto Pinochet 1973-90 inspired by the Chicago school and its economic theories (Figueroa et al., 2011: 70-1). Certainly, the first set of reforms was the privatization of public organisations, health care and pension systems, as well as the decentralization of public services to municipalities (Armijo, 2002: 269); whereas, the evaluation system emerged during the 1990s within the "Concertación's government" (Pimienta, 2002: 128).

Muñoz (2005: 2) argues that the return to democracy in the 1990s led to growing pressure for the reform of State institutions. An ex-minister interviewed by Muñoz pointed out these challenges were complex to address because the country's bureaucracy was part of the old authoritarian regime – despite a democratic President in power –, including most of Congress people in the Senate who lacked expertise to implement reforms. Moreover, these Congress people represented a majoritarian opposition to the President's party. For this purpose, the creation of the Ministry for the Presidency's General Secretariat (Ministerio Secretaría General de la Presidencia 'MINSEGPRES') represented a major innovation and acted as the coordinator for the modernization of the state following the end of Pinochet's rule.

#### **1990-94**

The post authoritarian period of modernization is divided into two, the first during the presidency of Patricio Aylwin 1990-94 and the second during the presidency of Eduardo Frei (1994-98). Clearly, a key concern in the Aylwin Presidency was to preserve the fragile democracy, avoiding a return to the authoritarian regime, and downsizing the bureaucracy (Pimienta, 2002, Tello, 2011 and Weissbluth, 2005). A critical period prevailed in the Aylwin administration, in which he strived to restore the balance of power, reduce the influence of the military – the former dictator still commanded the armed forces – and cut the public debt which had grown due to mismanagement in previous decades.

However, during that time there was not pressure from citizens to force the government to prioritise a State modernization agenda or any reforms. As Orrego (1998, cited by Armijo, 2002 and Tello, 2011) confirms, "There are no citizens' claims for State's modernization". Moreover, according to Fazio (1996) the actions taken by the government did not impact the economic model inherited from previous government. As the ex minister of DIPRES [Budget Office] declared, "The State legitimacy crisis jeopardized consolidation of democracy and governance. It forced the implementation of some reforms" (Armijo, 2002 and Tello, 2011). For instance, his government introduced innovative programmes to reduce poverty and inequality, confronted the past by seeking consensus across different sectors and political parties and stimulated the economic growth (Armijo, 2002, Fazio, 1996 and Pimienta, 2002).

The President undertook some important administrative initiatives such as the Performance Management Agreements of the Nuclear Energy Commission, the indicators of the Budget Law, the Modernization Commitments and the Project of Management Improvement. Reforms also had impact in terms of evaluation. In the beginning of this administration the Ministry's Goal System (Sistema de Metas Ministeriales) was created by the MINSEGPRES as the first governmental evaluation initiative<sup>130</sup> (CLAD, 2001), and most important concepts such as efficiency and results-based management spread across the government.

### **1994-2000**

Under the Eduardo Frei's 'Concertación' administration or Concertación de Partidos por la Democracia (Coalition of Parties for Democracy) an Inter-agency Committee of Public Management Modernization<sup>131</sup> (Comité Interministerial de Modernización de la Gestión Pública) was established. To some extent, it arose as a response to the State crisis with the aim of introducing a managerialist strategy instead of a major administrative reform (see Armijo, 2002: 278 and Muñoz & Stefoni, 2002). It was not until 1997 that a leading group of qualified officials with expertise on modernization issues implemented the Strategic Modernization Plan 1997-2000 (Plan Estratégico de Modernización de la Gestión Pública).<sup>132</sup> According to Figueroa et al. (2011: 83-4) there were three main different actors who played a key role for the State's modernization: international organisations, the "new" political classes, and the top executives running strategic public services.

Although the plan was not regarded as a comprehensive reform, some components<sup>133</sup> were linked to the paradigm of the New Public Management 'NPM' (Armijo, 2002: 293 and Waissbluth, 2005). Indeed, Waissbluth (2005: 10) argues that the ministry's (MINSEGPRES) goals (such as linking budget to goals and performance indicators and incentives to improve the quality and innovation of services); reflected the influence of reformists within the administration, which gave a sign to the citizens of such modernization. While there were some attempts to implement monitoring mechanisms such as indicators for the budgetary process and some indicators were introduced in the draft budget law in 1995 by the executive through the Budget Office (Dirección de Presupuesto 'DIPRES') (Pimienta, 2002: 128), they were not used as an objective mechanism to inform policy-makers.

It was not until 1998, when the first attempt to introduce evaluation as a formal process emerged with the purpose of improving public management. The Management Improvement Programme (Programa de Mejoramiento de la Gestión 'PMG') was enacted by the law<sup>134</sup> no. 19.553 to improve public utilities/services and agencies by linking policy goals to incentives 'performance bonuses' (in cash) incentivising officials responsible for programmes to accomplish the goals (see DIPRES, 2001a and Guzmán, 2005). This process was thanks to the support and training from the Inter-American Development Bank and the World Bank. This assistance

<sup>130</sup> See: <http://siare.clad.org/siare/innotend/evaluacion/chile2/todo.html>

<sup>131</sup> The Committee consisted of the Ministry of Finance, Presidency, Government, Internal Affairs, Work and of Economy.

<sup>132</sup> Before this programme, the modernization of the government was a concept boosted by opposition parties claiming for the downsizing of the government agencies and bureaucracy and the more transparency on remuneration systems (Armijo, 2002: 292).

<sup>133</sup> The legislation focused on six areas of reform: strategic management, transparency and probity of public management, service quality and citizen participation, human resources, State's institutional framework, communications and extension.

<sup>134</sup> Law no. 19553 of February, 1998 and Decree no. 475 of May, 1998.

addressed the limited experience of Chilean officials in evaluation subjects and focused on the 'logical framework approach' (DIPRES, 2001: 18-19; DIPRES n/d).

However, branches such as the legislature or judiciary were excluded from the evaluation process. Indeed, Armijo (2002) and Tello (2011) highlight the fact that this process was discussed outside the parliament without any consensus between the legislature, executive and citizenry. Given the economic stability, the legislature was not consulted on these reforms. As these authors suggest, perhaps a stagnant evaluation process emerged mainly because this policy would not produce results in the short-term, which made it less profitable and attractive for them (Congress).

## 2000-06

In the following administration, the president Ricardo Lagos 2000-06 continued with this administrative reform but undermined the Political Reform and Modernization of Public Management 'PMGP' (Política de Reforma y Modernización de la Gestión Pública) carried out by the previous government. The reform implemented in this administration included the international experience of countries such as New Zealand, the United Kingdom, Canada, Australia and the United States of America on the basis of the second generation of reforms framed in the Washington Consensus (see the review of literature). However, internal conflicts within the government for holding the power and leadership slowed the process of reform. In this context, the Project for State Reform and Modernisation 'PRYME' (Proyecto de Reforma y Modernización del Estado) was transferred from the Presidency to the Budget Office 'DIPRES' inside the MoF but progress on reforms was undermined by the lack of consensus amongst officials of both agencies.

It was only after several corruption scandals such as Coimas, MOP-GATE, MOP-CIADE, MOP-IDECON (Figueroa et al., 2011: 74) and Corfo-Inverlink, – indeed, the president Lagos was directly involved on them –<sup>135</sup> were exposed to the electorate that the pressure for reforms increased. The government embarked on discussions with opposition and business groups in favour to the PMGP. At this point, Armijo (2002: 296) states, citizens became more aware of the government's action and concepts such as modernization became a part of the public debate. Political-legislative agreements – after the corruption crisis – focused on the State's modernization (public management, decentralization and human and financial resources), transparency and the promotion of economic growth. These reforms represented a major achievement for democracy in the post-authoritarian regime.

Armijo (2002: 76) and Tello (2011: 259) highlight some factors, which might help to explain the success of these reforms compared with previous ones, notably the participation of technocrats or "tecnopols", highly qualified professionals who were technically equipped to formulate and implement public policies. These authors add that the effectiveness of this policy was due not only to the context of crisis in the Lago's presidency, but also to the continuity of such policy during the previous "Concertación" government and the president's support for it. The importance of the role of the president is a characteristic of "presidentialist" regimes, where the leader of the executive exerts such authority over the legislature through the majority of the party(ies). In such regimes (Duverger, 1992: 213), the parliament cannot act

---

<sup>135</sup> According to these authors, the "Coimas" case involved the bribery of a group of deputies who authorized – through the Ministry of Transport – the installation of an automobile's factory. The "MOP" cases are related to the Ministry of Public Works (MOP) and different companies to which the MOP paid and the public works never finished or they had done it at inflated costs/price.

independently as an intermediary between citizens and government as it can do in parliamentary or classic presidential regimes.

In Chile, these circumstances and the experience of an authoritarian legacy allowed these post-Pinochet administrations to promote a democratic government and implement reforms to overcome the events which occurred during that regime (Díaz, 2000 & Boeninger, 1997 cited on Muñoz, 2005: 3). This was in contrast to Mexico, where social movements such as those which emerged in 1968 (students' demonstrations) failed to mobilise the public into pushing for reforms such as greater democracy, openness and transparency of government. Here it is important to add that in Chile in 1998, the evaluation system was formally introduced by the DIPRES (Ministry of Finance), (Dussauge, 2012a: 185), which emerged from the PMG. Both the PMG and evaluation were coordinated by the DIPRES even though the modernization agenda – including evaluation – was coordinated by the MINSEGPRES (Presidency).

What was the role of the DIPRES in devising the policy evaluation? DIPRES itself was created in 1927, when the decree no. 1.924 enacted formal operation of the Budget Office 'DIPRES' (Dirección de Presupuestos). In 1975, the law no. 1.263 further established its budgetary authority "The Budget Office is the technical agency for allocation of public resources, tailoring budgeting and monitoring its execution... Additionally, for implementing tools for public management's improvement." Moreover recently, the office acquired more power when the head of the DIPRES, Mario Marcel 2000-06 created the Division for Management Control 'DCG' (División de Control de Gestión) to run the System of Evaluation and Management Control (Sistema de Evaluación y Control de Gestión 'SECG'). Since then, the DCG has been responsible for the evaluation of programmes, technical assistance and the improvement of public management.<sup>136</sup>

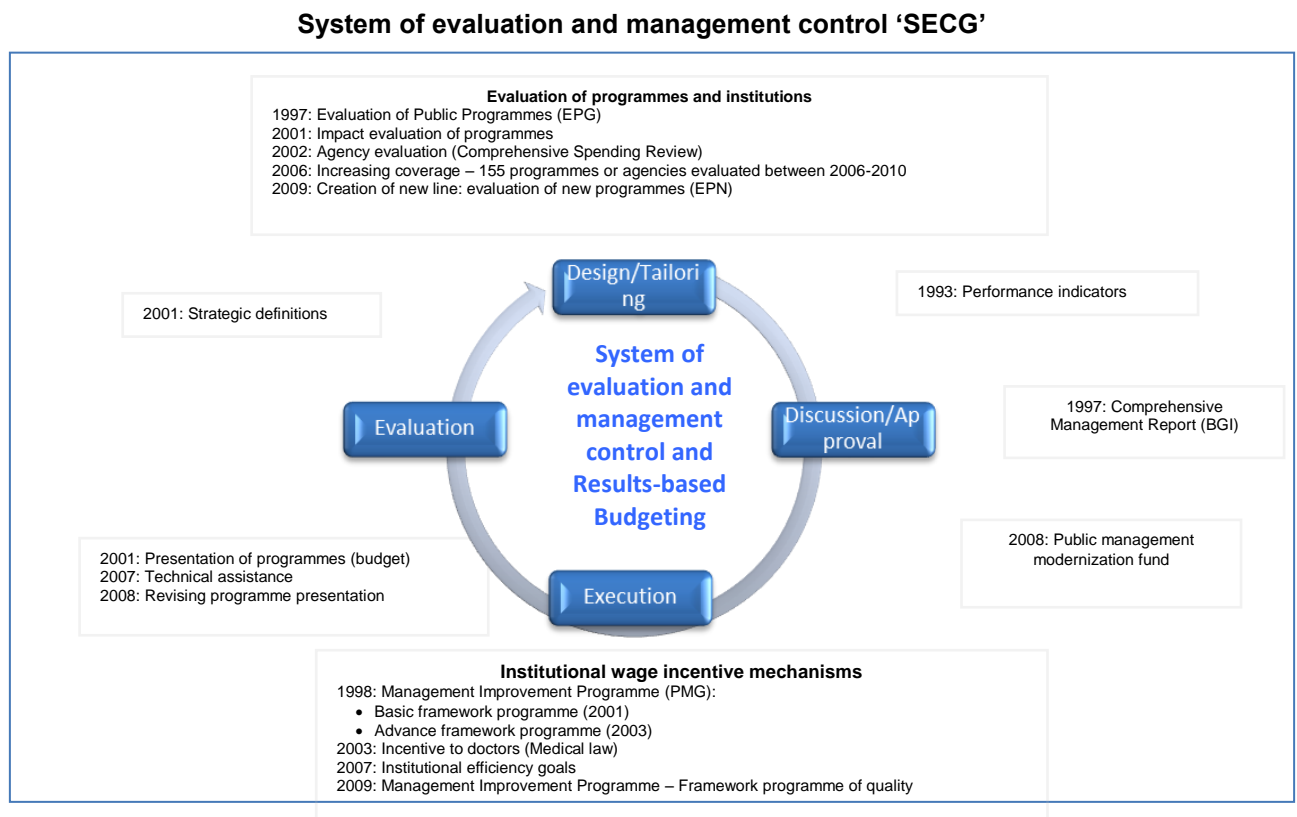
The two major achievements of the DCG were the 'SECG' and 'PMG'. However, the latter was revamped as a "system of benchmarking for the public sector" (Marcel, 2006 cited on Dussauge, 2013: 146) during Lagos' administration but eventually faced further changes during the Bachelet's presidency of 2014-18 (DIPRES, 2014). The DIPRES (along with the DCG) is responsible for the SECG, which consists of many evaluations and subsystems as shown in figure 7. The system embraces the following monitoring and control mechanisms: the performance indicators, strategic definitions (both closely linked), institutional efficiency goals, the PMG, and the Presidential Unit for Delivery Management UPGC (not included here and explained below).

It also includes some incentives (see quality of bureaucracy section in the chapter 'The effectiveness of policy evaluation in Chile'), which operate as a mechanism for encouraging the achievement of agencies' goals through the results of the PMG. Indeed, the incentive known as "public management modernization fund" was created to solve agency problems through incentives and feasible proposals tailored by universities or external consulting firms (DIPRES, 2008).

---

<sup>136</sup> See DIPRES' background <http://www.dipres.gob.cl/594/w3-article-3672.html>

Figure 7



Source: Budget Office (DIPRES), Ministry of Finance.

Regarding evaluations there are five main types:

- i) The evaluation of public programmes (EPG), which maps out the programme's design, objectives, internal organisation manual as well as its management and outcomes. This evaluation is based on the logical framework approach, performed by three external experts and often used in the draft budget formulation.
- ii) The impact evaluation (EI) aims to measure whether a programme/policy affects its beneficiaries based on a combination of qualitative and quantitative approaches comparing the beneficiaries with the non-beneficiaries.
- iii) The evaluation of new programmes (EPN) aims to provide a baseline for future performance evaluation based on a control group from the beginning for a period of 2 to 3 years. Coordinated by an international advisory panel and the University of Chile's Economy Department. This evaluation started to operate until the Bachelet administration.
- iv) The comprehensive spending review (ECG) or evaluation of organisational spending (EGI) (renamed afterwards)<sup>137</sup> evaluates the set of agencies of a given sector. It embraces a comprehensive number of factors such as objectives, management processes, organisational structures and functions, and services/products provided (see Dussauge, 2012a).
- v) The ex-ante evaluation emerged with the aim of gathering relevant information for the preparation of budget.<sup>138</sup>

<sup>137</sup> According to the DIPRES' office letter no. 1761, in 2011 emerged the *evaluación del gasto institucional* (evaluation of organisational spending 'EGI') replacing the comprehensive spending review (ECG) with the aim to get more accurate data of each sector (see annex 3).

<sup>138</sup> Formerly started in the Ministry of Social Development for social programmes. Later on, in 2014 the DIPRES continued this effort with the same aim of those of social sector but focused to the non-social programmes regardless of the fact that these were new, being reformulated or had poor performance (as result of evaluation in 2013). Despite all these efforts to introduce evaluations, in the end, it seems that some of these evaluations disappeared



Another element in the process is the summary report. Although, it is not another type of evaluation, according to the DIPRES (2008), this comprehensive management report is the basis for gathering information of evaluations for the draft Budget Law into a standard report distributed by Congress and made it available to the public.

The evaluations explained here are part of the SECG and were described in detail due to evaluations revised to test the effectiveness of evaluation belong to this system. However, the Chilean policy evaluation embraces a wider range of evaluation systems (see annex 7) such as the evaluation of schoolteachers' performance, the environmental impact evaluation and the evaluation of the social sector. In the section of policy evaluation framework of the chapter 'The effectiveness of policy evaluation in Chile' is explained the aim and context in which these were developed.

### **2006-10**

During the presidency of Michelle Bachelet, the SECG operated as a form of legitimization of public administration. It was praised by international organisations such as the World Bank (WB), Inter-American Development Bank (IDB), Organisation for Economic Co-operation and Development (OECD) and Latin American Centre for the Administration of Development (CLAD). As explained in the literature review chapter, the role of these organisations has a great influence exerting pressure over governments to adopt policy-transfer and promoting administrative reforms. Dussauge (2013: 61-5) noted that some developing countries in order to become members of such organisations "engage in a voluntary but constrained policy transfer";<sup>139</sup> whereas, others assume this influence to strengthen the capacity building of the State.

Some examples of these transfer processes are the public service agreements (PSA) – later called service delivery agreements –, citizen's charters and the Programme Assessment Rating Tool (PART),<sup>140</sup> which traveled from the UK and USA, across many countries including Chile, Mexico, Brazil, Australia, Netherlands, among others. A major change in the first Bachelet's administration was the introduction of the evaluation of new programmes (EPN) thanks to the influence of international organisations (Dussauge, 2013: 167).

### **2010-14**

In 2010, at the beginning of Sebastián Piñera's administration 2010-14, the president established the Presidential Unit for Delivery Management (Unidad Presidencial de Gestión del Cumplimiento 'UPGC') inside the MINSEGPRES and the Ministerial Coordination Division (División de Coordinación Ministerial 'UCM') to replicate the UK government's Delivery Unit. According to the IDB (Dumas et al., 2013), the Chilean government wanted the Unit to fulfil four key roles: i) strategic planning; ii) coordination of governmental action; iii) a follow up of the key government's goals; and iv) accountability.

---

afterwards (see policy evaluation framework section in the chapter 'The effectiveness of policy evaluation in Chile'). See DIPRES <http://www.dipres.gob.cl/594/w3-article-111377.html>

<sup>139</sup> This author highlights that mechanisms of influence of international organisations are coercive, normative, cognitive, executive, indirect political influence and the dissemination of best practices.

<sup>140</sup> This is a government's tool emerged from the New Public Management reforms pioneered in the UK with the aim of improving standards of public service and service delivered that citizens expect to receive and seeking citizen's empowerment (see [www.parliament.uk](http://www.parliament.uk)). Whereas the PART is a tool developed by the US Office of Management and Budget to assess the performance of programmes for further consideration during the annual budget process (see [expectmore.gov](http://expectmore.gov)).



Despite its short life from 2010 to 2014, the UPGC had positive results<sup>141</sup> such as the reduction of interministerial committees, institutionalization of the task of inter-agency coordination, the strengthening of the president's position, and the gaining of trust and credibility by the constituency.

### **The current state of affairs of policy evaluation**

In the end, the UPGC (MINSEGPRES) could not survive political tensions with the DIPRES during the second administration of Bachelet 2014-18. An assumption of the UPGC's dissolution is that the MINSEGPRES and DIPRES clash might be due to the prevailing role of the latter and similarities in their roles as coordinators of monitoring and evaluation practices, particularly those related to the UPGC's objectives. As some interviewees by Dussauge (2013: 170) argue, the UPGC has an unclear approach, "...it is not as different in comparison to the way we were working already. The focus is now on the product, the result to the citizen. They assume we previously cared only about procedures, but that was not the case". Currently, the UPGC has been abolished and no longer counterbalances the power of the Ministry of Finance (DIPRES) as had been intended while the DIPRES remains in place leading the evaluation.

Nor do the SECG (DIPRES) operates in the way that it was originally designed. Instead, a number of adjustments to the system have been introduced. For instance, there are not any evaluations e.g. EI, ECG, EPN and ex-ante performed or at least there is no data of them in the last years. This means that from the beginning of this presidency, the only evaluation performed across sectors is of public programmes (EPG), thus, it seems that the SECG operates with only one type of evaluation. Although, the ex-ante evaluation has been proposed during this period (2014-18) and mentioned as the best (benchmarking) practice emulating the same type of evaluation already run by the Ministry of Social Development, so far no evaluations of this type have been performed.

Moreover, the evidence indicates that evaluation is not as closely linked to the PMG as it used to be (see DIPRES, 2014). Instead, the coordination of evaluation is solely the responsibility of DIPRES without the involvement of the MINSEGPRES as was originally envisaged. Indeed, the DIPRES' report (DIPRES, 2014) affirms it has improved the PMG's mechanism and turned the management control and planning system into the "organisational performance monitoring system" (sistema de monitoreo del desempeño institucional) for monitoring outcomes and indicators. However, there is no reference of evaluation results linked to this new system. It seems they (DIPRES-PMG) took separate paths over time and are no longer part of an integral system.

The current president's agenda introduced significant changes to the policy evaluation as it used to be operated by establishing the National Evaluation Office, whose aim is to be the [only] agency responsible for evaluation and coordination of public agencies in the cycle of policy process. Indeed, the self-imposed goals are quite challenging. For instance, the attempt is to evaluate the 25% of the total of programmes each year. This means that by 2018 (at the end of present administration), all programmes will be assessed (see the government programme 2014-2018).<sup>142</sup> However, it is not yet clear whether this office will be placed separately

<sup>141</sup> <http://gestion2010-2014.cumplimiento.gob.cl/>

<sup>142</sup> <http://www.gob.cl/programa-de-gobierno/>

from the executive or whether it will continue to be supported by the Ministry of Finance (DIPRES).

## CHAPTER 5. The effectiveness of policy evaluation in Mexico

### Introduction

This chapter presents the results that test which factors are likely to influence the effectiveness of policy evaluation in Mexico. The first part is a case study of the health sector's evaluations performed from 2007 till 2014, in order to analyse different pathways taken by evaluations. It is focused only at the federal level due to the complexity of implementation at the state level and the time available for the research. The evaluations conducted are part of the Results-based Budgeting 'PbR/SED' process, which is the core of the Mexican policy evaluation, and the analysis is divided into two components. One is the review of programmes and whether evaluations performed led to changes classified as amendments, *status quo* or cancelation. The other examines the budgetary process at Congress to determine whether those evaluations were used for programmes' changes that is their effectiveness. Previously, it is explained how the federal budget establishes the connection between evaluation and the PbR/SED.

The second part provides a brief but comprehensive description of each variable, e.g. intra governmental relations, quality of bureaucracy, level of democracy, policy evaluation framework and degree of autonomy, in order to test their effectiveness. Also, in order to shed light on the practices between branches of government, the analysis draws upon other data, notably from interviews, which are needed to investigate the research question outlined in the methodology section. Thus, informal interviews with politicians and officials will be summarized.<sup>143</sup>

Lastly, this chapter brings together conclusions from the earlier discussion of 'Policy evaluation in Mexico' to analyse the particular forms of interaction between actors promoting evaluation. To enlarge the study of policy evaluation and not merely the PbR/SED, other evaluations performed by different agencies with different purposes will be noted in order to provide a comprehensive view of policy evaluation. The chapter concludes by raising some issues regarding the level of fragmentation of those institutional factors or variables, and placing them in order of importance regarding their impact on the effectiveness of evaluation.

### Results-based Budgeting 'PbR/SED'

In this section, the details of the documentary analysis of the PbR/SED evaluations results are presented. But first, an introduction to the process of planning the federal budget will offer a better explanation of how evaluation is associated with the PbR/SED.

How is the federal budget made?

Most public spending and country's finances are included in the budget.<sup>144</sup> The public budget reflects the priorities of constituencies, choices about what government will and will not do, but also political preferences influenced by the power of individuals and organisations interested in specific outcomes (Rubin, 2010: 345). For example, to accomplish a myriad of priorities in areas such as energy, health, education, and so on a large budget is needed. Thus, the way the Mexican government later divides such

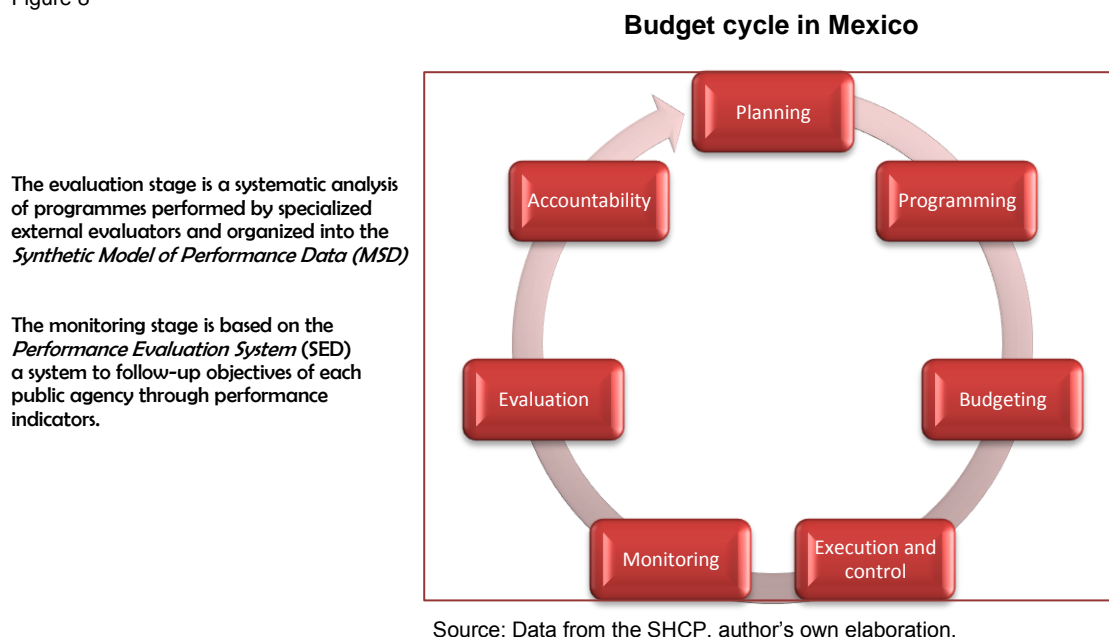
<sup>143</sup> The names of the people interviewed were changed in order to keep confidentiality of personal data.

<sup>144</sup> Except for some financial funds such as PIDIREGAS, IPAB, national debt and loans by state and local authorities. These loans do not require the legislature approval due to the autonomy of the states.

large amounts of money per area is into “budgetary programmes” to ensure better control and accountability regarding the budget.

A budgetary programme embraces the funds allocated per programme in areas such as poverty reduction, states or municipality finance, infrastructure, the oil industry, etc. According to the SHCP, the budget cycle has the following seven stages (see figure 8).<sup>145</sup> Regarding evaluation, the system called Synthetic Model of Performance Data (Modelo Sintético de Información del Desempeño ‘MSD’) is the database of budgetary programmes’ performance. The strategy is hereafter known as PbR/SED (Results-based Budgeting/Performance Evaluation System).

Figure 8



### Results of the programmes assessed

Therefore, the core of policy evaluation is conducted within the context of the PbR/SED strategy and evaluations are mostly government-ordered per year and developed by external evaluators, which are independent from the government such as academic institutions, NGOs, independent consultant or consultancy agencies. The period of time under review was determined when the strategy initiated in 2007. The health sector is the focus of this examination of the effectiveness of evaluation. As mentioned in the methodology chapter, this is a key sector in the modern welfare State, in terms of budgetary importance, public interest and political focus (Bovens et al., 2001: 606). Bearing this in mind, this study will examine the results of evaluations regarding the extent to which they change the policy or programme itself.

This study is more focused on the evaluations rather than on an exhaustive analysis of the Mexican health sector. This sector represents one of the largest sources of public resources of the Mexican government spending. For instance, in the Expenditure Budget of the Federation of the fiscal year 2015 (PEF 2015), this sector represented approximately 14% of the total budget. This means that the number was around of \$511 millions of mxn pesos equivalent to \$33.3 (millions US dollars)<sup>146</sup> plus those resources of medical services by PEMEX. This represents the third largest source of the

<sup>145</sup> Available at: [www.transparenciapresupuestaria.gob.mx](http://www.transparenciapresupuestaria.gob.mx)

<sup>146</sup> On June 2015, the exchange rate was of \$15.30 mxn pesos per \$1 us dollar.

budget, which was slightly less than the education and energy sector, which had the largest amount of the federal budget compared to the total.<sup>147</sup> In broad terms, according to the OECD data, in 2010 it accounted for 6.2% of Gross Domestic Product.<sup>148</sup>

These figures illustrate the importance of this sector in terms of resources but also the vast amount of services and programmes offered by it. Indeed, the sector comprises 268 budgetary programmes, plus 2 strategies and 3 budget accounts,<sup>149</sup> being 273 in total. In comparison with the 889 federal programmes<sup>150</sup> currently operating by the government, the health sector represents 31% of this total of programmes, which is representative in terms of number for the purpose of this study. It comprises the Ministry of Health (SSA), the Mexican Social Security Institute (IMSS), the Institute for Social Security and Services for State Workers (ISSSTE), as well as the health services offered to a minor population by the Ministry of National Defense (SEDENA), the Ministry of the Navy (SEMAR), and PEMEX (the State oil company).

As mentioned in the methodology chapter, to examine the trends of decision-making over time as a result of evaluation (the effectiveness of evaluation), it is important to identify whether the programme improvement could be attributed to evaluation. This analysis, therefore, looks at changes in programmes, which could take various forms such as amendments or modifications, cancelation or elimination and the *status quo* or unchanged categories. It also examines the allocation of financial resources to identify whether they increased, decreased or were unchanged after evaluations took place.

### **The effectiveness of evaluation in the health sector**

Table 3 annexed shows all evaluations performed between 2007 and 2014, based on the PbR/SED strategy. Of the total of 273 potential programmes mentioned earlier, only 45 have been assessed,<sup>151</sup> which is equivalent to 16% of the total – at the time of the study, one evaluation was still in progress to sum up 45 programmes assessed – and some of them have performed more than one of the different types<sup>152</sup> of evaluations available (depending upon the need of each programme/policy).<sup>153</sup> The following chart shows relevant results.

<sup>147</sup> See [http://www.apartados.hacienda.gob.mx/presupuesto/temas/pef/2015/docs/tomo\\_1/tomo\\_1\\_i14.pdf](http://www.apartados.hacienda.gob.mx/presupuesto/temas/pef/2015/docs/tomo_1/tomo_1_i14.pdf)

<sup>148</sup> Health at a Glance 2013: OECD.

<sup>149</sup> Estrategia Integral de Asistencia Social Alimentaria 'EIASA' (Integral strategy for food security) by SSA-SNDIF, Sistema de Protección Social en Salud (System for social protection in health) by SSA, Ramo 19 *Aportaciones a seguridad social* (Social security contributions), Ramo 23 *Provisiones salariales* (Wage provisions) and Ramo 33 *Fondo de Aportaciones para los Servicios de Salud 'FASSA'* (Fund of health services contributions).

<sup>150</sup> See <http://www.transparenciapresupuestaria.gob.mx/>

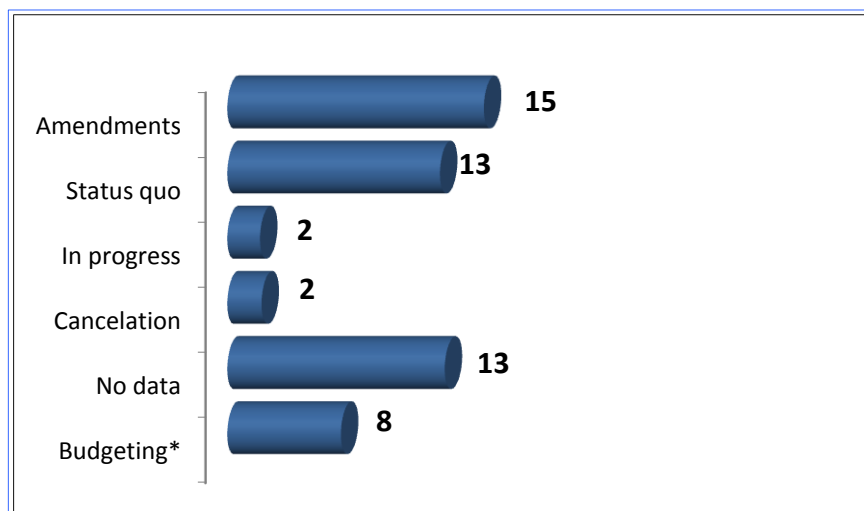
<sup>151</sup> Data provided by the SHCP in 2013, available at: [www.transparenciapresupuestaria.gob.mx](http://www.transparenciapresupuestaria.gob.mx)

<sup>152</sup> According to the evaluation guidelines there are different types of evaluation: design, consistency and results, indicators, process, impact, strategic, complementary, and specific (Specific Evaluation of Performance), see chapter 'Policy evaluation in Mexico'.

<sup>153</sup> Besides that, during 2014, new forms of assessment were included, which are not considered as evaluation due to the methodology used. These forms are the monitoring and evaluation card (M&E) and diagnostic applied to some programmes included in the Annual Evaluation Programme 'PAE' 2014. For this purpose, 54 programmes recently performed the M&E (51) and diagnostic (3). However, despite evaluators made recommendations there is no evidence on how these recommendations are inserted into the mechanism to following up results, and their methodology is not considered as a formal method to assess a programme.

Chart 1

### Effectiveness of evaluation in the Mexican health sector 2007-2014



Source: data from the SHCP, SFP and CONEVAL.

\* Programmes selected for budgeting process belong either to the amendments or *status quo* categories.

The classification of “no data” refers to those programmes in which there is no information available about them or any action taken by policy-makers. Thus, it is not possible to track neither the recommendation nor the evaluation.

Regarding these 45 programmes evaluated, the amendments are the starting point to look at. A programme is classified as an “amendment” when policy-makers have adopted at least 30% or more of the total recommendations suggested by evaluators in order to improve it in some way. The findings indicate that 15 (33%) of the total of 45 programmes evaluated achieved these results for improving them and continuing the strategic planning process. Some examples of these amendments are the programmes *Atención a la salud pública* (Provision of healthcare services) E001 and *Prevención y atención contra las adicciones* (Prevention of addictions) E025 operated by the IMSS and the SSA, respectively.

The former programme of *provision of healthcare services*’ aim is to reduce morbidity and mortality of diseases and those related to reproductive health. After the EED (Specific Evaluation of Performance) in 2009 and the EED in 2010 took place, all recommendations regarding the programme’s performance in terms of coverage, reduction of mortality of cancer and inclusion of new indicators were implemented.<sup>154</sup> Another example of this category is the programme of *prevention of addictions*, where most of the recommendations suggested by the evaluation of design performed in 2011 such as implementation of operating rules and improvement of indicators were accomplished; only one recommendation is still being implemented.<sup>155</sup>

Regarding the “termination” form, although two programmes of this sector were canceled, there is no evidence their termination was due to evaluation results. For instance, in 2007 the programme *Cruzada Nacional por la Calidad de los Servicios de Salud* (National crusade for the quality of health provision) operated by the SSA S041 performed an evaluation of design and ECR (Evaluation of Consistency and

<sup>154</sup> See CONEVAL’s report. Available at: <http://www.coneval.gob.mx/Evaluacion/Paginas/Evaluacion.aspx>

<sup>155</sup> See Ministry of Health’s report. Available at: [http://portal.salud.gob.mx/codigos/columnas/evaluacion\\_programas/cenadic.html](http://portal.salud.gob.mx/codigos/columnas/evaluacion_programas/cenadic.html)

Results) in 2007, however, there is no data of such evaluations. Besides that, in 2007 the programme changed its name to *Sistema Integral de Calidad en Salud* responsible by the same ministry (SSA) S202 but there is no evidence that those changes were affected in some way by evaluations.

In other cases, some programmes appeared in the budget only a few years before they were eliminated. Some examples are the programme *Construcción de centros de especialidad para pacientes ambulatorios y de unidades de atención hospitalaria* (Construction of medical specialized centers) ran by the SSA B001 or the programme *Subsidios y ayudas para gastos de funeral* (Grants for funeral expenses) by the IMSS J004, with no mention by the SHCP in the subsequent draft budget of the possible reasons of cancelation (see table 3 annexed).

Regarding the programme *Financiamiento equitativo de la atención médica* (Equitable funding of health provision) run by the SSA U004, all recommendations (6) of the EED 2008 were implemented in the same year. However, it was eliminated and integrated into another programme, the so-called *Seguro popular* (Health insurance) U005 operated by the same ministry (SSA), in order to improve the indicators of the latter (as pointed out by policy implementers). While it seems evaluation had an effect on such decision, there is no explicit indication this was the case.

Following the classification, conversely, in 13 programmes representing 29% of the total of programmes evaluated, the *status quo* category prevailed. For example, the programme *Reducción de mortalidad materna* (Reduction of maternal mortality) by the SSA U007 only considered 3 of 16 recommendations made by evaluators from the two evaluations performed in 2010 (design) and 2011 (strategic).<sup>156</sup> Apart from that, during the review of evaluations performed some issues arose. For instance, there are two programmes with the same key code “budgetary programme” (S201) and objectives. The *Seguro médico para una nueva generación* (Medical insurance for a new generation) was run by the Ministry of Health from 2008 until 2012, but in 2013 the programme merged into the *Seguro médico siglo XXI* (Medical insurance XXI) inside the same ministry sharing the same key code but without evidence of changes due to evaluations or other factors.

The weakness identified in these programmes and also linked to the *status quo* is that there was no data available to classify them in each of these categories. Other examples are the programmes *Fortalecimiento de las redes de servicios de salud* (Reinforcement of health provision networks) operated by the SSA U006 and *Mejoramiento de unidades operativas de servicios de ingreso* (Improvement of medical units) run by the IMSS E010, which were canceled with no mention of possible reasons for their elimination. Indeed, neither developed any of the evaluations designed by the coordinators of policy evaluation, in order to address the changes to these. A first assumption is that sometimes the lack of data regarding government actions and decision-making led the discussion to the *learning* categorization.

In other words, it might be that one or some of these canceled programmes reflected a learning process. Alternatively, perhaps the decision to cancel was the result of low

---

<sup>156</sup> See Ministry of Health's report. Available at:  
[http://portal.salud.gob.mx/codigos/columnas/evaluacion\\_programas/cnegsr.html](http://portal.salud.gob.mx/codigos/columnas/evaluacion_programas/cnegsr.html)



State administrative capacity to use the data from evaluations to support decisions such as this. The latter is in line with Howlett & Ramesh (1995: 178) reasoning where “simpler forms of formal and informal evaluations expect to occur without the necessity for any learning to actually occur within the state itself”.

Lastly, the ECR performed in 2014 of the programme of *social infrastructure projects of assistance and social security* operated by the IMSS, which should have been executed in the same year (2014), is still *in progress* according to the agency. Thus, there will not be any actions taken neither comments made by policy operators until results are revealed.

### **The effectiveness of evaluation in the Results-based Budgeting (PbR/SED)**

The analysis of budgeting is more extensive due to the complexity of the dynamics of the relations between actors and actions inside the executive power and the legislature. To examine the effectiveness of evaluation in the PbR/SED it is important to identify any connections between those changes in programmes due to evaluations and the allocation of financial resources, as well as any changes in the policy sector. Thus, the review of the evaluations performed indicates that they only had an effect on the budgetary process tailored in 2011 in eight programmes (see table 4), despite the attempts of the LFPRH (Federal Law of Budget and Fiscal Responsibility) to link governmental planning, design, and implementation of public programmes with the budgetary process through evaluation and the ‘SED’.

The document called Sistema de Evaluación del Desempeño<sup>157</sup> states that the main objective of the PbR/SED is that the “budgeting process is driven by results and boosting a new model to allocate public resources through evaluation of budgetary programmes...” Nonetheless, a review of the draft budget since 2007 shows that only in the budgetary process of 2011, the SHCP attempted to plan the budget based on the SED’s results,<sup>158</sup> as enacted in articles 25 and 110 of the LFPRH, and hence provide useful data to Congress in order to allocate public resources objectively. In 2011, the allocation of public resources to eight programmes (see table 4) was modified based on the results of evaluations and not only based on the draft budget of the previous year. Thus, this draft budget was presented by the executive (SHCP) as the first exercise in allocating resources objectively but also considered by the legislature for the same purpose (as shown in table 4).

However, it should be borne in mind that the programming of the budget is based on data from the year before. Indeed, due to the process implicit on developing evaluations, the data used is from two years earlier.<sup>159</sup> That is to say, the federal budget 2015 is programmed during 2014 based on data of 2013, while evaluation’s results are based on data from 2012. Indeed, a preliminary finding is that such gaps on the time of data collected are the main reason for not using evaluations. As a politician<sup>160</sup> who was interviewed affirmed, asymmetrical data received by Congress during the budget cycle is not useful for programming and budgeting. Here, what is important is whether the programme’s budget shrinks or increases, and what modifications to it were proposed.

<sup>157</sup> The Performance Evaluation System was published by the SHCP in 2007.

<sup>158</sup> See ‘Proyecto de Presupuesto de Egresos de la Federación para el ejercicio fiscal 2011, Exposición de motivos 2011 – Avances para mejorar la eficiencia y la calidad del gasto público -’, SHCP.

<sup>159</sup> See Evaluation of the SED by the Superior Audit Office of the Federation (2012) for additional information.

<sup>160</sup> Interview #16. 25/march/2014.



Certainly, in the fiscal year FY 2011 the SHCP stated that six programmes achieved their goals and proposed an increase in budget due to their results, while only two programmes saw their resources reduced based on their evaluation results. For instance, the programme's *Seguro popular* (Health insurance) aim is to protect households from the burden of out-of-pocket spending for medical care. Based on the external evaluation, the SHCP pointed out that the programme is well designed, it has indicators to measure results and achieved its annual goals. Thus, in the FY 2011 its resources were increased by 14% compared to 2010.

Table 4

Allocation of financial resources based on evaluations in Budget (PEF) FY 2011

| Public agencies  | Key budget | Budgetary programme  | Type of evaluation  | Budget 2010 (million mxn pesos) | Draft Budget 2011 (million mxn pesos) | Budget approved 2011 (million mxn pesos) | % of increase/reduction of public resources  | Comments   |
|--|------------|--|---|---------------------------------|---------------------------------------|--|--|--|
| Ministry of Health 'SSA'   | E010       | Formación y desarrollo profesional de recursos humanos especializados para la salud* (Training and professional development of health specialized human resources) | Ev. of design 2009, EED 2009  | 1,863.9                         | 2,873.3                               | 2,867.2                                  | In the PPEF 2011 public resources were increased by 65% compared to 2010 by the SHCP, whereas Congress approved 64%. | All recommendations (3) of the Ev. of design 2009 and EED 2009 were the same and implemented.<br><br>The SHCP mentions the programme achieved some progress on indicators.   |
| Ministry of Health 'SSA'   | E022       | Investigación y desarrollo tecnológico en salud* (Research and health technological development)   | Ev. of design 2009, EED 2009  | 1,197.3                         | 1,332.5                               | 1,406.7                                  | In the PPEF 2011 public resources were increased by 10%, whereas Congress approved 15%.                              | All recommendations (3) of the Ev. of design 2009 and EED 2009 were the same and implemented.<br><br>The SHCP mentions the programme achieved some progress on indicators.   |
| Ministry of Health 'SSA'   | P016       | Prevención y atención de VIH/SIDA y otras ITS (Care and prevention of HIV/AIDS and other diseases)   | Ev. of gender 2009  | 150.4                           | 165.8                                 | 243.8                                    | In the PPEF 2011 public resources were increased by 9%, whereas Congress approved 62%.                               | In the Ev. of gender 2009, evaluators pointed out five recommendations; however, these were not implemented by the health sector and not mentioned in the mechanism to follow-up evaluation results.<br><br>The SHCP mentions the programme achieved progress on indicators.   |
| Ministry of Health 'SSA' and the National System for Family Integral Development 'SNDIF' | S150       | Programa de Atención a Familias y Población Vulnerable* (Programme to vulnerable people and families)  | EED Specific ev. of performance 2009  | 382.0                           | 397.7                                 | 397.7                                    | In the PPEF 2011 public resources were increased by 4%, the same amount approved by Congress.                        | There was only one recommendation of the EED 2009 and it was implemented.<br><br>The SHCP mentions the programme achieved progress on indicators.  |
| Ministry of Health 'SSA' and Ministry of Social Development 'SEDESOL'                    | S174       | Programa de estancias infantiles para apoyar a madres trabajadoras (Childcare centers for supporting working mothers)  | Diagnostic of gender equality 2008 (Ev. of gender 2009), Ev. of perception 2009, Specific ev. of performance (EED) 2009 | 124.3                           | 202.4                                 | 202.4                                    | In the PPEF 2011 public resources were increased by 61%, the same amount approved by Congress.                       | It doesn't have indicator of budgetary process.<br><br>In the Ev. of gender 2009, evaluators pointed out five recommendations; however, these were not implemented by the health sector and not mentioned in the mechanism to follow-up evaluation results. In the Ev. of perception 2009 there are no recommendations (4) in the mechanism to follow-up evaluation results. Only one recommendation of 2 of the EED 2009 was implemented.<br><br>The SHCP mentions the programme achieved progress on indicators. |
| Ministry of Health 'SSA'   | S200       | Caravanas de la Salud* (Health caravans)   | Diagnostic of gender equality 2008 (Ev. of gender 2009), Specific ev. of performance (EED) 2009, Complementary ev. 2009 | 734.3                           | 726.9                                 | 726.9                                    | In the PPEF 2011 public resources were reduced by 1%, the same amount approved by Congress.                          | In the Ev. of gender 2009, evaluators pointed out five recommendations; however, these were not implemented by the health sector and not mentioned in the mechanism to follow-up evaluation results. Only one recommendation of the EED 2009 was implemented; however there are more pointed out by evaluators (6). In the Complementary ev. 2009 only 3 recommendations of 33 were improved.<br><br>The SHCP mentions the programme's diagnostic does not have positive results and needs to measure impact.      |
| Ministry of Health 'SSA'   | S202       | Sistema integral de calidad en salud (Integral system of health quality)   | Specific ev. of performance (EED) 2009  | 164.3                           | 162.6                                 | 166.5                                    | In the PPEF 2011 public resources were reduced by 1%, whereas Congress approved an increase of 1%.                   | The recommendations (3) of the EED 2009 are not mentioned in the mechanism to follow-up evaluation results.<br><br>The SHCP mentions the programme needs to  |

|                                 |      |                                    |  |         |         |         |   |  |
|---------------------------------|------|------------------------------------|--|---------|---------|---------|---|--|
|                                 |      |                                    |  |         |         |         |   | improve its indicators and mechanism to evaluate results.  |
| <b>Ministry of Health 'SSA'</b> | U005 | Seguro Popular* (Health insurance) | Specific ev. of performance (EED) 2009 | 48842.1 | 57267.8 | 56946.6 | In the PPEF 2011 public resources were increased by 15%, whereas Congress approved 14%. | There are no recommendations (3) as result of EED 2009 in the mechanism to follow-up evaluation results.<br><br>The SHCP mentions the programme achieved progress on indicators. |

Source: Data from SHCP (2010, 2011).

\*These programmes are considered as part of the social sector and coordinated by the CONEVAL.

In contrast, results were not positive for the programme Caravanas de la salud (Health caravans), which aims to promote access to medical services in geographically remote areas where health care services are unavailable. The programme did not achieve its goals regarding coverage, conditions of medical units and provision of medicines; consequently, its resources were reduced by 1% compared to previous year. Additionally, as shown in table 4, Congress approved a different amount in the final budget from that proposed in the executive's draft budget, which could be observed in the budget 'PEF' 2011. This indicates an attempt by the legislature to look at evaluation results for budgeting, thereby operating a strategy of results-based budgeting.

Dissimilar to this effort is the report<sup>161</sup> of the Superior Audit Office of the Federation (Auditoría Superior de la Federación 'ASF'), which claims that the finances of 17 programmes were allocated without any specific criteria for performance. The office points out that the PbR/SED did not identify efficiency, economy, efficacy, quality of programmes and social impact of finances as enacted by law. It also failed to identify the extent to which the evaluations were used for budgetary decision-making. However, currently they include most of these criteria in each programme's indicator except for those data regarding budgetary decision-making.<sup>162</sup> Another recent report (Account of Public Treasury 2013) of the ASF confirms the strategy (PbR/SED) has had few results:

distributional mechanisms [public resources] do not incentive the achievement of goals nor do they sanction inefficiencies. In this sense, the PbR/SED strategy almost had zero progress regarding public expenditure at federal level. ASF (2013: 50, 57)

Indeed, the office suggests training for local agencies supporting the strategy's implementation along with effective sanctions to those (states and municipalities) that do not accomplish the accountability of resources transferred. The evidence collected indicates few attempts to establish evaluation. Such discontinuities in the results of policy evaluation should be analysed in detail focusing on different phases of interventions in order to explain the gaps of the phenomenon. In other words, over time there is no linearity in the decision-making by either branch of government, as observed here it is more an isolated effort to use evaluations than to close the public policy cycle. Thus, the case study cannot address all the actions and decisions in one single event to identify a unique causal chain to explain it.

Here the problem of effectiveness is explained through diverse variables and under different circumstances. In analysing these attempts, for instance, the statement of purpose for the 2009 draft budget mentions some effects of evaluations. An example is the programme *Financiamiento equitativo de la atención médica* (Equitable funding of health provision) of the Ministry of Health (SSA) U004, which was merged

<sup>161</sup> See Evaluation of the SED by Superior Audit Office of the Federation, 2012.

<sup>162</sup> Another result observed is the difficulty for monitoring evaluation results of the PbR/SED from its initial phase, it means when evaluator(s) publish(es) a final report until policy makers take public action in the following years. Another complexity is to find evidence of evaluation within the legislature. This becomes a challenging issue, which requires an exhaustive documentary analysis of legislative materials such as daily debates (chamber of deputies), evaluation reports, the draft budget and the one approved, as well as audits.

into the programme *Seguro popular* (Health insurance) U005 and *Caravanas de la Salud* (Health caravans) S200 to improve their indicators. However, there are further reasons to affirm that evaluation has not played a formal role in the decision-making budgeting process, as the CONEVAL's report<sup>163</sup> states regarding the Ramo 33 (budget account),

the resources (economic) are paid out based on those received in previous years, when resources executed in every state depend on political arrangements instead of analysis of their needs... it seems there is no efficiency or equity criteria and it is assumed that the state's needs do not change over time.

The utilisation of evaluations by Congress is not limited to budget decisions. Evaluation could be used to support sector policies such as to target specific health needs of population, increase coverage of programmes, link different sector programmes to particular municipalities, etc. However, a review of daily debates under two different administrations from 2011 to 2013, in terms of the political party in power (PAN and PRI) undermines the potential value of evaluation over the health sector policy making process. In 2013, the initiatives of the law<sup>164</sup> of retirement pension and unemployment should have benefited the health sector, however, it did not mention whether evaluations could have influenced over health policy changes. This is no different of what the Centro de Estudios de las Finanzas Públicas<sup>165</sup> (Public Finances Studies Centre) 'CEFP' (2013a; 2014) indicates.

The comparative analysis of the draft budget and decrees of 2012, 2013 and 2014 indicates the importance of establishing mechanisms to modernize the budget cycle and use of data "gradually" from evaluations in the budget decision-making. Alongside, in the PEF (budget) 2012, Congress pointed out the obligation of the SHCP to evaluate the PbR/SED, including costs and whether these resources are properly allocated (CEFP, 2013). Thus, the exercise to allocate resources in 2011 by the legislature seems an isolated event, indicating that the intended role of evaluation has not been achieved. Demonstrating this branch has powerful reasons to keep old patterns in decision-making, even in the face of efforts to reform the budgetary process.

A comparison between the budget cycles in 2012 and 2013 shows that public resources of the health sector increased by 5.6% without any consideration of evaluations. This translates into 418.7 mdp (million mxn pesos) equivalent to \$27.2 (millions us dollars)<sup>166</sup> in 2012, while in 2013 it was of 443.7 mdp equivalent to \$28.9 (millions us dollars), with a further increase of 485.2 mdp equivalent to \$31.6 (millions us dollars) in 2014. Another example is the discussion of the PEF 2013,<sup>167</sup> where a senator requested from Congress more resources for vulnerable people. While it might be assumed that this decision would be supported by performance data, his petition was supported on health statistics with no mention of all evaluations having taken place.

So far, this study has provided an empirical review of whether evaluations have been taken into account by both branches. An *a priori* summary of outcomes indicates

<sup>163</sup> Informe de Evaluación de la Política de Desarrollo Social en México, 2008, p. 112, available at: <http://www.coneval.gob.mx/contenido/home/2509.pdf>

<sup>164</sup> CEFP (2013b).

<sup>165</sup> The CEFP is a body inside Congress acting as a supplier of opinion studies, analysis and forecasting, especially those related to public finances.

<sup>166</sup> On June 2015, the exchange rate was of \$15.30 mxn pesos per \$1 us dollar.

<sup>167</sup> Senator Francisco Salvador López Brito (PAN). Daily of debates, 28/09/2012. Mexican Congress.

that the executive is more influenced by the evaluations to take action than Congress. Indeed, it can be affirmed that the trend is to increase public resources systematically without considering the importance of the evaluation even in programmes with no visible results<sup>168</sup> such as the *programme of care and prevention of HIV/AIDS and other diseases* or the *childcare centers for supporting working mothers* (see table 4).

Similarly, inside the executive, actors are [narrowly] pursuing the objective of greater utilisation of evaluation results. However, there are other implications when evaluation refers to the formulation of budget by the MoF. In order to infer the answers, avoiding an imposed theory or hypothesis (George & Bennett, 2004: 91) a question emerges: *Why it is not available any supporting document - apart from that draft budget 2011 mentioned above - to imply to what extent the evaluation led decision-making in the tailoring process?* The variance of utilisation may be in instance best described in terms of other factors and context in which the policy is embedded.

#### *What explains the effectiveness in policy evaluation?*

To explain variation of the dependent variables within this country, the following section focuses primarily on such institutional (political) factors, which according to the literature are salient attributes of an effective policy evaluation. The main reason to keep a separate chapter of the variables per country is due to the extensive analysis of data needed to explain such variations, which could be difficult to understand if both policies are scrutinised together. However, the comparative study across the Mexican and Chilean policies is done in another chapter.

The section also analyses the gaps between those who operate programmes (executive) and decide upon expenditure (legislative) on the one hand and current practices of performance management on the other. For this purpose this section also uses another data source, the in-depth interviews with a sample of key informants (politicians, officials) representing actors involved in the development, implementation, monitoring, evaluation and auditing of policy to explore causal mechanisms of decision-making and identify the potential impact of the variables over public sector performance. Moreover, this section brings together the conclusions of the preceding chapters (methodology and policy evaluation in Mexico) to analyse particular forms of governance in Mexico. In order to do this, the variables here discussed are placed in order of their importance to influence the effectiveness of policy evaluation.

### **Intra-Governmental Relations (IGR)**

Intra-Governmental and Interagency Relations (IGR) constitute the most important factor in explaining the effectiveness of evaluation. The relationship known as Interagency operates between actors inner ministries, state enterprises and/or agencies inside the executive, whereas patterns of interaction between the executive (ministries/agencies) and legislative (Congress) are called Intra-Governmental Relations. Regarding the results of evaluations, performance audits and analysis of the health sector in Mexico, the variable to determine the interrelation among public organisations is divided into:

---

<sup>168</sup>The programme of HIV performed two evaluations (diagnostic of gender 2008 and ev. of gender 2009), which were merged into one. Of the five recommendations pointed out by evaluators none were not implemented by the health sector. Nonetheless, in the PPEF 2011 its resources were increased by 9%. Whereas the programme of childcare performed ten evaluations, in which, most of the recommendations 55 of 67 in total were not implemented. In the PPEF 2011 its resources of this programme were increased by 61%.

- 1) Forms of coordination: strong, weak or minimum.
- 2) Information: availability of data, timing, reliability and quality.
- 3) Economic resources to support policy evaluation: even or uneven

The forms of coordination to support effectiveness are minimal, so according to annex 1 the variable is coded “0” when these relationships result in minimum or weak coordination. During the implementation stage of evaluation, policy makers of programmes operate as administrators of data, whose job is to follow instructions by sector coordinators (SHCP, SFP, CONEVAL),<sup>169</sup> although they (coordinators and policy makers) are located at the same level as the executive. The task of implementation looks more complicated when actions need to be executed spanning two or more organisational settings (O’Toole, 2007: 143), such as integration of evaluation reports, which are developed by these three coordinators and the evaluation process itself. This means that the form of coordination is weak and the ministry (SHCP) responsible of the PbR/SED operates in its formal role as the main coordinator of this strategy.

Conversely, some officials interviewed believe that the coordination between agencies and the strategy coordinators has improved, although they believe the SHCP should be responsible for leadership and there is a lot of room for improvement. At this level, there is a pattern of authority lead by the organisation that controls public resources. Indeed, lead coordinator agencies apply the formal mechanisms of interagency coordination such as the PAE (Annual Evaluation Programme) to assure horizontal relationships and to strengthen a higher level of accomplishment by those programmes evaluated. The scenario of coordination between branches seems to be the minimum required to strengthen the evaluation. The situation appears worse regarding the relationship between the executive and Congress, as 50% of those interviewed of each branch did not know – or did not say – what form of relationship exists between them.

Only a few indicated that there was a relationship between Congress and the SHCP. Moreover, the relationship between the executive and legislative is constrained until the budgetary process starts and the evaluation results are sent by the executive to the legislature as required by law. Another example was found in the statement of purpose of the draft budget<sup>170</sup> in 2009, which notes that Congress (deputies) should have access to the SED for supporting their duties. However, there is no evidence this ever actually happened. In this context, the availability of information for those key stakeholders is essential for effective policy decision-making at any level of inter-agency relationships. Interviews conducted revealed some concerns about data availability as well as about timing, reliability and quality of data.

From the total of 19 interviews, those with politicians at Congress confirm that even when information of budget is available, it is of little use for decision-making because it is based on data that is two years out of date, which minimises the usability of evaluations. This outcome shows that the variable of availability of data is present and coded “1”, whereas the timing of such data is absent and coded “0”. Regarding quality, whereas officials in the executive consider evaluations to be of good quality

<sup>169</sup> From 2007 to 2009, the SFP did not have a special unit responsible for the PbR/SED strategy until 2010 when the Evaluation Management and Government Performance Unit (UEGDG) came into sight. Before that, evaluation was performed by the Unit of Control and Evaluation of Public Management (UCEGEP) inside the SFP. While in the SHCP, the Unit of Political and Budgetary Control (UPCP) used to lead the strategy until 2012 when the Performance Evaluation Unit emerged, although, in 2014 budget of the PbR/SED was allocated to the former (UPCP).

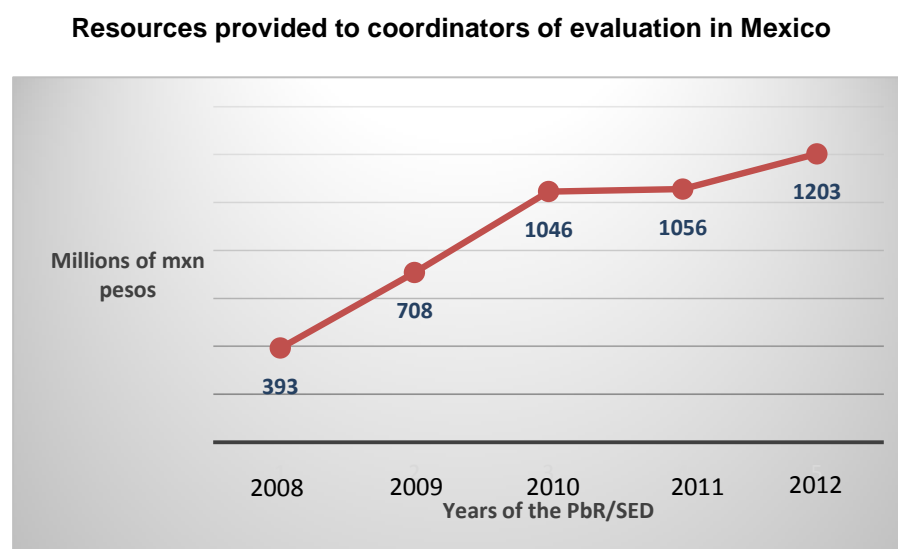
<sup>170</sup> See “PPEF 2009, exposición de motivos T.I.”, pp. 25-26, 47.

and reliable, for politicians, evaluations are perceived of dubious reliability because the executive coordinates them. Members of Congress therefore do not consider evaluations trustworthy, even though these are performed by external evaluators (such as academics or NGOs).

Consequently, the quality of the information is questionable due to the data provider. Indeed, Congress people regularly use studies developed by the CEFEP (Public Finances Studies Centre) as three politicians<sup>171</sup> affirm. However, this finding is puzzling since the CEFEP's studies make use of the same data, (i.e. the external evaluations' data coordinated by the SHCP, SFP, and CONEVAL). Thus, following the analysis of the variable measuring the reliability and quality of data, it is clear this variable should be and coded "0".

The third indicator is related to the economic resources. From the stakeholders' view, these represent the necessary tool for having direct control and coordination over actors when executing any policy. A review of the budgets performed in these years – since the strategy started – shows an incremental trend as it follows:

Chart 2



Source: Data from 2008 to 2012 by the ASF. The chart figures are in Mexican pesos (millions).

According to the ASF,<sup>172</sup> from 2008 to 2012, the coordinators (SHCP, SFP and CONEVAL) received \$4.4 (millions mxn pesos) equivalent to \$0.28 (millions us dollars)<sup>173</sup> to implement the PbR/SED, including the funds provided by international organisations such as the Inter-American Development Bank (IADB) and International Bank for Reconstruction and Development (IBRD). This amount

<sup>171</sup> Interview #14.13/march/2014; interview #16.25/march/2014; and interview #17.27/march/2014.

<sup>172</sup> See Performance Evaluation System's report by the ASF in 2012. The graph includes resources of four budgetary programmes linked to the Performance Evaluation System by each organisation such as the SHCP (P002, R005), the SFP (P005) and the CONEVAL (P003), as well as the Budget Account (Ramo 23) and funds by the Inter-American Development Bank and International Bank for Reconstruction and Development.

<sup>173</sup> On June 2015, the exchange rate was \$15.30 mxn pesos per \$1 us dollar.



represents 0.025% of the federal budget in those years. As chart 2 shows there has been an increase over time, only the SFP has reduced its finances.<sup>174</sup>

Regarding public resources, the continuity of policy evaluation requires a large amount of financial support to develop the PbR/SED, for which, coordinators have been ensured an even distribution of such resources to place the indicator as an outcome “1” present. As explained above, there was a commitment made by the executive but also by the legislature. In other words, despite misutilisation of evaluation data by Congress, this financial support was committed to the evaluation process through the approval of the budget presented by the executive (SHCP), as shown in chart 2, which responds effectively to the attempts of the executive to develop the PbR/SED. For this, the legislature approves funds to finance the operation of the evaluation (pay external evaluations) through their coordinators (SHCP, SFP and CONEVAL) and policy-makers inside the agencies.

Generating a successful policy evaluation means coordination among interdependent actors, facing current relationship set by legal authority. As O’Toole (2007: 145) says “B cooperates with A because B feels it is an obligation to do so”. The main results evidence the importance of the IGR for supporting continuation of evaluation to the end. At last, although the policy is designed at central level, it must be implemented by officials at horizontal or vertical level. Thus, the centre provides the finance but the other parts of government provide the intention and action.

### **Quality of bureaucracy**

To measure the second variable the following sources will be used:

- 1) Government effectiveness (The Worldwide Governance Indicators ‘WGI’)
- 2) Evaluation staff unit in the Mexican government: background, skills, and experience through interviews.

Mexico has a long way boosting a civil service (Pardo, 1999) in an environment where political control has been a key for stability and good governance under the post-revolutionary administrations. Although the Mexican government has promoted the goal of a career civil service<sup>175</sup> since 2003, this does not ensure that personnel have become more skilled (see Guerrero, 2007, Pardo, 2005, 2007). Therefore, an effective policy evaluation requires the involvement of skilled and experienced cadres in the administration. Let us explore how the quality of both bureaucracy and the evaluation coordinators affects the effectiveness of evaluation. In 2013, regarding the Worldwide Governance Indicators<sup>176</sup> Mexico was placed in the 50<sup>th</sup> and 75<sup>th</sup> percentile range in terms of the overall indicators.

Concerning the indicator *government effectiveness*, which measures the quality of the civil service amongst others dimensions of governance, as table 5 shows, the result is an average of 61 in the scale of 0-100 from 2007 to 2014. This is reflected in attempts to improve the quality of officials, for example in the recruitment of more

<sup>174</sup>In 2007, the SHCP attempted to give greater authority and power to the CONEVAL through the creation of another council to replace the CONEVAL but also to undermining the functions of the SFP, which caused friction between the two ministries (see presidential term 2006-12 in the chapter ‘Policy evaluation in Mexico’).

<sup>175</sup>Keeping in mind that the establishment of the civil service in the Mexican government is limited to the executive branch, excluding the legislative and judicial the adoption of the service in the states and municipalities depends upon the interest and support of local political actors.

<sup>176</sup>The Worldwide Governance Indicators measure on six broad dimensions of governance through a six percentile ranges in the scale 0-100<sup>th</sup> of 215 countries over the period 1996-2013, however, here is only included the year 2013 (see methodology chapter).

experienced cadres in evaluation i.e. the CONEVAL's personnel, the ministries (SHCP, SFP) and agencies, whereas in the legislative – at least – there is a will to improve.

Table 5

Government effectiveness (quality of bureaucracy) in Mexico

| Country/year  | 2007      | 2008      | 2009      | 2010      | 2011      | 2012      | 2013      | 2014      |
|---------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| <b>MEXICO</b> | <b>59</b> | <b>61</b> | <b>60</b> | <b>60</b> | <b>63</b> | <b>63</b> | <b>63</b> | <b>61</b> |
| <b>Chile</b>  | <b>87</b> | <b>84</b> | <b>85</b> | <b>87</b> | <b>86</b> | <b>86</b> | <b>86</b> | <b>84</b> |

Source: Data from the World Bank (worldwide governance indicators WGI).

As one deputy<sup>177</sup> suggested, once the re-election of the members of Congress takes place, there will be more experienced cadres in the legislative, but the results will not be observable for the next 10 or 15 years. For instance, concerning the people working as coordinators of the PbR/SED, the evaluation of the SED by the Superior Audit Office of the Federation (2012: 54) points out that 141 officials are responsible for the PbR/SED.<sup>178</sup> Moreover, the analysis of such staff units operating the PbR/SED based on 19 interviews shows that most officials and politicians in the executive and legislative have a masters degree (68%), 10 years or more of experience working in government (73%), and have evaluation experience ranging from 1½ to 18 years. These results corroborate that the outcome regarding the quality of bureaucracy is present and coded “1”.

In spite of this, most of those interviewed emphasize the need for experts in this field including qualified external evaluators. Such experience should be reflected in the decision-making process. However, inside the legislature, the results (chart 1) showed that the existing Congress people is still subject to personal loyalties, patronage networks and lack of expertise, factors that reduce the possibilities to ensure that the skills and experience of actors help on improving the public sector. Indeed, some of those arguments described by Arellano (2000), Guerrero (2007) and Pardo (1999, 2005, 2007) remain as valid in those years [referring to old regime practices] e.g. meritocratic practices, reliability of administration and structure of the system, the old spoils systems returning and lack of transparency amongst others.

For instance, it is widely known that elected public officials grant public positions as political “booty” for their party supporters. An argument is that centralized administrations gain power<sup>179</sup> through the bureaucracy, establishing ostensibly objective rules to appoint new officials to perform professionally. In regimes like this, it is important for the president to keep some power through those officials who have to be loyal to him/her. These flaws co-exist (overlap) with those attempts to establish a formal civil service and might explain the problem in consolidating reform and keeping separately from the power of the State.

This hierarchical and archaic model is a legacy of public administration practices, at least in Mexico, where top bureaucracy is highly well paid and positions are promised and committed even before governments come to power. Evaluation thus involves a clash of views between administrative officials and elected members over the boundary between technical decisions and political ones. Another consideration is the

<sup>177</sup> Interview #16. 25/march/2014.

<sup>178</sup> From the total of 141 officials responsible of the PbR/SED, there are 65 (46.1%) people who work in the CONEVAL, 50 (35.5%) in the SFP and 26 (18.4%) in the SHCP.

<sup>179</sup> As an organ of repression, power represses nature, class, and individuals (Foucault, 1980: 90).



gap on bureaucratic influence over the legislative regarding policy evaluation. By exploring bureaucratic influence in the policy process, it could be affirmed that in contexts where performance and results-based or oriented management reforms are undertaken, there is a positive effect over making bureaucracies more effective (Nicholson-Crotty & Miller, 2012: 348).

However, as explained in the previous section of the “IGR”, the linkages between branches of government operate on the basis of legal authority and the executive bureaucracy influence over legislative process is narrowed until the budgetary process starts. Moreover, in less professionalized agencies – such as the Mexican Congress –, Nicholson-Crotty & Miller (2012: 366) argue that they did “do not find evidence that bureaucratic effectiveness has a larger impact on bureaucratic influence in less professionalized chambers”.

Given these facts, context matters greatly and makes it difficult to manoeuvre a civil service promptly due to political parties, which need those positions to repay favours. Many positions in the hierarchy of government are linked to the governing political party and co-exist with the civil service. In this case, the asymmetries of experience and technical knowledge of the bureaucracies and politicians across the executive and legislature have affected the effectiveness of evaluation, mainly because these asymmetries produce inaction for decision-making once it takes place.

The results seem as underestimated by both branches for further utilisation. The average of 7 years of experience in evaluation of people interviewed is not clearly linked to its effectiveness. This means, there is a need of staffed members with a high level of know-how to operate sophisticated evaluation and monitoring systems such as the PbR/SED, to perform efficiently data gathering, data-entry and analysis of evaluation. Thus, it can be affirmed that the quality of bureaucracy is necessary for achieving effectiveness but closely linked to political factors, at least in Mexico. A more meritocratic civil service plays a central role in making policy evaluation effective and factors explained above remain as potential obstacles to its effectiveness.

### **Level of democracy**

Democracy comprises not only free and fair elections, the right to vote, and freedom of expression; it also refers to the capacity to implement policies effectively within a system of checks and balances (Dahl, 1971: 3 cited in Lijphart, 2002: 46 and EIU democracy index, 2014). As explained in the methodology chapter, to measure this variable the Economist Intelligence Unit (EIU) is included, which is based on the UNDP's report and Latin American Barometer, and adds that concentration of power, social and economic inequalities, insecurity and violence endanger the democracy of Latin America.

To analyse this function and whether policy evaluation is imposed upon or engaged with stakeholders, the *democracy index* has four categories of democratic system:

- 1) Authoritarian regimes
- 2) Hybrid democracy
- 3) Flawed democracy
- 4) Full democracy

In 2014, the democracy index placed Mexico at number 57 in the ranking with an overall score of 6.68, regarded as a flawed democracy. The score of the country has

stayed at the same position since 2006 when the list was first published. Equally important is the result regarding codification of this variable, for which, the absence of a full democratic regime is either equivalent to a code “O” or an outcome not attained (see annex 1).

An example concerning the functioning of government (a category of the democracy index)<sup>180</sup> under a flawed democracy is through day-to-day routines, in which the bureaucracy has significant chances to influence and implement programmes and policies, but also to exert discretionary action for purposes beyond their formal authority. From one perspective, the exercise of discretion could be desirable for promoting democratic governance by bridging gaps between citizens and elected officials (Meyers & Vorsanger, 2007: 157). Conversely, it might undermine or jeopardize desirable goals, the effectiveness of policy and efforts to improve democratic governance.

It should be borne in mind that evaluation becomes more effective in full rather than flawed democracies. Evaluation is not a necessary concomitant of a democratic system, but it could emerge under diverse circumstances and it has been imposed upon different actors over time and the same political arguments. At a first glance, the presidentialism regime uses monitoring systems as a form of control of the organisations and their resources. For example, in Mexico during the mid 80s these were used for strengthening the planning process enacted in the Planning Law and the National Democratic Planning System, in which evaluation of government performance was implicit in the end of the process.

However, those measures were performed as result of the economic crisis, which required the reduction of government structures rather than as a process where evaluation helped to determine agencies' efficiency, for transparency and accountability. Indeed, in those years the 969 out of a total of 1150 state enterprises were downsized (see 1982-88 presidential term in the chapter 'Policy evaluation in Mexico'). Additionally, in those years the evaluation was under the responsibility of several ministries such as the Presidency, National Heritage, Programming and Budget, General Controller of the Federation, Public Administration and of Finance and Public Credit. The priority given to policy (evaluation) depends upon different actors and how much political power they exerted over the public administration.

It is possible that the incentive to fully implement evaluation was ineffective due to lack of political consensus with those assessed and most importantly due to the consequences of bad or poor performance. To push forward democracy in these emergent economies, evaluation becomes a powerful tool to ascertain whether the government's actions are in accordance to those planned. The evaluation process should also serve to create or maintain trust in officials or bureaucracy through accountability and transparency of the public administration operation, with the aim of achieving an effective and efficient performance. This extends to politicians, who gain legitimacy and trust from constituencies via public action and effective evaluation.

Taking this into account, evaluation can be seen as a public matter that concerns both the governors and governed. Some explanations could be offered for the persisting modesty of citizens' demands (Manning, 2001: 302) in terms of service

---

<sup>180</sup> Other categories of this index are electoral process and pluralism, civil liberties, political participation, and political culture.

quality. Therefore, the need of evaluation to acknowledge programmes' performance has been underestimated by constituencies. The evolution of evaluation over time shows it was undertaken by officials and politicians at the central level in the absence of much interest from the citizenry in their governments' performance though with increasing engagement from interest groups over time.

Equally important is that consolidation of democracy brings up the significance of 2000 in the political life of Mexico. In that year, for the first time, a non 'PRI' president came to power alongside the "collapse" of presidentialism. Nonetheless, the figure of the president still remained as the key decision-maker, a fact which helped ensure further improvements in government and public policies. Interest in evaluation increased further due to both the world economic crisis of 2008 and 2009 (see Wiesner, 2011), and the international tendency to promote it and monitoring systems as ways of controlling public resources.

For instance, although the Mexican government had not dismantled its authoritarian structure (Reyna, 2006: 131) in those years, it reinforced the need for a system of checks and balances, amongst others. Therefore, all the innovative actions performed by government were boosted thanks to the legacy of presidentialism exerted by the PRI party during 70 years. As noted by Reyna (2006: 131) 'the Mexican government has not dismantled its authoritarian structure although it has undergone a political change, the public debate is more open and citizens tend to be more participatory'. An example was the Presidential Goals System (SIMEP) during the administration of Fox (2000-06). Its effective implementation was because the executive agenda was focused on it and the president monitored it, even more than the PbR/SED or the SICEG, which used to be the most relevant component of policy evaluation before the PbR/SED was implemented.

Another example of the influence of political actors rather than democratic factors on the development of evaluation can be seen during the administration of Calderon from 2006 to 2012. Although, this president came from the same right party (PAN) and it was assumed that he would continue with most of the public policies implemented by his predecessor, the evaluation took another path due to the lack of consensus and joint action to continue previous policies regarding evaluation. Another example regards a further set of reforms including the New Public Management 'NPM' carried out in those years. Although the NPM was not directly linked to the reforms performed by the Mexican administration (see chapter 'Policy evaluation in Mexico'), it attempted to alter the managerial jargon and administrative practices.

The government moved towards an administrative reform as Arellano & Guerrero (2003: 11) note, "through a 'neutral' one", which could be fully implemented in order to promote policy evaluation and government's efficiency. Thus, the NPM paradigm influenced the adoption of evaluation and specific actions to make it more feasible. For instance, the programme 'Prospera' was pioneered due to its innovative model of impact evaluation. These positive findings enlightening the policy process rewarded the programme, with more financial resources and coverage was increased to benefit more households that eventually influenced the spread of evaluation across the Mexican public administration. Moreover, evaluation proves the importance of its effectiveness, in which the evidence comes as the "objective" authority for further improvements in public policy.

To sum up, it could be assumed that Mexico is still in a transition to democracy, with some gaps in the functioning of government that render it a flawed democracy. Thus, democracy was not determinant of the adoption of evaluation based on the historical context and political circumstances in Mexico but was relevant for its effectiveness. The transformation of the government has been difficult to achieve in the Mexican state due to the presidential control over the governmental apparatus (Arellano & Guerrero, 2003: 2-3), characteristic that prevails in hybrid regimes. Nef (2007: 325) states that “incomplete transition, restricted democracy and the receiver state have had significant effects upon the administrative systems...” and they also reduce the capacity of the State.

However, the fact of being a flawed democracy has not been an obstacle to promote the effectiveness of evaluation from the central level. A comparison of non-democratic and democratic administrations before and after 2000 corroborates an increasing trend of effectiveness after this transitional period. In the later years of the 90s, public organisations only justified programmes' costs, whether or not they produced the expected results from a cost-benefit analysis perspective, as Mejía (2005: 17) argues. Although evaluation was constantly mentioned (see chapter 'Policy evaluation in Mexico') it did not evoke a rapid response from those evaluated to transform recommendations into actions to improve their programmes.

In those years, evaluation was considered as a mechanism of control focused on audits and on the oversight of public resources, whose reports summarize the findings of audits instead of a performance measurement model using a results-based approach. In this century, evaluation gained acceptance inside the government branches, mainly due to political factors and other external factors. An important factor was the pressure exerted by international organisations such as the OECD among others, leading transformation of government but also as a result of the academics demanding the government's accountability, transparency due to the world economic crisis between 2008-09. Later on, the evaluation's effectiveness acquired importance thanks to programmes such as 'Prospera', which demonstrates the importance of policy take-up recommended by external evaluators.

Regardless, it could not achieve effectiveness due to narrow demand-led by those interested, especially by constituency. The challenge might be is to design new institutions, strengthening capacity-building and the role of the state, rather than focusing on policies, as Hague & Harrop (2004: 319) suggest “in established democracies, institutions process new policies; in new democracies, the policy is to develop new institutions”. Currently, the risk is that the Mexican administration adopts a *status-quo* approach and retrograde tendencies regarding the system of checks and balances prevail, with consequences for the effectiveness of policy evaluation.

### **Policy evaluation framework**

Implementation of complex policy initiatives is challenging in various ways. It requires effective coordination, network management, regulation, and information systems that reliably update stakeholders and policy implementers of the status of programmes' outputs. To measure the balance between design and operationalization and how these impact the effectiveness of evaluation, the variable is divided in two with some sub-variables:

- 1) Coherence in the design:
  - 1.1) Matching policy evaluation objectives and outcomes.
  - 1.2) Overlap with regulations.
  - 1.3) Holding hearing of interest groups (consensus between actors).
- 2) Structure of monitoring and evaluation system:
  - 2.1) Technical operationalization (easy/difficult).
  - 2.2) Perception of operationalization.
  - 2.3) Methodological rigor.

As mentioned in the methodology chapter, this variable is related to the formulation stage of the policy process and has an effect on the effectiveness of evaluation. Indeed, evaluation is a source for 'objective' decision-making to lead to changes in policies/programmes. However, objective decisions could be erroneous because evaluation is a product of the design of policy-making created by human beings' values, which are subjective (Bobrow & Dryzek, 1987 cited in Parsons, 1995: 564 and Miller, 1984). Thus, an ill-designed policy leads to unexpected or non-wanted outcomes, as Knill & Tosun (2012: 162) conclude, "distributional conflicts might lead to the formulation of ill-designed policies that are characterized by inaccurate assumptions about the causal relationship between policy problems and politically adopted remedies". In this sense, all outcomes of this variable are absent and coded "0" due to the following reasons.

Ideally, *coherence* in evaluation design should strike a balance between gathering and obtaining relevant information on the one hand and enabling the evaluated to carry out their tasks on the other. Firstly, an examination of the survey and observational data indicates that policy evaluation's objectives and how these match with outcomes as established in regulations have different perspectives which are desirable but do not guarantee *coherence* in the design of policy.

On the one hand, at the macro level the PbR/SED strategy shows coherence with the National Development Plan (see PND 2013–18), which also applies at the meso level where the ministries and CONEVAL are vertically integrated to sectors, regions and municipalities through various regulations. On the other hand, scaling up to policy evaluation as a comprehensive model, the scenario changes drastically at both horizontal and vertical level. At the micro level there is no correlation between individual (officials) performance and major goals established in programmes, at the institutional and national levels, nor with regulations, evaluation and monitoring systems coordinated by other public organisations.

Given this context, regulations essentially provide a legal framework for implementation, avoiding discretionary judgments by policy implementers, but also risk making the evaluation process difficult to monitor. As Simon (1983 cited by Meyers & Vorsanger, 2007: 155) argues, "political efforts to control discretion through the promulgation of detailed rules and procedures often produce the contrary result..." Looking at evaluation as a fully integrated system, these systems are not harmonized and there is some overlapping. It appears that the evaluation process is characterized by a fragmentation of regulations.

Secondly, the SFP executes a myriad of monitoring and evaluation systems such as the budgetary-programmatic evaluations every three months, coordinates the performance evaluation subsystem of officials (Career Civil Service), the performance agreements of the state enterprises jointly with the SHCP, the performance model of

control and oversight bodies (MIDO), and the system for government control and evaluation (SICEG). Certainly, all these systems that belong to the SFP are in the waiting list for relocation (see chapter 'Policy evaluation in Mexico'). Thirdly, the evaluations of the social sector coordinated by the CONEVAL, which results are linked to the PbR/SED are more monitored by Congress than evaluations of other sectors.

In the fourth place, the existing evaluation of the education sector is led by the National Institute for Evaluation of Education (Instituto Nacional para la Evaluación de la Educación 'INEE'), a decentralized agency of the Ministry of Education for measuring students' performance<sup>181</sup> and other educational issues is not connected to the main strategy (PbR/SED). Neither the environmental impact evaluation performed by the Ministry of Environment and Natural Resources (SEMARNAT), whose aim is to be a preventive tool of the environment's sustainability.

Aside from that, since 2009 the National Statistics Office (INEGI) has been producing data on government performance – census of local government's structure and indicators on public safety and administration of justice – but without any linkage to the PbR/SED. Another actors performing evaluation & monitoring activities are the National Institute of Climate Change and Ecology (Instituto Nacional de Ecología y Cambio Climático 'INECC',<sup>182</sup> which was created in 2012 through the enacted General Law of Climate Change.

The most recent evaluations performed were introduced by the Ministry of Energy (Secretaría de Energía 'SENER'). In the manual of procedures of the General Office of Social Impact and Surface Occupied enacted by the end of year 2016,<sup>183</sup> it was established the obligation to perform evaluations of social impact, and consultation of indigenous people and communities, amongst others. Lastly, another actor in the legislative branch is the ASF, which has recently undertaken evaluations<sup>184</sup> and aims "To assess goals, objectives set in federal programmes on the basis of strategic indicators approved in the budget to verify their performance and legality of public resources"<sup>185</sup> (see annex 4).

Evaluation thus has been implemented in diverse political scenarios embedded in a variety of monitoring and assessment activities. The government's efforts to establish evaluation through the PbR/SED have produced the opposite of a comprehensive policy, leaving aside strategic political actors to increase its effectiveness such as the INEE, INECC, SENER, NGOs or ASF. Indeed, the latter exerts similar functions in terms of evaluation and its regulation seems to overlap with the PbR/SED in a way that is financially inefficient.

In spite of this, there is no indication that results-based budgeting attempts to include these outputs as it should. This is mainly because there is some key data developed by these agencies waiting for stakeholders and politicians to use it for further improvement in government action. Regarding the structure of the PbR/SED as a system it has a complex design. Firstly, the mechanism to follow-up evaluation

<sup>181</sup> E.g. Programme for International Student Assessment (PISA) coordinated by the OECD.

<sup>182</sup> Official Journal, 06/06/2012.

<sup>183</sup> Manual of procedures of the General Office of Social Impact and Surface Occupied, 24/11/2016.

<sup>184</sup> Evaluation of the SED in 2012.

<sup>185</sup> It is enacted by the Audit and Accountability Law of the Federation [Ley de Fiscalización y Rendición de Cuentas de la Federación], Official Journal 18/06/2010.

results operates as a separate module, coordinated by the CONEVAL, SHCP and SFP. In addition, the same recommendations are registered into the Programme for a Close and Modern Government (Programa para un Gobierno Cercano y Moderno 'PGCM'), which is coordinated only by the SFP. This procedure duplicates activities to those programmes' implementers across federal public administration<sup>186</sup> because they have to report the same information in different formats and units inside the same ministry (SFP).

Secondly, the system underlines – in the discourse – the need for results over procedures but the MIR's methodology is based on the 'logical framework approach', which requires a set of indicators at four levels: impact, results, services and activities, emphasizing inputs rather than results, increasing the complexity of the evaluation system.<sup>187</sup> Thirdly, according to the OECD study,<sup>188</sup> the capacity and functions of authorities such as the SHCP and SFP are perceived similarly and even overlapping in operation:

Mexico is unusual among OECD countries in having three institutions with overlapping roles and responsibilities for these reforms: the Ministry of Finance, the Ministry of Public Administration, and CONEVAL... In most OECD countries, only one institution has the lead responsibility, although that institution varies by country (OECD, 2009b: 116).

In regard to the coordination of evaluation, the review of roles across branches shows that it is complicated to identify their sphere of responsibilities and scope of action. Yet, these ministries have not clearly defined the number of budgetary programmes under supervision of each one. Only the CONEVAL, which is responsible of the social sector, recently published a complete catalogue of the social programmes under its supervision.<sup>189</sup> Notwithstanding this responsibility it is assumed that the social sector should include the entire health sector's programmes reviewed here. However, the council is only responsible for a few of them while the others are coordinated by the SHCP such as the programmes *Construction of medical specialized centers* or *Reduction of maternal mortality*, which should be considered as key programmes of the social sector.

Lastly, the involvement of interest groups in the design of the policy framework is the usual form of any inclusive policy process in a democratic government. Interest groups include NGOs and target groups whose behaviour is to alter government action (Wu et al., 2010: 69) as policy implementers through using public hearings statutorily defined, along with political, administrative and judicial activities to represent diversity of interests and expertise (Wu et al., 2010: 68) and thereby to ensure a flexible and feasible policy framework. Moreover, it guarantees implementation of the policy over public administration according to the proposal put forward by the executive. Far from consultative devices, however, the actual practice analysed here is an authoritative decision by those implementers along with the executive and legislative branches avoiding negotiations, discussions, empathy with those affected, and compromises with powerful groups.

---

<sup>186</sup> According to the "Programa para un Gobierno Cercano y Moderno 2013-18", evaluations and recommendations are perceived as a bureaucratic procedure rather than actions to improve the programme. Official journal, 30/08/2013.

<sup>187</sup> The utilisation of indicators, which represent the most objective technical criteria feasible for evaluation has faced some difficulties in their design due to the complexity of programmes. In some cases, such as the Ministry of External Affairs, the results are not tangible outcomes, instead consisting of political negotiations or agreements between such ministry and other agencies. Thus, the outcomes are inputs of another process even of another public agencies.

<sup>188</sup> The OECD issued the Review of Budgeting in Mexico in 2009 but this study was presented by the Presidency Office on January 7th, 2010.

<sup>189</sup> In 2015, according to the CONEVAL, there were 235 social programmes under its coordination.

The second variable refers to the structure of the monitoring and evaluation system in which ideally should prevail a coherent and easy operationalization by users. However, officials and politicians surveyed about this, point out to a difficult, rigid and unfriendly operationalization system. They also refer to the problem of operationalization by policy implementers to the myriad of systems to report indicators and evaluation results e.g. the Performance Evaluation System (SED), the mechanism to follow-up evaluation results (SAS), the Integral Data System of Revenue and Public Expenditure [Sistema Integral de Información de los Ingresos y Gasto Público 'SIIWEB'], the Synthetic Model of Performance Data (MSD) and the "Portal Aplicativo" (PASH).

In theory, the SED should be the only database system as conceptualized in regulations but in practice, the range of systems can be considered as the side effects of poor design, in which implementers were not able to develop a major data system to manage the vast amount of information produced by diverse agencies. This brings practical difficulties for technical operationalization and monitoring purposes, which are different from those objectives established in the Results-based Management<sup>190</sup> (RBM), underlining results rather than procedures. The review of the SED by the ASF (2012: 79) confirms the position of this research about the complexity to follow recommendations once these are published in the mechanism<sup>191</sup> by programme's operators.

For example, as an external user it was encountered some difficulties in finding data in the health sector. As mentioned in the first section, there is no data available for two programmes *Reinforcement of health provision networks* and *Improvement of medical units*. In other cases, the results of evaluations are only published in the CONEVAL's website and not available in the Ministry of Health's website. In more specific cases, the recommendations do not appear in the mechanism to follow-up evaluation results such as in the EED 2010 (specific evaluation of performance) of the programme *IMSS-Oportunidades*. In other cases, recommendations were reformulated by agencies and different from those reported by evaluators such as in the EED 2012 of the programme *of people with disability* as well as in the process evaluation 2012 of the programme *of childcare centers*.

Data collection through these systems is challenging and not easily available at central level. Complexity increases when data of budgetary-programmatic evaluations, performance agreements, evaluation of officials and economic evaluation of incomes and expenses refers to because information is not open to society via the official websites. These results suggest that internal conflict is present mainly because different authorities attempt to put different systems into operation without consensus. Based on the idea that data means power, those to whom database belongs to are sufficiently powerful to exert control over the rest and to restrict the accountability of government performance in a discretionary way. Policy evaluation determines not only what will be done or provided to whom, but also the resources, regulations and data system that stakeholders should have at their disposal.

At the time of the study, the implementation of policy evaluation framework was making slow progress toward balancing the system, the scheme of regulations with

<sup>190</sup> See Results-Based Management handbook, United Nations Development Group, available at: <http://www.undg.org/docs/12316/UNDG-RBM%20Handbook-2012.pdf>

<sup>191</sup> The mechanism to follow-up evaluation results and/or "Synthetic Model of Performance Data".



the modernization processes undertaken by the public sector characterized by traditional legalistic obligations and complexity of systems. Taken together, all these factors have hindered the effectiveness of policy evaluation.

### **Degree of autonomy**

As explained in the methodology chapter, in general terms, autonomy occurs when a public agency can decide by itself which matters are considered the most important (Verhoest et al., 2004: 14) and decisions concerning policy-making are independent and not subject to influence by external factors and actors. The emergence of “independent” agencies in Mexico such as the CONEVAL or INEE represents the government’s commitment to build credibility in its institutions. This is based on the premise that relative independence of the evaluator will have an impact on objectivity, transparency and accountability of results obtained, as well as amongst scholars, citizens and even officials.

In the context of evaluation, autonomy implies the independence to decide which programmes will be evaluated, the type of evaluation required without the coordinator’s approval, and the delivery times of results from external evaluators. It also refers to self-ruling separately from the top of the power (government). Ideally, it seems more a bottom-up rather than top-down approach (Knill & Tosun, 2012: 155), where implementers have more flexibility and autonomy to adjust policies based on evidence. Thus, to rank an agency’s autonomy, this variable is divided into the following three dimensions:

- 1) Policy autonomy: low and high.
- 2) Financial autonomy: low and high.
- 3) Legal autonomy: low and high.

Here, there are two approaches selected regarding autonomy in evaluation. The first one refers to the real scenario and political environment where evaluation takes place. For Karlsson & Conner (2006: 232-4), this is the value-sensitive evaluator position, in which the evaluator should not stay separated from politics at all and must take responsibility for the technical aspects of how the evaluation is done, and its ethical aspects and democratic values. Following this idea, the evaluator should consider countries’ issues, such as Mexico’s hierarchical bureaucratic structures, where the autonomy of the agencies does not guarantee homogeneity of the public policies or programmes implemented. This might appear to be the key reason for governments to move to more centralized control and evaluation systems.

Although these moves could hinder any attempt to empowering interagency relationships, and could be contradictory to managerial movements such as the NPM (Barzelay, 2001, Hood, 1995 and Radin, 2007), which emphasizes the decentralization of systems, disaggregation of units in the public sector and inclusion of private sector forms of management, amongst other things, centralization is the key for operationalization. In Mexico, while the CONEVAL and INEE have attempted to be autonomous it is not sufficient for the purpose of autonomy. Certainly, in 2007, two initiatives were put forward for public debate to replace the CONEVAL<sup>192</sup> in order to grant greater functions, authority and most important autonomy to the agency coordinating the policy evaluation (see 2006-12 presidential term in the chapter ‘Policy evaluation in Mexico’). However, it is important to bear in mind that

---

<sup>192</sup> The creation of the National Evaluation Council for Public Policies and the National Evaluation Council of Government Management, the former submitted by the SHCP and the latter by a deputy at Congress.

the CONEVAL is only responsible<sup>193</sup> for approximately 20% of the total of 960 budgetary programmes.<sup>194</sup>

In this context, the attempt to give “more” autonomy to agencies is characterized by the level of decision-making competencies (Verhoest et al., 2004: 104). According to these authors’ premises (see methodology chapter), the SHCP, SFP, CONEVAL or INEE have achieved management and policy autonomy. The former refers to aspects of management decisions and the way they are taken, whereas the latter is related to individual decisions of general rules or policy instruments. However, in spite of recent reforms to the Constitution<sup>195</sup> yielding to the CONEVAL an own legal personality (autonomy), this still operates under supervision of the SHCP, and its partly funded by central government (financial autonomy), and does not decide on individual applications of general rules (policy).

In this regard, while the legal initiatives discussed in Congress in 2007 (see chapter ‘Policy evaluation in Mexico’) aimed to increase the leading role of the CONEVAL over the PbR/SED, they had not produced the results expected and its autonomy is still limited as mentioned in the previous chapter. As for the INEE, its autonomy’s status is alike to the CONEVAL in terms of operating under supervision of the Ministry of Education despite being a public, decentralized agency with financial, technical, and operational autonomy. Keeping in mind that financial autonomy is self-sufficient to operate resources, the budget remains as a tool to impose control over inter and intra governmental relationships, as Rubin (2010: 346) states “budgetary decision making provides a picture of the relative power of budget actors within and between branches of government”.

Thus, a discretionary role prevails in Mexico in spite of the narrative that autonomy is key for the evaluation process. This statement is supported by interviewees’ responses, which mostly point out to autonomy as a condition needed rather than actually existing in the evaluation. According to officials interviewed who refer to autonomy as something essential that it has been achieved over time, but definitely not independent from the executive branch, as one informant<sup>196</sup> said, “autonomy is [only] the agencies’ decision to hire an external evaluator to evaluate programmes”. Moreover, they consider the SHCP necessary to lead evaluation and standardization due to its technical experience. One interviewee<sup>197</sup> mentioned “control of the SHCP is important to operate evaluation, to following up recommendations by public organisations” and “if it (SHCP) was autonomous it should not has power over programmes to make changes...”

This perception is contradictory to that expressed by politicians from the legislature. For them, autonomy is still something aspired to, preferably separately from the executive, as an interviewee<sup>198</sup> suggested, “it should be the fourth branch power”. Indeed, another politician<sup>199</sup> surveyed affirms that autonomy in Mexico is a “myth” and should be the exception not the rule as nowadays “they (autonomous agencies) should not longer be employees of the executive”. For these reasons, the evaluation

<sup>193</sup> In 2015, according to the CONEVAL there were 235 social programmes.

<sup>194</sup> Figures are based on data provided by the SHCP in 2013, available at: [www.transparenciapresupuestaria.gob.mx](http://www.transparenciapresupuestaria.gob.mx)

<sup>195</sup> Official Journal, 10/02/2014.

<sup>196</sup> Interview #7. 15/march/2014.

<sup>197</sup> Interview #2. 25/february/2014.

<sup>198</sup> Interview #15.25/march/2014.

<sup>199</sup> Interview #17. 27/march/2014.

agency and agencies coordinating the policy evaluation have a minimum or low level of autonomy in all dimensions (policy, financial and legal) and consequently assigned a code “O” as annex 1 shows.

Another perspective in the debate on autonomy is the value-neutral evaluator position outlined by Karlsson & Conner (2006: 232), who recommends keeping evaluation away from political influences.<sup>200</sup> The aim of this position is to emphasize the autonomy of evaluation in relation to political institutions and powerful interests in society. In the end, politics is power, and power could be a source of corruption, which jeopardizes objectivity, impartiality, and rationality of facts. Under this approach, the relative independence and credibility of the evaluator should be located in the legislative branch or through an agency operating independently from the executive.

The legislature is representative of constituencies’ demands and must be the oversight body of government’s (executive and judicial) actions. Furthermore, the focus on the legislature leaves aside the potential role of interest groups and organisations influencing public policy on behalf of constituencies such as the NGOs. Unlike political parties, these pressure groups have autonomy from the government “their work quietly, their activity is nonetheless pervasive in established democracies” (Hague & Harrop, 2004: 166). They are necessary, especially in activities related to budgeting, adopting a pragmatic view questioning what government does or proposing alternatives in a predictable way. Regarding results-based budgeting, they should operate as a society’s watchdog of the State’s activity.

Since the PbR/SED has developed, some NGOs have been involved following the budget’s outputs. “Curul 501”, “Borde político” and “Visión legislativa” represent such pluralism exerting strong influence over a responsive government (Hague & Harrop, 2004: 177), helping to link the space between society and the State in matters of public expenditure,<sup>201</sup> and representing a strong pressure where decisions are formed. Indeed, in recent years these NGOs have been upheld by the SHCP in order to deploy some pressure over Congress to use evaluation.

Notwithstanding the relevance of autonomy by the agency performing evaluations and coordinating the rest of the public administration, it becomes necessary to exert some control and political power over the rest at least in contexts such as the Mexican. At this point, it is important to note that autonomy is considered as a proxy of leadership and power due to the complexity for measuring them separately, whereas control is the inverse of autonomy (Verhoest et al., 2004: 106), a scenario in which the government could take control of agency behaviour to influence its decisions. Thus, it is not enough having agency’s autonomy in regimes with hierarchical bureaucracies such as Mexico. This statement is similar to that pointed out by Cabrero et al. (2003), where public agencies are developed in a particular context, have their own patterns of interaction and power struggles, and certain traditions, myths and organisational routines arise within them, each one with different meaning.

---

<sup>200</sup> In terms of the selection of the evaluator, funding, information given to the evaluator, scope or approach, methods, subject pool selection, instruments used, data analyses, interpretation of data, recommendations and information disseminated (Turpin, 1989: 55 cited in Karlsson & Conner, 2006: 232)

<sup>201</sup> For this, they developed a website called “Open Budget” (Presupuesto Abierto) to following up the budgeting process. Available at: <http://www.presupuestoabierto.mx/>

For example, while policy makers on the executive consider evaluations of good quality and reliable, for politicians' evaluations are perceived as being of dubious reliability because the executive coordinates them. Congress people interviewed (42% of the total) confirmed they did not consider evaluations trustworthy although these are performed by autonomous (external) evaluators e.g. academics, NGOs. Essentially, these are strong reasons to classify this factor at the end of the variables hindering the effectiveness of policy evaluation and for explaining evaluation use. In this sense, the assumption that credibility, objectivity, transparency and accountability of evaluation results will depend upon the extent to which the agency that coordinates evaluation is operationally independent from a hierarchical supervisory body is valid depending on where one sits.

Furthermore, the effectiveness is more engaged to the political arenas where decision-makers, the intragovernmental relations, power and political control are determinant for success. Lastly, in 2015 there was a change on the way the budget was designed, and the executive proposed the zero-based budget for the FY 2016. If the Ministry of Finance (SHCP) wants to be innovative, the evaluation should be included as enacted by law (LFPRH) utilising the tool to support decision-making on allocation of public resources. However, the results of this budget will not be included in the analysis due to limitations of time. The analysis of institutional and political factors examined above takes the discussion to the need of a more detailed comparison of variables across countries to determine which factors are the most relevant and the circumstances surrounding the effectiveness of evaluation.

## **CHAPTER 6. The effectiveness of policy evaluation in Chile**

### **Introduction**

This chapter focuses on the analysis of the Chilean policy evaluation and its effectiveness. The methodology utilised for this purpose, which was already mentioned in previous chapters is based mainly on the federal level as well as it was in the chapter about Mexico. Here indeed, the strategy was developed by the central government and there is no evidence about whether municipalities are implementing some sort of evaluations at local level. The starting point is the explanation of the budgetary process and how the federal budget is made on the basis that evaluations are performed under the results-based budgeting as the core of policy evaluation.

The second part describes in detail whether those programmes achieved effectiveness once they have been assessed based on whether [or not] changes are classified as amendments, status quo, learning process, cancelation or no data available. It is important to bear in mind that the revision is limited to the health sector programmes from 2007 till 2014 as explained in the methodology section in order to have the same point of comparison with those of Mexico. Both countries combine performance monitoring indicators and in-depth evaluations. One of the assumptions to be analysed in the following sections is that presumably the policy evaluation has a better performance when there is expertise in the field. For this, the discussion not only focuses on the results but also on the political factors directly related to the success or failure of evaluation, this is, its effectiveness.

In this context, the study of the DIPRES (Hawkesworth et al., 2013: 7) between 2000-09 found that 7% of evaluations led to the termination of programmes, while the 91% of those results led to some modifications such as the programme redesign, minor changes and reassignment of the programmes to other agencies (Arenas & Berner, 2010: 69 cited in Hawkesworth et al., 2013: 7). Thus, the following sections will compare such results. Some references to the quality of evaluations are mentioned but not analysed. Alongside this, the review of budget discussion at Congress from 2007 to 2014 is presented, mainly to verify whether those recommendations made as a result of evaluations are used for decision-making and such programme improvement could be addressed to evaluation.

Lastly, this chapter will discuss the level of influence of each political/institutional factor such as intra governmental relations, quality of bureaucracy, level of democracy, policy evaluation framework and degree of autonomy in favour of effectiveness. Also, the chapter brings key aspects of the preceding chapter 'Policy evaluation in Chile' (historical perspective), which contribute to analyse such variables. The analysis is mixed with key data of informal interviews with politicians and officials<sup>202</sup> to shed light on current practices of decision-making regarding the effectiveness of evaluation. Concerning this, the number of interviewees is low due to limited access to conduct them face to face. Some conclusions of this chapter are the noticeable period of time when bureaucracy led the policy evaluation to an effective process positioning the country operating a M&E system at the forefront of the international scenario. However, the results over time evidence some asymmetries of the institutional factors reviewed here jeopardizing this leading position.

---

<sup>202</sup>The names of the people interviewed were changed in order to keep confidentiality of personal data.

## **The budget process: “Results-based Budgeting”**

To start the discussion, it is important to bear in mind that the health sector represents a major area of public resources inside the Chilean government. As mentioned in the methodology chapter, in 2013, this sector represented 7.5% of the total health expenditure as a share of the Gross Domestic Product (GDP), whereas the media of the OECD countries was around 9.3%.<sup>203</sup> In the Budget Law of the FY 2015, the amount was of \$6,128 millions of cl pesos equivalent to \$9.7 millions of us dollars<sup>204</sup> (including health agencies of the Army) representing approximately the 18% of the total budget (\$34,788 millions of cl pesos equivalent to \$55.4 millions usd).

Indeed, the draft budget law 2014 states that of every two (Chilean) pesos of increase in the budget, one goes to the health sector,<sup>205</sup> and during the budget law 2014 and 2015, this sector represented the third largest source of the budget, which was less than the education and work sector (see DIPRES). These figures demonstrate the priority of the sector in the budget, which comprises the National System of Health Services and the Health Services to the Armed Forces. In terms of programmes, the sector offers a large number of services. However, in contrast to the Mexican case, in Chile there is no evidence about the total of programmes currently being operated by the government.<sup>206</sup>

Regarding the budgetary process, it is important to explain how the federal budget is made in Chile because similarly to Mexico and other countries evaluation is linked to the results-based budgeting (see figure 9). In Chile, the MoF formulates the draft budget law in the period between July and September from the previous fiscal year and sends it to Congress to be enacted no later than September 30<sup>th</sup>. Once the draft budget arrives to Congress, it is revised by the Special Budget Committee<sup>207</sup> and turned to the Lower Chamber (Chamber of Deputies). The preliminary version is sent to the Senate (Upper Chamber) for discussion and approval. Once both chambers approve it, this is passed to the executive and the president enacts it as the Budget “Law” by Constitution (art. 72). The promulgation of the draft by the president should not exceed 60 days from the moment when it first arrives to Congress.

Regarding the results of evaluations of the so-called “public programmes” should feed the budgetary process inside the government in two ways. Firstly, as an input for the tailoring of the budget performed by the executive (MoF), process in which evaluations are needed to identify the performance of each public programmes and agencies to decide which of them needs more resources or at contrary, these could be reduced due to poor performance. Secondly, during the revision by Congress, who also have priorities in terms of citizenry demands and sector performance e.g. education, health, security sectors.

---

<sup>203</sup> Health at a Glance 2013: OECD.

<sup>204</sup> On June 2015, the exchange rate was of \$627.64 clpesos per \$1 us dollar.

<sup>205</sup> Initiative of draft budget law 2013 of 30/09/2013.

<sup>206</sup> The number of programmes currently operated by the government was requested to the Ministry of Health (via transparency Law) However, the response (office letter no. 1606, 28/10/2014) referred to the website already reviewed, which does not show the exact amount of programmes.

<sup>207</sup> According to Congress Law (art. 19) both finances commission of low and upper Chamber integrate the committee; the former has 13 deputies included the president of the commission, while the latter has 5 deputies members.

Figure 9



Source: Data from the DIPRES, author's own elaboration.

### **The effectiveness of evaluation in the health sector**

According to the budgetary process, between 2007 and 2014, there were 54 programmes of the health sector in operation. During that time, 13 programmes had been assessed, that is to say, the 24% of programmes of the health sector were evaluated (see chart 3). Unlike in Mexico, in this case, each programme had performed only one evaluation from the different five types available. As shown in chart below, by looking up at the *amendments* in such programmes it is indicated that 4 of the 13 programmes evaluated took these results for further improvements on them. For instance, the programme *tratamiento de la obesidad* (treatment of obesity) under responsibility of the FONASA, Undersecretary of Public Health and Undersecretary of Healthcare Network had performed an evaluation of public programmes (EPG). This process resulted in 7 recommendations suggested by evaluators, 4 of them were implemented while 3 were not even mentioned in the report to following up results.

Another examples of these amendments are the programmes *comisión de medicina preventiva e invalidez* (commission of preventive medicine and disability) and *programa de prevención y control del VIH/SIDA* (prevention and control of HIV/AIDS) both operated by the Undersecretary of Public Health. In the first case, the commission's aim is to guarantee access (beneficiaries) to healthcare protection. In 2008, a comprehensive spending review (ECG) was performed from which most (15) of the 21 recommendations suggested by evaluators were implemented. Whereas the HIV programme's aim is to reduce morbidity, transmission, discrimination and mortality of diseases related to HIV-AIDS. Derived from the evaluation of public programmes (EPG) in 2010, there were implemented 10 of 12 recommendations proposed by evaluators.

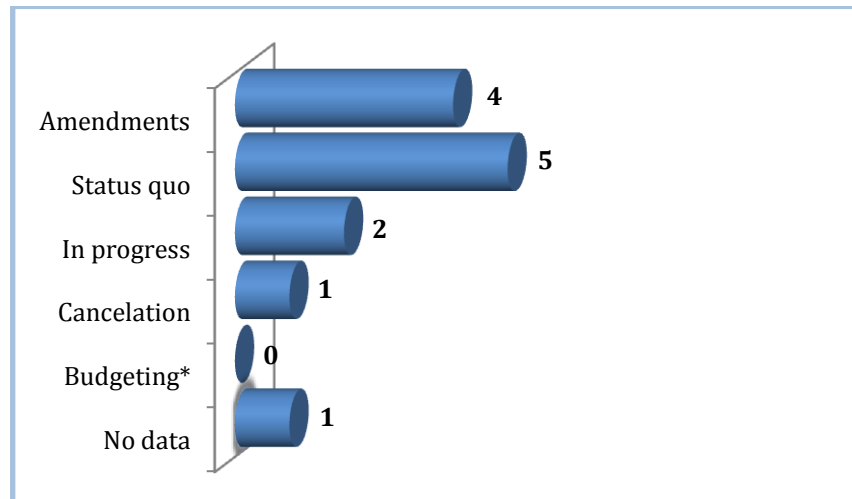
Continuing with the review of categories, from the 13 programmes assessed there are 5 classified in the *status quo*. As mentioned in the methodology chapter, a programme is placed in the *status quo* category when recommendations taken for further improvement are less than 30% of the total suggested. For instance, the programme's National Health Fund's (FONASA), which aims at optimizing the health finances sector and reinforcing auditing activities, in 2011 performed a comprehensive spending review (ECG). The report issued by FONASA after evaluation does not mention specific actions as a result of 17 recommendations suggested by evaluators (see table 6 annexed).



Another example is the programme *campaña de invierno* (winter campaign), which until 2007 was run by the Undersecretary of Public Health but in 2008 was reallocated to the Undersecretary of Healthcare Networks within the same ministry. There was no evaluation performed in those years to evidence that such reorganisation was due to it. This means that policy-makers often made decisions without any support by objective data.

Chart 3

#### Effectiveness of evaluation in the Chilean health sector 2007-2014



Source: data from the DIPRES.

\*Programmes selected for budgeting process belong either to the amendments or *status quo* categories.

The classification of "no data" refers to those programmes in which there is no information available about them or any action taken by policy-makers. Thus, it is not possible to track any recommendation neither the evaluation.

Regarding the *elimination* category, only the programme *de apoyo al recién nacido* (newborn health support) was canceled and merged into the programme *de apoyo biopsicosocial* (of bio psychosocial support) as result of evaluation.<sup>208</sup> However, the latter programme was not included in the budget 2007-2014 (see table 6 annexed) and the former still had funds in the budget law 2014. The condition of this programme is similar to Zaltsman's (2009: 460) interviewees of DIPRES who alluded that "political considerations make it usually very difficult to terminate programs or to cut down their budgets", even in programmes with negative evaluations.

A reason why termination of a programme/policy is rare<sup>209</sup> might be due to the tendency of survival inherent in those agencies known as "dynamic conservatism", in which, administrators act as "antitermination coalitions" to preserve organisational routines (de Leon, 1978 cited on Dahler-Larsen, 2000: 82) because they [evaluations] threaten the stability of administrators (Vedung, 1997 cited on Dahler-Larsen, 2000: 84). Yet, de Leon adds that termination is always a political decision, which becomes more difficult when it refers to a comprehensive policy (sector) than a specific programme.

<sup>208</sup> See Ministry of health's response of the evaluation of public programme (EPG) in 2011. Available at: [http://www.dipres.gob.cl/595/articles-76586\\_doc\\_pdf.pdf](http://www.dipres.gob.cl/595/articles-76586_doc_pdf.pdf)

<sup>209</sup> In his study-survey of 275 municipalities in Denmark about evaluation practices by external consultants, de Leon shows that adjustments of programmes are near to 80% whereas termination rates are extremely low (near 1%).



Moreover, the tracking of health sector programmes indicates that some were cancelled without evidence of the reasons of termination. For instance, the *Programa de desarrollo indígena* MIDEPLAN-BID (programme of indigenous development) and another classified as “others”, which belong to the National Health Fund (FONASA), only come into sight in the budget for two and three years respectively and silently disappeared afterwards. Other examples are the programmes *Secretaría y administración general hacienda* (finances administration) and *Portal integrado de comercio exterior* SICEX (international trade website) of the Institute of Public Health, which only operated during 2011. To sum up, from the total of 54 programmes of the health sector, 12 were cancelled during the period 2007-14 without any evaluation or supporting argument for the action taken.

Concerning the total of programmes assessed, only the programme *red de urgencia* (emergency network) is classified as *no data*, which means there is no information of the programme operation, neither of its evaluation. According to the office letter of the DIPRES (see annex 3) in 2013, the programme run by the Undersecretary of Healthcare Networks should have performed an evaluation of organisational spending (or comprehensive spending review ‘ECI’). However, there is no such programme in the budget 2007-2014, neither data of evaluation results in the DIPRES’ webpage. In this, there is a lack of transparency by operators to mention either the reasons why evaluation had not been performed or to publicize its results.

The following programmes are classified as *in progress*. In the first place, the three recommendations of the programme *bono AUGE* are still on the waiting list and despite the operators agreed with these suggestions and are working on them there is no evidence of action taken. These refer to the need of: i) improvement of the system to register recipients and patient queues; ii) improvement of mechanisms of coordination by interconnecting diverse health services (interagency coordination); and iii) improvement of the healthcare services educating recipients about the services provided.

In the same situation is the programme *training of medical specialists and subspecialists*, which has six recommendations in the process to be implemented. These refer to: i) the identification of the gap between both programmes for better planning in the medium-term; ii) the redesign of the programme components; iii) the development of a mechanism to link supply to demand; iv) the development of a cost study of the programme; v) the strengthening of the monitoring and control mechanisms; and vi) the establishment of a division between roles and responsibilities of both programmes. The agency only issued the evaluation report but there is no evidence about any action taken by operators.

Besides these significant results in the effectiveness of evaluation, there are some inconsistencies between programmes evaluated and those pointed out in the budget. For instance, the programmes *tratamiento de la obesidad* (treatment of obesity), *comisión de medicina preventiva e invalidez* (commission of preventive medicine and disability), *compras a privados: programa prestaciones valoradas* (purchase from private sector), *programa de formación de especialistas y subespecialistas* (training of medical specialists and subspecialist) and *programa de prevención y control del VIH sida* (prevention and control of VIH) have developed an evaluation. However, none of them are classified in the budget as a specific programme.

Moreover, it is difficult to track variations on programmes due to evaluations because there is no such level of disaggregation in the budget. Unlike Mexico, in Chile it is not possible to compare the draft budget – per programme – prepared by the executive and changes made by Congress before being approved. This is similar to Zaltsman's (2009: 461) statement,

And, according to most of the interviewees, such recommendations exert great influence on the president's budget bill and, thereby, on the national budget law as well. Nevertheless, none of these effects could be detected in the agencies' budget documents.

The fact that the budget is quite aggregate in terms of programmes,<sup>210</sup> seldom provides strong arguments for establishing causal links between branches of government and institutional arrangements (formal and informal) such as power, party systems or control over budget. Some examples of these inconsistencies can be found in the programmes coordinated by the health care centers, which in some cases they are not the same than those established in the budget as it was previously mentioned. According to the formal request made through the transparency Law to the health care center of the region of Maule of the total of programmes coordinated by it, this argued of having 31 programmes (see annex 5) ordered by the Ministry of Health. From this total, 23 are different from those stated in the budget such as the programmes of solid residues (residential, hazardous and non-hazardous), potable water, dental health, and prevention of tuberculosis, of cardiovascular disease, of cancer, among others.

Likewise, in the health care center of Osorno 20 of a total of 27 programmes are different from those mentioned in the budget (see annex 6). The reason for this might be that there are programmes operated only at the local level (state and municipalities). Furthermore, it was requested to the Ministry of Health – via transparency Law – about how many and which are the programmes comprised in the health sector and the amount of money spent by DIPRES to coordinate such evaluations. However, during the time this study was taking place it had not replied yet. Apart from the budgetary discussion and regarding the ex-ante evaluation, there is no data (results) about them at least in the health sector, in spite of a DIPRES' top official presented some results during the annual meeting in the OECD,<sup>211</sup> saying that around 88 of non-social programmes have performed an ex-ante evaluation during 2008-2013. Thus, the focus of this type of evaluation might be of the non-social programmes but it is not clearly defined in the methodology.

Differences could be also found in the programme *operational contingencies*, which is *sui generis* in terms of operation. The Undersecretary of Public Health operated this programme only during 2007. Nevertheless simultaneously, from 2007 to 2014 it became visible as part of other programmes such as *atención primaria* (primary healthcare), *prestaciones valoradas* (value benefits) and *prestaciones institucionales* (organisational benefits) operated by the FONASA, a public agency that operates separately from such undersecretary.

However, during the same period from 2007 to 2014, this programme also appears separately and directly operated by the ministry of health. This means, it belongs to other programmes including these previously mentioned and nine more such as

<sup>210</sup> In the budget law, most of the programmes are grouped and these are not labeled in sub-programmes.

<sup>211</sup> See the 9th annual meeting of the OECD Senior Budget Officials in 2013 and the presentation by Ms. Paula Darville, Head of Management Control Division in the Ministry of Finance. Available at: <http://www.oecd.org/gov/budgeting/9thannualmeetingoftheoecdseiorbudgetofficialsnetworkonperformanceandresultberlin7-8november2013.htm>

*subsídios de reposo maternal y cuidado del niño* (subsidies of maternity and childcare) and *subsídios de enfermedad y medicina curativa* (subsidies of preventive medicine and disease). These programmes varied each year from 2007 to 2014 or were not included in the budget each year. Thus, discrepancies on data complicate the task of tracking results because the responsibility of the programme is not clearly defined.

Lastly, there are no programmes classified in the status of *learning*. The reasons might be the same pointed out in the methodology chapter such as the difficulty to measure them or the few incentives to encourage its use. Furthermore, some weaknesses to investigate the effectiveness of evaluation are related to the programme structure. The budget law does not clarify which programmes are operating separately, instead they are all integrated into few of them, which make the task of analyse each programme tougher. In order to distinguish the effects of each programme, these will be narrowly defined rather than classified as budgetary programmes similar to the budgetary programme structure in Mexico. This affirmation agrees to the OECD's report (Hawkesworth et al., 2013: 9), which states that the government implemented a reform in 2011 to define a budgetary programme and designing sub-programmes. However, the gap in the classification of programmes still prevails.

### **The effectiveness of evaluation in the Results-based Budgeting**

In this context, attaining successful policy evaluation means the effectiveness for taking up results of evaluations and performance indicators for the budgetary cycle every year and change in the policy sector. In general, the effectiveness has contradictory perceptions regarding the budget process. On the one hand, the study of the World Bank (Rojas et al., 2005: 71) mentions that evaluation is present at Congress during the budgetary process. Officials running the programmes claim those evaluations' results are considered either for the preparation of budget or for the decision-making of sector policies, in the 60% and 50%, respectively.

On the other hand, the same study highlights that the use of evaluations by Congress is limited in practice (Rojas et al., 2005: 53), which blocks the budgetary process cycle in some way, especially the decisions based on objective criteria that Congress should made such as to validate or reject the budget allocated to public programmes or sector policies by the executive. At the end, there are no programmes classified in the status of budgeting. That is to say, there is no evidence during the budgetary discussion and the allocation of financial resources of any implication of a policy or programme, neither of changes in the policy health sector due to evaluations made either by the executive or the legislature. Thus, the analysis of budget allocations due to evaluations is constrained to the resources assigned to the health sector as follows:

Table 7

## Financial resources of the Chilean health sector

| Year | Draft budget       | Budget approved    | Comments   |
|------|--------------------|--------------------|--|
| 2007 | \$1,145,997,836.00 | \$2,219,272,236.00 |  |
| 2008 | \$1,329,054,588.00 | \$2,531,153,989.00 |  |
| 2009 | \$1,546,233,506.00 | \$2,971,985,285.00 |  |
| 2010 | \$1,915,181,367.00 | \$3,331,429,547.00 |  |
| 2011 | \$2,381,133,468.00 | \$3,921,803,728.00 |  |
| 2012 | \$4,293,285,000.00 | \$4,301,644,923.00 | The Chamber of deputies, legislature 359, session no. 89 of 04/10/2011 proposed the amount of money pointed in the draft budget. However, in other session the Chamber increased the amount to \$4,301,644,923.00, and during the last session of the Chamber of deputies, legislature 359, session no. 116 of 28/11/2011, \$2,580,534,544.00 were approved. There is no clue of the reasons of such decrease in the budget. |
| 2013 | \$2,748,069,959.00 | \$4,713,209,055.00 | In other session, the Chamber of deputies increased the draft budget to \$7,852,675,026.00 (Special report of budget account 16 of the health sector of the draft budget law 2013) but at the end, the total of resources approved was of \$4,713,209,055.00   |
| 2014 | \$3,039,310,294.00 | \$5,217,428,088.00 |  |

Source: Health sector budget, Ministry of Health, author's own elaboration.

Concerning this sector, the review of budget discussion at Congress provides some examples of these assumptions. In the draft budget law 2007 sent by the executive to Congress,<sup>212</sup> the former requested an increase on budget to this sector of 18.8% compared to the previous year. However, it did not mention the results of the sector to support the petition, notwithstanding, several deputies rejected the draft.<sup>213</sup> In those budget discussions, a deputy<sup>214</sup> pointed out that evaluations were not considered in the draft and it lacked of mechanisms of transparency and accountability. Additionally, another deputy<sup>215</sup> argued there were 400 suggestions proposed by Congress to improve the mechanisms of evaluation and control, which have not been implemented yet.

In the draft budget law 2008, a deputy<sup>216</sup> warned out about the need to make the budget process more transparent and to increase the number of evaluations of social programmes, "Indeed, there are programmes which results are negative, even though they are still operating". Another example was found in the draft budget law 2009, in which the Special Budget Joint Committee [Comisión Especial Mixta de Presupuestos] was committed to reinforce evaluations but did not mention the strategy to do it. In those sessions, a deputy argued that the health committee does not have criteria to assess, lacks of systems to analyze data and evaluation mechanisms to follow-up evaluation results.<sup>217</sup> Certainly, in the draft budget law 2010 extra resources were given to improve the evaluation and management systems.<sup>218</sup>

In the same year, the Special Budget Committee (see art. 19 of Congress Law) was formally constituted – despite having been operated time before – to revise and

<sup>212</sup> Chamber of deputies, legislature 354, session no. 97 of 14/11/2006.

<sup>213</sup> At the end of the budgetary cycle 2007 the initial amount of health sector budget was approved \$1,145,997,836.00

<sup>214</sup> The deputy Becker a party member of "Alianza por Chile" in the Chamber of deputies, legislature 354, session no. 97 of 14/11/2006.

<sup>215</sup> The deputy Cardemil of the Chamber of deputies, legislature 354, session no. 97 of 14/11/2006.

<sup>216</sup> The deputy Delmastro in the Chamber of deputies, legislature 355, session no. 100 of 13/11/2007.

<sup>217</sup> The deputy Melero in the Chamber of deputies, legislature 356, session no. 105 of 19/11/2008 and 20/11/2008.

<sup>218</sup> Special report of budget account 16 of the health sector of the draft budget law 2010.

inform about these agreements regarding the budget. Over time, it seems that evaluation became more relevant for Congress people, at least in the political discourse.

Another key point took place in the draft budget law 2012,<sup>219</sup> the chief officer of the Undersecretary of Health Networks pointed out that results of evaluations were important for decision-making regarding allocation of the health care centers. Apparently, the evaluation as policy was boosted in the budget decree and the discussion emphasizes the fact that evaluation occurred in specific administrative process such as the performance evaluation of officials or the environmental impact evaluation (system). However, the review of budget discussion did not mention any of those evaluations, which had already been performed by external evaluators.<sup>220</sup>

In 2014, the evaluation took a positive turn and those (evaluations) promoted by the DIPRES were used in the budget discussion as a tool for decision-making in different areas than the health sector. For example, the Budget Subcommittee's<sup>221</sup> report indicates that the programme *Bonificación a la contratación de mano de obra* (Bonus for workforce hiring) run by the Ministry of Work had unsatisfactory results and overlapped to another programme *Servicio Nacional de Capacitación y Empleo* (National Service for Training and Employment 'SENCE'). Thus, the former was cancelled in 2013 and part of its resources will be reallocated via the SENCE. Another example is located in the social sector. A deputy<sup>222</sup> requested to the Ministry of Social Development a report of the evaluation of policies and programmes' performance including their impact in the budgetary process for better decision-making.

One more case was the programme *de coordinación reforma judicial* (judicial reform coordination), for which, the Budget Committee<sup>223</sup> claimed that its continuation during 2014 would have depended upon the results by the impact evaluation to prove a satisfactory performance but there is not results about actions taken yet. However, concerning the health sector there was no mention about any changes due to evaluation results. Certainly, as previously shown in table 7 and similar to the Mexican case, there is a trend to increase public resources every year with no implications of the evaluation as a tool to support decision-making process.

#### *What explain the effectiveness in policy evaluation?*

To explain variation of the dependent variables within this country, the analysis will focus on the institutional (political) factors placed in order of their relevance influencing effectiveness, including interviews of those key informants. However, it is important to point out that the sample obtained is smaller compared to the Mexican chapter due to external factors explained in the methodology chapter. Additionally, relevant facts and features from a historical perspective inducing or threatening the aim of evaluation are also included. Similarly to Mexico, the main reason to keep a separate chapter of the variables per country is due to the extensive analysis of data needed to explain such variations. However, the comparative study across the Mexican and Chilean policies is done in the following chapter.

<sup>219</sup> Special report of budget account 16 of the health sector of the draft budget law 2012.

<sup>220</sup> Chamber of deputies, legislature 359, session no. 117 of 29/11/2011.

<sup>221</sup> Report of the third Special Budget Joint Subcommittee of the account 15 of the Work Ministry; bulletin no. 9.117-05, 07/10/2013.

<sup>222</sup> The deputy Silva in the budget discussion, see report of the first Special Budget Subcommittee of the account 21 of the Social Development Ministry; bulletin no. 9.117-05, 18/11/2013.

<sup>223</sup> Report of the Special Budget Joint Committee of the draft budget law 2014; bulletin no. 9.117-05, 25/11/2013.

### **Intra-Governmental Relations (IGR)**

The first factor has three sub-variables:

- i) Forms of coordination: strong, weak or minimum.
- ii) Information: availability of data, timing, reliability and quality.
- iii) Economic resources to support policy evaluation: even or uneven distribution.

To start the analysis, it is important to say that regarding the evaluation at intragovernmental level, both the judicial and legislative branches are excluded to perform any evaluation (e.g. object under evaluation), consequently to be accountable of their performance. In general, the first assumption is that forms of coordination prevailing in the Chilean administration are weak and coded “O” because the outcome is not existent. This feature is recognized as the opportunistic approach or exchange noted by Kaufmann et al. (1985: 801), where the asymmetry of roles prevailed between principals and agents, superiors and subordinates, the rulers and the ruled as explained in the methodology.

These asymmetries along with particular interests and values in budgeting negotiation jeopardize optimal relations boosting informal relations between those coordinators of policy evaluation (principal) and users (agent) such as policy implementers (government agencies), and NGOs. An example of this economic theory of principal-agent (Ross, 1973)<sup>224</sup> was during the stage of modernization of the first democratic regime of Aylwin. In those days, it seemed that evaluation was strong and the relation between the executive and legislative was based on mutual consensus both acting as coordinators. Indeed, officials interviewed<sup>225</sup> assure that relations between coordinators and operators of programmes and/or Congress are effective in terms of coordination.

In 1996, the Chilean Congress realized the importance to establish an evaluation system by law similar to other countries such as the National Audit Office in the United Kingdom or the General Account Office “GAO” in the United States of America. But it was until 2002, when deputies approved the creation of the Department of Evaluation of Law<sup>226</sup> with the aim to assess regulations enacted by Congress creating a network of non-governmental organisations interested on evaluation topics.<sup>227</sup> Yet beyond their mutual consensus, a counterargument by Tello (2011: 248-51) indicates that the lack of intragovernmental coordination was one of the major issues to operate such network. Similarly to Aylwin’s administration 1990-94, the parliament along with the political sphere were excluded of the modernisation process accomplished by the executive, reason to call this process as extra-parliamentary.

While both branches started from very similar points, the coordination took separate pathways after the authoritarian regime stepped down from power. The executive began implementing minor management changes – with the same results – in the administration such as the goals system and other performance measures as it was explained in the first section. Mainly, the evaluation was developed in Frei’s administration 1994-2000 thanks to the leadership of a small group of expert officials

---

<sup>224</sup> See methodology chapter for principal-agent theory.

<sup>225</sup> Interviewees #20, 21, 22 and 23.

<sup>226</sup> Proyecto de acuerdo (draft of agreement) no. 69 (see Memoria anual 2010-11 of the Chilean Congress).

<sup>227</sup> Regardless, many evaluations were performed in the meantime between 1995 and 2002 some coordinated by the DIPRES and others by the Undersecretary for Social Evaluation (Subsecretaría de Evaluación Social) in the Ministry for Social Development (Ministerio de Desarrollo Social).

who introduced the modernisation agenda in the governmental routine but without any convergence with the legislature.

Another example of these asymmetries took place inside the executive. During the presidency of Lagos 2000-06 some clashes between the MINSEGPRES and DIPRES arose. In spite of the fact that both agencies had a role coordinating similar monitoring and evaluation mechanisms, as it was explained in the historical section, the former attempted to keep control of the Ministry's Goal System (re-launched in 2000 as the System for Monitoring Government Programming), while the latter ran the SECG.

In the end, agencies followed DIPRES' requirements and such outcome was or "might" have been due to its power in terms of control of the budget. As one MINSEGPRES' official interviewed by Dussauge (2013: 152) affirmed, "the Finance's dictatorship" prevailed, even over the ministry supported by the presidency. The enhancement of evaluation mechanisms and availability of data do not mean decision-makers will take these results for further improvements. At least not at Congress, where the analysis of budgetary process indicates that evaluation is present but not as a mechanism for supporting decisions regarding budget in the health sector despite some evidence of good results in other areas such as the social or work sectors. Although it [Congress] agrees about the quality and timely nature of data, it does not seem to use it (Rojas et al., 2005: 53). At least until 2014, the evaluation was only in the rhetoric during the budgetary process as explained in the last section.

The second sub-variable is the availability of information (evaluation) in time for the planning process of the following year and the budgetary process. In some cases data is not available for such process. According to the paper "Evaluación de programas e indicadores de desempeño" (Evaluation of programmes' and performance indicators) by DIPRES (2001), the indicators resulting from the evaluation of public programmes (EPG)<sup>228</sup> and the comprehensive management report (BGI) are considered for the budgetary process, specifically in the draft budget. The DIPRES' paper argues that data was used for this purpose; however, there is no evidence to affirm this currently occurs.

Of 13 programmes of the health sector assessed, more than a half have performed an evaluation of this type (EPG) and should be considered for decision-making. However, the review of these 7 programmes assessed shows that 3 of them have not informed [yet]<sup>229</sup> in the report to following up results about changes in the programmes resulting from evaluation. Thus, this valuable data to formulate the budget based on the sector performance did not use it due to its availability and coded "0". This means, data is out of time when it is needed for the tailoring budget process in the following year.

The last sub-variable is related to the economic resources to support policy evaluation. According to the DIPRES presumably resources are those necessary to fulfil evaluation requirements. For this, the World Bank's study (Rojas et al., 2005: 82) agrees that the Chilean evaluation system has a low cost compared to other

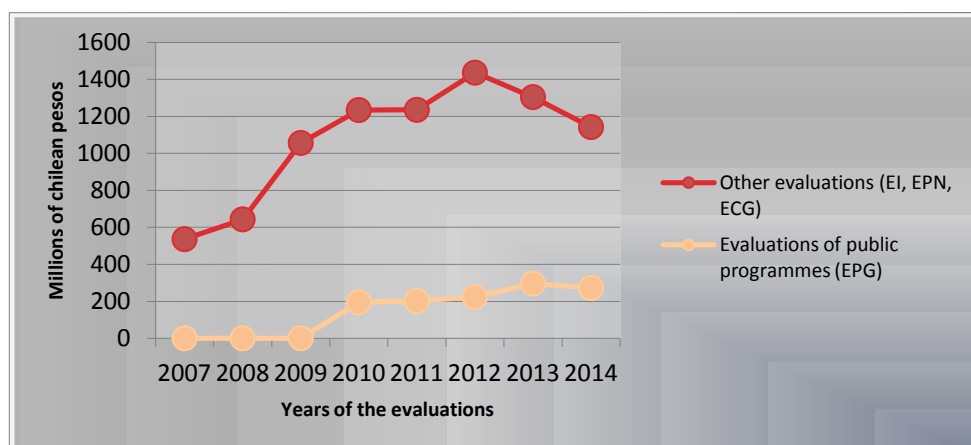
<sup>228</sup> For this purpose, the EPG has an Interministerial Committee integrated by the Ministry for the Presidency's General Secretariat (MINSEGPRES), Ministry of Cooperation and Planning (MIDEPLAN) and the Ministry of Finance through the Budget Office (DIPRES) to coordinate and assure evaluations have been performed. But the latter is responsible of the methodologies, provision of resources and reception of evaluation reports, as well to their inclusion in the budgetary process.

<sup>229</sup> At the time this thesis took place.

countries such as Australia or Canada.<sup>230</sup> The examination of resources provided for implementing evaluation across public administration has shown these are directly administrated by the DIPRES, keeping in mind that evaluations of public programmes (EPG) is the most frequent compared to other types of evaluation. The table below provides in detail a comparison of resources given for such purposes.

Chart 4

#### Resources provided to evaluations across the Chilean public administration



Source: Budget Office's budget (DIPRES), Ministry of Finance, author's own elaboration.

Regarding these results, and despite there is no data available from the EPG from 2007 to 2009, there is an incremental trend on resources, especially for the EPG – its subsistence might be due to being the one used for the budgetary process –. Conversely, there is an important reduction observable in 2013 and 2014 to the set of EI, EPN and ECG's evaluations. This agrees with the fact noted in the policy evaluation framework chapter about not having data since 2011 of any of these health sector evaluations (EI, EPN and ECG) performed. In the same period reviewed, resources for implementing other evaluations (impact, new programmes and comprehensive spending review) had an important increment from 2009, the double if 2007 and 2009 are compared.

Although, at the beginning an even distribution of public resources ensured operation of the policy evaluation, there was a reversible tendency from 2013 and 2014, when resources dropped drastically to \$870,169.00 Chilean pesos in 2014 (see table 8). An increase or decrease of public resources in any area, programme or sector corresponding either to the support or lack of interest by policy-makers and politicians in the programme referred. Thus, this uneven distribution of resources over time is equivalent to a code "O". Regarding resources granted to the DIPRES as coordinator of evaluation, there is no data in the budget to make assumptions and there is still not an official response by DIPRES.

<sup>230</sup> Indeed, in 2004 the DIPRES developed an indicator to measure the cost-effectiveness of the system "Percentage of budget assessed versus budget to be assessed", see DIPRES.



Table 8

**Comparison of resources to perform evaluations**

| <b>Year</b> | <b>Budget executed for EPG</b> | <b>Budget executed for other evaluations</b> |
|-------------|--------------------------------|--|
| 2007        | No data                        | \$536,211.00                                 |
| 2008        | No data                        | \$643,711.00                                 |
| 2009        | No data                        | \$1,056,383.00                               |
| 2010        | \$195,855.00                   | \$1,039,688.00                               |
| 2011        | \$202,318.00                   | \$1,033,000.00                               |
| 2012        | \$223,076.00                   | \$1,213,652.00                               |
| 2013        | \$296,352.00                   | \$1,008,420.00                               |
| 2014        | \$272,924.00                   | \$870,169.00                                 |

Source: Budget Office's budget (DIPRES), Ministry of Finance, author's own elaboration.

Following this, the reduction of resources is not congruent with the intention of the current administration to create the National Evaluation Office to evaluate every year 25% of the programmes. This means, by 2018 at the end of present administration, the total of programmes should be evaluated (see the government programme 2014-2018).<sup>231</sup> Thus, distribution of resources reflects the degree of support for the policy evaluation, which consequently affects its effectiveness.

In principle, the experience across countries (see OECD DEREc)<sup>232</sup> indicates that evaluation needs some leading (O'Toole, 2007: 149) in terms of agency (ies), strategy and actors to keep it as the centre of attention over all the public administration's issues. This is particularly so when the task is handling a common programme that involves operations across organisational [intra or inter] lines (O'Toole, 2007: 143). The evidence confirms the less support (including budget) to evaluation from those who are leading, the less commitment from those evaluated for taking-up recommendations.

**Quality of bureaucracy**

To measure the second variable the following sources are utilised:

- i) Government effectiveness (The Worldwide Governance Indicators 'WGI')
- ii) Evaluation of staff unit in the Chilean government: background, skills, experience through interviews.

To start the analysis, it is relevant to bear in mind that for achieving the effectiveness of evaluation it is essential that those coordinators and implementers have the skills required to lead the strategy, as well as the know-how and technical knowledge to fulfil the needs of the evaluation and monitoring systems, data analysis and decision-making process (see methodology chapter). Thus, the quality of bureaucracy and the ability of the government to implement a proper civil service guarantee effectiveness in great extent. For example, during the 70s, a considerable number of people went out to prestigious universities in the USA or Europe to undertake postgraduate studies (Bardon et al., 1985: 1-2).<sup>233</sup> Eventually, this will contribute to specialization of the high-bureaucracy in political studies for greater management of the public

<sup>231</sup> <http://www.gob.cl/programa-de-gobierno/>

<sup>232</sup> <http://www.oecd.org/derec/>

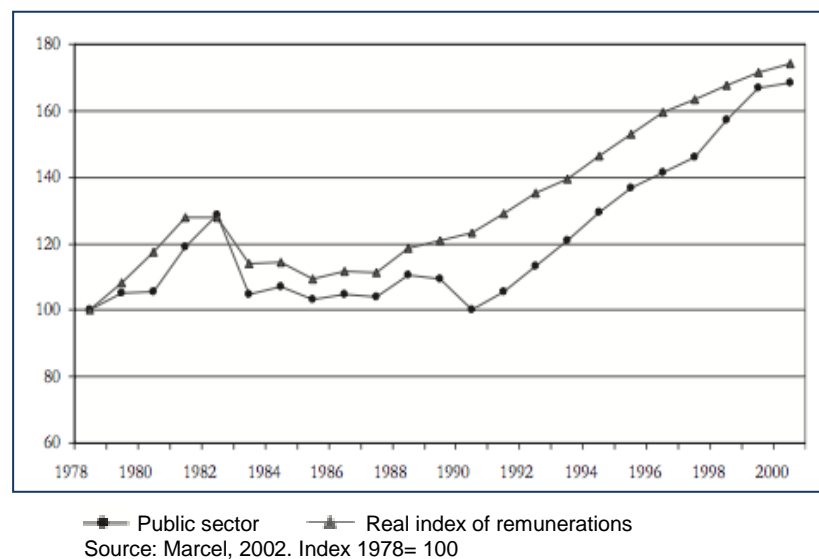
<sup>233</sup> Indeed, as long as many people studied out the country as those international universities sent academics for teaching in the most important universities of Chile e.g. the Universidad de Chile and the Universidad Católica de Santiago.

administration and to empower a group of academics, who have influenced the political debate since the decade of the 1970.

A fact emerged from the diagnostic of the situation in Chile shows that the bureaucracy was mistreated for a long time, being characterized by unfit officials for positions required and poor salaries gradually reduced (Waissbluth, 2005: 14). The salaries of the public sector were lower compared to the private sector for a long time (see chart 5). Remunerations dropped drastically during the military regime and later on in the 80s along with the State capability to operate the sector (Marcel, 2002). These eventually facilitated the rise of inefficient bureaucratic structures such as the trade unions in the education and health sectors, with a damaged reputation perceived by the society.

Chart 5

#### Remunerations (real) in the Chilean economy and public sector 1978-2000



At this point it is important to refer to salaries as a necessary condition for the improvement of the quality of bureaucracy, only to the extent that these salaries are linked to the capacity of the workforce to offer better services and for better decision-making. Thus, the new deal law (further explained) along with fair and worthy wages implemented in the last years contributed to improve the quality of the Chilean bureaucracy. Waissbluth (2005: 6) even says that the monthly nominal salary in Chile has increased from \$950 USD in 1987 to \$1900 USD in 2003, which means the double in almost two decades. Indeed, Marcel (2002: 259) claims during those decades, salaries had recovered in around 70%. Certainly, in the 90s, the PMG was implemented with the aim – among others – to improve their salaries linking remunerations with the improvement of utilities and services (Arenas & Berner, 2010: 75, Zaltsman, 2009: 454, 459), with the consequent enhancement of public sector and officials' compensations.

The amount per year when the programme started was of 3% and 1.5% of remunerations to officials on the extent that the agency had accomplished 90% or less of the 90% of organisational goals, respectively (law no. 19.553). Eventually, the law no. 20.212 of 2007 modified the percentage up to 7.6% (see table 9). The latter authors add that after ten years since this programme was launched it has been

consolidated as a mechanism to articulate the human resources policy, remunerations and public management improvement.

Table 9

#### Evolution of incentives in Chile 1998-2010 through the PMG

| Goal    | 1998-2003 | 2004 | 2007 | 2008 | 2009 | 2010 |
|---------|-----------|------|------|------|------|------|
| 90-100% | 3.0%      | 5.0% | 5.7% | 6.3% | 7.0% | 7.6% |
| 75-89%  | 1.5%      | 2.5% | 2.9% | 3.2% | 3.5% | 3.8% |
| < 75%   | 0%        | 0%   | 0%   | 0%   | 0%   | 0%   |

Source: DIPRES, 2014

A third fact took place after the crisis of corruption mentioned in the first section during the presidency of Lagos (2000-06). The Civil Service Office was created through the Law no. 19.882 of Senior Public Executive<sup>234</sup> [Ley de Alta Dirección Pública] enacted in 2003 (see Figueroa et al., 2011, Guzmán, 2005), which eventually helped to enforce the quality of bureaucracy. All the above circumstances could explain to some extent the increase interest for boosting the evaluation in Chile, mainly after the 90s, where public administrations were strengthened thanks to the administrative reforms carried out by the new democratic regimes. The results from the four interviews show these officials operating the PbR have a master degree and up to 15 years of experience working either in the public service or evaluation activities.

For Waissbluth (2005: 10-11) the severe crisis of the public sector, which occurred in this century under the administration of Lagos 2000-06, brings up the opportunity for an agreement between the government and the opposition to implement many reforms including the *ley del Nuevo Trato* (New Deal law) to define the selection process for public officials.

Table 10

#### Government effectiveness (quality of bureaucracy) in Chile

| Country/year  | 2007      | 2008      | 2009      | 2010      | 2011      | 2012      | 2013      | 2014      |
|---------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| <b>CHILE</b>  | <b>87</b> | <b>84</b> | <b>85</b> | <b>87</b> | <b>86</b> | <b>86</b> | <b>86</b> | <b>84</b> |
| <b>Mexico</b> | 59        | 61        | 60        | 60        | 63        | 63        | 63        | 61        |

Source: Data from the World Bank (worldwide governance indicators WGI).

Regarding the Worldwide Governance Indicators (WGI),<sup>235</sup> in general the country was placed in the 75<sup>th</sup> and 90<sup>th</sup> percentile range over the period of 2013, whereas the result of the indicator *government effectiveness*, which measures the quality of the civil service amongst others, was of 86 in the scale of 0 to 100. As shown above, the quality of bureaucracy is present and coded "1" due to results of the country are above 50.

This far, actions taken to boost the quality of the bureaucracy are currently addressed to the executive. Bureaucracy expertise plays an important role. During

<sup>234</sup> Chilean National Congress Library, 11 of June 2003.

<sup>235</sup> The Worldwide Governance Indicators measure six broad dimensions of governance through six percentile ranges in the scale 0-100<sup>th</sup> of 215 countries over the period 1996-2013, however, here is only included 2013 (see methodology chapter).

Frei's administration (1994-2000), the officials' educational background was from prestigious universities and expertise in the private sector, thus, their political capital was decisive in dealing with politicians for the State's modernization (Figuerola et al., 2011: 84).<sup>236</sup>

This group was lead by Mario Marcel<sup>237</sup> a well-recognized leader with a top-educational background, whose role as budget director and principal actor inside the DIPRES (Armijo, 2002, Figuerola et al., 2011, Guzmán, 2005, Tello, 2011, Waissbluth, 2005) enabled him to carrying out the modernization agenda across the public administration. Indeed, in his preliminary paper delivered to the CLAD, Waissbluth (2005: 12-3) affirms that thanks to those "intrapreneurs gerenciales" (executive entrepreneurs) inside the DIPRES and the Ministry for the Presidency's General Secretariat 'MINSEGPRES' in the 90s, whose role was known as the "reformers engine" of the modernization reform, helped to run it less slowly than it was.

However, the decision to make effective the evaluation in a systemic way is not only made by the side of the executive, it also depends upon the legislature. Thus, the quality of the politicians at Congress is fundamental to perform their responsibilities e.g. to improve sector policies and the budgetary process. In this sense, experienced politicians are essential to pursue the effectiveness of evaluation. Nevertheless as Armijo (2002) argues, in those years, the legislature did not tackle modernization reforms due to low impact of the issue. Also due to lack of experienced cadres, as this author suggests most of the politicians, who remained at Congress after the military regime, came to this new era of governance without any knowledge about modernization reforms and how to improve the way budget has been tailored and allocated.<sup>238</sup> Thus, since the last two decades the executive has led the implementation of policy evaluation and carried it out across public agencies and the legislative. These asymmetries in the quality of bureaucracy of branches are might be the reason of the gap in the effectiveness.

In the end, the taking-up of recommendations has been influenced "slightly" positively by the bureaucracy expertise, despite Congress did not make effective use of evaluations for a long time. The review of the budgetary process shows that in recent years the legislative agenda has been marginally influenced by bureaucracy (executive), highlighting the need of implementing the results-based budgeting without relevant results. In Carpenter's (2010: 15-16) study of power inside the pharmaceutical regulation this facet is denoted as the "gatekeeping power", which becomes visible by highlighting the importance of using evaluation results in the budgetary process. For this author, power appears in the ability of those who hold it to define what sort of problems, debates and agenda should be structured,

...power exists not only in broad formal authority to direct the behaviour of others (directive power) but also in appearances that are less obvious: the ability to define what sorts of problems, debates, and agendas structure human activity (gatekeeping power), and the ability to shape the content and structure of human cognition itself (conceptual power). (Carpenter, 2010: 15)

<sup>236</sup>These authors cite Domínguez (1997) to refer to such experienced executives as the "tecnopol".

<sup>237</sup>In his presentation "La dirección de presupuestos avances 2000-2005 y desafíos futuros" in 2006, Marcel argues that in the beginnings of the 90s, the DIPRES had 148 officials (75 with degree and 2 postgraduates) and one decade later this number increased to 205 officials (140 with degree and 20 postgraduates) (Marcel, 2006). In the 90s, in general, the proportion of officials working in the Chilean public sector in relation with the workforce of the country was of 5.4% (Marcel, 2002: 229), whereas compared to Mexico for the same period was of 11.5% (Rojas, 2006: 319).

<sup>238</sup>The influence of other external factors such as the political environment where evaluation was implemented in the 90s was not favourable because there was no consensus between the legislative, executive and citizenry.

This means, in the day-to-day routine, the bureaucracy shapes and performs policies or programmes, and field implementers or policy makers have the opportunity to articulate the intended policy as it was enacted by regulations (law, bill, and act) into tangible outcomes for society. They have the power to put it [evaluation] in action. So, even at this level, on the extent they have power as these executive entrepreneurs mentioned above had, they could influence the agenda in favour of the evaluation. In regimes like the Chilean, the political support needed to spread policies across the public administration could be found at the executive, mainly because this branch is backed up by the power of the president and because there had been results despite misutilisation by Congress.

Then the capacity of the bureaucracy is relatively strong in terms of enforcing the effectiveness of policy evaluation, and it even seems like disregarding the relationship with Congress. In the meantime, the legislature has limited its participation as a user of data of these results instead of coordinator and decision-maker. Based on the latter, it is assumed that effectiveness depends more on the commitment and strong support from the executive rather than on the legislative to exert pressure over the rest – public administration – for taking-up actions as result of evaluations.

### **Level of democracy**

To analyse this indicator, the index developed by the Economist Intelligence Unit (EIU), which depends upon the characteristics of the country and could be classified into the four categories such as an authoritarian regime, hybrid, flawed or full democracy. Of 167 countries of the index, Chile is placed in number 32 of the ranking with an overall score of 7.8, which classifies the country as a flawed democracy. However, this indicator is coded “0” and consider as an outcome absent. For Heady (2000: 516), the connection between these political regimes explains the variation of bureaucracies as well. Some assumptions come up such as the fact that Chile went through over decades of military regime, which pushed those administrations to establish democracy during the 90’s, mainly to focus on assuring minimum standards of governance (Figueroa et al., 2011: 71) in order to close this chapter of the history.

Definitely, there is a strong connection between democratic governments and effectiveness. Interestingly to say it does not mean that there were not any attempts to establish evaluation in Chile before, especially in the previous years of the transitional administrations to democracy. As it was previously affirmed, evaluation emerged under diverse circumstances and non-democratic governments as well. The evolution of evaluation across time shows that effectiveness mostly consolidates into democratic environment. For example, after the fall of the military regime and despite Aylwin’s presidency (1990-4) emerged from democratic elections, key actors e.g. Congress people remained from authoritarianism. Therefore, this might be the reason why institutions had the less enhancements and evaluation was rarely mentioned.

However, it is possible to find evidence in the political discourse since the democratic governments came up to the power. For example, the need of a system of checks and balances (Arenas & Berner, 2010, DIPRES, 2005), might be as a result of citizens’ request to reduce inequalities. However, regarding society participation, the citizenry’s claim has been modest in terms of a system of checks and balance (evaluation). Indeed, the Chilean’s society was more concerned about the transformation of government in terms of authoritarianism, human rights and

freedom. Instead of this, the interest to reinforce democracy (evaluation) arose inside the government's agenda.

The fact the modernization of public management has taken place in most of the advanced democracies gives insight that these conditions are necessary but not determinant for building the capacity of the State. An example was during the Pinochet's government where the executive led a transformation of the State through the downsizing of a number of state enterprises. However, the major attempts occurred when this military regime was over. In fact, the PMGP<sup>239</sup> (including evaluation) was boosted by the government of Frei (1994-2000) as a necessary tool for economic development and changing the State's machinery, not only for getting back the democracy (Figueroa et al., 2011: 71-2), even for the enrichment of public management. For these authors, the latter joint of governments (authoritarian and democratic) influenced political and management decisions, which came up with concepts such as efficiency, efficacy, quality, evaluation and transparency. Eventually, it will help to the stability of the political system, governance and consolidation of the Chilean's democracy.

### **Policy evaluation framework**

The variable to measure the balance between design and operationalization and how these impact the effectiveness of evaluation is divided in two variables with some sub-variables:

- 1) Coherence in the design:
  - 1.1) Matching policy evaluation objectives and outcomes.
  - 1.2) Overlap with regulations.
  - 1.3) Holding hearing of interest groups (consensus between actors).
- 2) Structure of monitoring and evaluation system:
  - 2.1) Technical operationalization (easy/difficult).
  - 2.2) Perception of operationalization.
  - 2.3) Methodological rigor

The first variable is the coherence in the design in the three levels of evaluation. However, all sub-variables are coded "0" due to the following reasons. At the macro level, the policy reviewed here is coherent with the Chilean government programme 2014-18. Indeed, the president's agenda proposes the creation of an agency (National Evaluation Office)<sup>240</sup> responsible of evaluation and to coordinate public agencies in the cycle of policy making. The latter is linked to the "system of evaluation and management control" as the mechanism to improve the quality of public expenditure as well as the performance of public agencies for supporting the decision-making using the concept of results-based budgeting. Indeed, this system (see annex 7) includes several other elements such as the performance indicators, evaluations and organisational wage incentive.

Regarding incentives, these are linked to the PMG (Management Improvement Programme) through five major areas such as the planning/management control<sup>241</sup> for linking goals to incentives (see DIPRES, 2001a, 2012, 2014) as explained in the section of quality of bureaucracy. These subsystems seem to match perfectly in terms of coordination at central level by the Ministry of Finance (DIPRES). On the one hand,

---

<sup>239</sup> Political Reform and Modernization of Public Management.

<sup>240</sup> <http://www.gob.cl/programa-de-gobierno/>

<sup>241</sup> This planning/management control area includes the monitoring of organisation's performance system.



the system of evaluation guides the government-ordered evaluation across the public administration at the central (macro) level and the regions located in the meso level.<sup>242</sup> On the other hand, in the micro level the PMG relates the accomplishment of objectives to remunerations but also leads the officials' performance evaluation (individual). However, there are some agencies acting separately to this framework, which are not interconnected with the main system.

For instance, some evaluations performed by other agencies such as the evaluation of schoolteachers' performance developed by the Ministry of Education.<sup>243</sup> This was enacted by the Law no. 20.501<sup>244</sup> and run between a group of scholars, professionals and psychologists from the Catholic University of Chile who operate through the "communal" evaluation's coordinator in municipalities. Besides, there is the national evaluation system of learning results (SIMCE) coordinated by the Agency of Education Quality<sup>245</sup> to measure students' performance. Both agencies belong to the Ministry of Education. A third example is the Environmental Evaluation Service<sup>246</sup> (SEA) responsible by the Ministry of Environment, which develops the environmental impact evaluation in the public and private sectors. Not all these evaluation practices are included in the SECG as an integral system, which converge the 'vertical-horizontal' levels of evaluation.

Furthermore, according to the IDB's study (Dumas, 2013: 31) it seems to prevail an "inflación de monitoreo" (over monitoring) from those coordinators such as the DIPRES, UPGC and the Ministry of Social Development enquiring data from agencies which deliver public services. These asymmetries make the task to determine the impact of each coordinator performing their roles difficult (see annex 7). Regarding the interest groups, the review of the literature and the DIPRES and Congress' official documents do not mention a particular interest of inclusion of the NGOs or citizenry in the policy evaluation process. As mentioned in the chapter of Mexico, their involvement is a common form of any democratic government. Certainly, a main objective of a system of checks and balance is the transparency and accountability to keep the citizenry informed about government performance. For this purpose, these groups represent the interest of the society to ensure its accomplishment.

However, only during the administration of president Lagos 2000-6 it was identified the think tank Centre of Public Studies (CEP). This had a relevant role influencing the "Acuerdo político-legislativo para la modernización del Estado, la transparencia y la promoción del crecimiento" (Political-legislative agreement for the State modernization, transparency and the promotion of growth). In 2003, according to Figueroa et al. (2011: 86), this centre also promoted the Senior Public Executive system described in the previous section. Another example also referred by this author is the department of industrial engineering at the university, which helped in tailoring and amending some of these policies and developed some research studies of specific sectors for a short period of time. In the end, the fact that only one agency operates the policy allowed those coordinators of evaluation to attain effectiveness more easily.

<sup>242</sup> The country is divided in 15 regions, 54 provinces and 346 "comunas" (municipalities).

<sup>243</sup> Through the Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP).

<sup>244</sup> Library of Chilean National Congress, 08/02/2011.

<sup>245</sup> <http://www.agenciaeducacion.cl/>

<sup>246</sup> <http://www.sea.gob.cl/>

Regarding the structure of the monitoring and evaluation system there are some gaps identified. Firstly, keeping in mind this system uses the same format to homogenize the activity of following up recommendations, the exercise of tracking them was complex. In some cases, the recommendations diverge once the programmes' operators take them for further changes. For example, in the comprehensive spending review (ECG) 2010 of the national supply centre, one recommendation was handled by the agency but there was no relation between the original version and the report to following up results. In the original version the external evaluator suggested "To include the option of contracts management at the same level of importance of the others..." While the report mentions that the action taken by policy implementers was "Developing a report to determine the level of technical knowledge that professionals responsible of contracts should have..." (see final report in DIPRES).

Thus, even though both (recommendation and action) are related, the action taken does not help to solve the problem. Another issue is the IT platform, which still needs further improvements to facilitate technical operationalization, especially once recommendations are updated by operators. The evidence indicates that there are still some reports missing to following up results such as the evaluations of public programmes 2008 and 2013 of the programmes *of obesity* and *purchase to private sector*, respectively, and the comprehensive spending review 2010 of the programme *for assessing the experimental health centres*. In general, data is not easily available, for example, how many programmes belong to the health sector, the cost of each evaluation performed and criteria to select which programme should be assessed. As result this subvariable is coded "O" as an outcome absent.

Concerning the criteria, the DIPRES noted that the size, amount of budget, performance indicators, and programmes, which are facing issues and those becoming larger, are the main criteria (Rojas et al., 2005: 63) to select them as a priority for evaluation. There is no document available referring to this,<sup>247</sup> nor the general perception about the lack of criteria to select programmes to be evaluated neither by Congress nor by ministries. Another gap is the amount spent by the DIPRES to execute its duties as coordinator of the policy. Furthermore, there is no data of any impact evaluations (EI) and comprehensive spending reviews (ECG) performed since 2011. In the case of the ECG, along with the evaluations of new programmes (EPN) both were cancelled by the DIPRES (Dussauge, 2013: 168) but there is no official document to validate that. Then some assumptions arose, might be that cancelation is the result of some evaluations or reports (Hawkesworth, 2013, Muñoz, 2005, Rojas et al., 2005 and Zaltsman, 2009), which highlighted the need to do some improvements.

However, as mentioned in the first section, in 2011, the DIPRES replaced the ECG by the evaluation of organisational spending (EGI) to improve it (see annex 3) but without results available yet. Regarding the ex-ante evaluation, results are awaiting due to its recently introduction (2014) as a former evaluation to fit into the SECG's scheme of evaluations. At this point, the evaluation of public programmes (EPG) is the only constantly reporting results over the public administration. Moreover, the review of the set of EI, EPN and ECG's evaluations of all sectors reinforces these assumptions, which could be generalized to the entire public administration. For instance, the EI only appears until 2011 for assessing two programmes of agriculture

---

<sup>247</sup> For in-depth revision of the quality of evaluations see Rojas, 2005.



and education while the others (ECG, EPN) were cancelled. These results are contradictory with the executive's attempt for boosting evaluation through the National Evaluation Office.

However, at the time this study took place it is difficult to stand a position due to the recent creation of this office with no observable results yet. Viewing these issues on the operation of the system and despite centralization of the evaluation, the perception of agencies running programmes is positive. Firstly, due to technical prestige of the DIPRES (Rojas et al., 2005: 54) but secondly, due to the synchronized constitution and budget legal framework, which is not commonly found in OECD countries (Dussauge, 2012a: 191). Moreover, the perception about operationalization seems to be effective according to those interviewed and coded "1" as a tangible result. For instance, an interviewee<sup>248</sup> mentioned that the IT website is feasible and easy to access to any user. These features "seem" to keep the balance amongst actors and regulations avoiding overlapping between them like in the Mexican case. However, variables altogether indicate the asymmetries in the system producing a weak policy evaluation framework.

### **Degree of autonomy**

The variable is divided into three dimensions for ranking agencies:

- i) Policy autonomy: low and high.
- ii) Financial autonomy: low and high.
- iii) Legal autonomy: low and high.

Verhoest et al. (2004) highlight different degrees of involvement<sup>249</sup> of policy autonomy. This notion is closely linked to the freedom and control capacity to manage their resources and for those who tailor it [evaluation] to decide on applications of regulatory framework in general. For these authors (2004: 105), the more or less policy autonomy is linked to the level of detail of policy decisions. When the agency that coordinates evaluation belongs to the government as here (DIPRES), the scenario to achieve the effectiveness could change depend upon the level of autonomy of this agency from other factors. In this case, the autonomy of policy decisions regarding evaluation is minimum and constrained to the extent to which the agency's head responsible of the programme or policy could appoint the external evaluators to perform evaluations government-ordered at the central level. As shown in annex 1, a code "O" was assigned to all levels of autonomy due to the low performance.

Ideally, the agency's freedom should be for dealing with a sort of resources such as personnel and operational management e.g. the process and procedures for organisations' performance measurement. However, the freedom given to these agencies responsible of programme/policy is limited to become independent from the executive either to decide or not on individual applications of rules and regulatory framework. This means, to operate without getting approval by the agency (DIPRES) at central level in activities such as hiring the external evaluators, amongst others. Indeed, the general perception of interviewees<sup>250</sup> is that autonomy of institutions is on

---

<sup>248</sup> Interviewee #20.

<sup>249</sup> The concepts of managerial and policy autonomy are linked and involved in the government decision-making: i) by itself without asking the agency for advice; ii) after consultation of the agency; iii) based on a proposal of the agency; iv) together after negotiation; v) after consultation of government/within strict rules set by government; vi) without consultation of government and not restricted by any rules set by government. See Verhoest et al., 2004: 105.

<sup>250</sup> Interviewees #20, 21, 22 and 23.

the extent the external evaluators are independent from the government, neither to fully yield for collecting financial resources (Verhoest et al., 2004: 114) as the maximum level desirable to have financial autonomy.

Nevertheless, it is important to keep in mind a topic recently arisen. In the second administration of Michelle Bachelet 2014-18, the government programme points out to the creation of the National Evaluation Office, which among other goals, aims at the developing evaluations,<sup>251</sup> establishing the guidelines and criteria of evaluations, and following up recommendations of evaluations. According to the presidency, the top official responsible of this office will be appointed by the Senior Public Executive system. With this political decision, the government attains to legitimize the autonomy of evaluation throughout an agency specialized in this topic. However, a gap identified here is that the National Evaluation Office's target will be to develop evaluations by itself. The mere fact that these evaluations are executed by the same agency, which coordinates them across public administration, threatens autonomy, mainly, because the executive plays a double role of judge and jury.

The risk observed here is when decision-making came from a particular interest of the leader of the executive (president), the top board of the ministry or group might control it in some way. This takes the discussion to the next variable (financial autonomy). The risk when an organisation is financially dependent by governmental funding is that these decisions and processes - policy autonomy - are restricted to a superior authorization. This places the agency at the minimum level as Christensen (1999) noted, generally, where grants are provided by the central government discretionary regulations may apply. Lastly, the legal autonomy is closely linked with the latter and remains at the minimum level despite some independence granted.

On the one hand, the government operates as coordinator of the evaluation through the DIPRES, which makes the decision-making process easier. This means, this agency decides by itself without previous consultation of any other public agency, the president, neither Congress. On the other hand, when the parliamentary action is not needed the centralisation of decision-making competencies is implied (Verhoest et al., 2004: 106), which potentially threatens the taking-up of recommendations. In other words, inaction by Congress as the representative body of citizenry jeopardizes the accountability of government and the effectiveness of policy evaluation. After the review of the limitations described here, a last point comes up. The autonomy in all three dimensions (policy, financial and legal) is more restricted to those evaluated than to the agency leading policy evaluation.

Summing up, the analysis of these variables show that the evaluation process requires interaction between organisations, actors and institutional arrangements, recognizing that different inter and intra organisational patterns are essential toward an effective evaluation. Regardless, on the relevance of a myriad of factors, this research highlights the importance of intra organisational cooperation, mainly between executive and legislative to influence the flow of actions during the stage of taking-up results. As it was analysed here, the political dimension of the IGR is also related to the evolution of evaluation.

---

<sup>251</sup> The goal is to evaluate each year the 25% of budget programmes, so, during the administration of 4 years' time all programmes will be evaluated. See the economic programme: the strategy of State modernization in the Chilean's government programme 2014-18. Available at: <http://www.gob.cl/programa-de-gobierno/>

At some point, it seems that evaluation works more effectively in non-democratic governments as this case during the military regime, because it is easier to control all branches by the president. However, there is no evidence from these years about how evaluation influenced the improvement of programmes and the strategic planning process, mainly, because there was no transparency and accountability to track how effective the policy evaluation was.

Finally, the following chapter will develop extensively a comparison between these variables across Chile and Mexico to determine which are the most influential in the policy evaluation.

## **CHAPTER 7. A comparative analysis of the institutional factors influencing the effectiveness of evaluation**

### **Introduction**

This section explores utilisation from another perspective. As mentioned in the review of the literature, utilisation varies according to the settings in different policy sectors and countries. Looking at the reasons why the effectiveness of evaluation was achieved or not in the programmes assessed, the aim of this exercise is to assess whether and how the independent variables explain variation in the effectiveness of evaluation and which of these variables has the strongest explanatory power. The method is through the analysis of a sample of all the programmes evaluated and previously reviewed in the chapters on evaluation in the Chilean and Mexican health sectors. It also seeking to explain which country has been more effective using evaluations over time.

The chapter is divided into the following sections. The first section introduces the categories of recommendations according to the complexity of the task to be solved. The second and third part investigate the recommendations of the programmes of the Mexican and Chilean health sectors classified either as *amendments*, *status quo* or *cancellations* in order to classify them according to the categories and complexity established. The next section explores the importance of institutional factors in explaining effectiveness. The comparative analysis of these institutional factors across and within countries might help to elucidate the intricacies in the utilisation (effectiveness) of evaluation and to answer the central question *Whether policy evaluation is effective across these countries?*

### **Categories of recommendations**

Despite the difficulty to measure the complexity of the tasks (Askim, 2007: 464), the following pages will give some indications of whether these results affect the programmes/policies by tracking the adoption of the evaluators' recommendations. The analysis identifies eight categories of recommendations.

- i. *Design of the programme.* This group includes modification and improvement of the indicators of the MIR (Indicators for Results Matrix), the method of measurement, a diagnosis of the problem that the programme would tackle, the regulation, guidelines and operating rules (Reglas de Operación 'ROP'), quantification of the programme's recipients, inclusion of gender perspective, implementation of a strategic plan in the short, medium and long term, to improve the programme's goals and objectives and to increase the objective of the programme established.
- ii. *Process of the programme and healthcare services.* This covers interrelating services and recipients' database, increasing the capacity of recipients of the medical units and coverage of the programme, strengthening and simplification of the programme, opening hours services, homogenisation of procedures, an increase in programme resources (finance and infrastructure), and an improvement in the system to register recipients.
- iii. *Activities linked to the staff and programme operators.* These include training, supervision and diagnostic analysis of personnel, increase in staff size, and implementation and/or standardisation of an incentive and compensation programme for staff.

- iv. *Development of a specific evaluation or analysis.* These would include costs, impact, cost-analysis, results, cost-effectiveness.
- v. *Coordination and synergies.* These could cover such factors as competences between programmes, staff, systems, agencies and branches related to the programme at local, regional and federal level.
- vi. *IT systems.* These could refer to the digitalization of procedures, development and improvement of IT systems and services provided, and implementation of monitoring systems.
- vii. *Survey.* This would entail the development of a satisfaction survey for the programme.
- viii. *Publicity.* This would include publicity, promotion and media coverage of the programme and its results, transparency and accountability, evaluation culture.

Of these recommendations, the category related to the design of the programmes is the only one in which implementation is relatively straightforward, perhaps because these are largely desk-based and there is no need for extra funds, coordination or other resources to accomplish them. The other seven groups are more complex tasks due to the costs, time, coordination, finances, infrastructure and human resources needed to implement the recommendations. The following section reviews the programmes evaluated. In the Mexican case nine programmes were randomly selected due to the extensive nature of the recommendations made by some evaluations. For example, the programme *health caravans* has generated nine evaluations and more than one hundred recommendations were proposed by evaluators. In the Chilean case all programmes (10) were included in the analysis.

As part of this comparative analysis of recommendations achieved and not achieved, it is essential to bear in mind that the evaluated programmes are classified as *amendments*, *status quo* or *cancellations*.<sup>252</sup> The database used to determine whether or not recommendations have been improved is the SECG in the Chilean case and the mechanism to follow-up evaluation results in the Mexican case as already explained in the methodology chapter. Moreover, unlike the analyses made in previous chapters, for the purpose of this exercise, the progress of recommendations has been updated here until 2016 in order to have an accurate overview of such changes.<sup>253</sup> See annex 8 for a classification of each recommendation along with the cases used for comparison.

### **Evaluation in the Mexican health sector**

The first part of this review concerns policy evaluation of the Mexican health sector. The examples analysed here start with the programmes classified as *status quo* due to the small number of recommendations attained by policy makers. We then turn to cases where significant amendments were made to programmes and, finally, programmes which were cancelled.

#### *Status Quo programmes*

The programme *Emergency healthcare* operated by the ISSSTE has undergone three evaluations. The evaluations EED (Specific Evaluation of Performance) 2009 and EED 2010 were merged and conducted simultaneously. While this evaluation only made one recommendation – to include indicators (MIR) of impact, services provided,

<sup>252</sup> The category 'in progress' is not included due to the evaluations' results, which have not been publicized yet.

<sup>253</sup> In many cases, when agencies update information of programmes' recommendations, they do not mention date of implementation of each one, making difficult to establish a distinction between those recommendations implemented in 2014 to 2016.

management and quality of service – it was initially rejected. The agency argued that the programme was not bound to comply with enforced evaluation results (even though the Mexican government's regulatory framework for policy evaluation indicates that all programmes are required to make recommended improvements).<sup>254</sup>

Eventually, in 2010 the recommendation was implemented along with another recommendation, which sought to modify the goals according to the reality of the programme, as part of a separate cost-effectiveness evaluation. The latter evaluation made a number of other recommendations but only one (to implement a pilot to standardise triage model in emergencies) is in the process of being implemented.

Performance evaluations (EED) in 2009 and 2010 of the programme *Health services provided across different medical sites* operated by the SSA, made four key recommendations, none of which were implemented. As in other cases, the agency replied that the programme was not required to adopt evaluation recommendations. With regard to other recommendations the agency claimed that the problems identified were costly to solve and needed interagency coordination between the regional hospitals, indicating that the complexity was also a justification for rejection of the recommendations. However, more recent reports (agency statement) suggest that the recommendations will be eventually adopted.

An evaluation of the programme's design in 2010 presented five recommendations, three of which (relating to the design and improvement of indicators and quantifying the levels of potential recipients) were adopted, possibly because they were primarily desk job tasks and relatively easy to implement (though other largely desk job recommendations were not achieved). There was another evaluation of gender developed in 2010 but it was performed to a set of 20 programmes in coordination with the National Women's Institute (INMUJERES). Because INMUJERES was responsible for addressing these recommendations, these were not included.

The *programme for vulnerable people and families* operated by the SSA and SNDIF underwent four evaluations. An evaluation of consistency and results (ECR) performed in 2007 made seven recommendations, four of which were accepted. Again these mainly focused on desk-based tasks such as developing indicators, gauging potential demand and conducting national and international research. Other recommendations relating to costs management, strategic plans and synergies with other programmes were rejected because of complexities relating to coordination between different levels of government and agencies. The agency argued that the programme had constantly modified its operation due to changing requirements and that there was no time to make changes prior the elections of 2012. Instead, the agency suggested that the next administration should launch a medium term programme to address the issue.

A similar response was given to the recommendation regarding an analysis of the synergies between the programme and others. In this case the operators responded that the recommendation needed an interagency coordination and suggested solving

---

<sup>254</sup> The mechanism to follow-up evaluation results 2011 invalidated the previous two issued in 2008 and 2010. The former noted that the mechanism is compulsory to all public agencies and state enterprises operating budgetary programmes and those who have performed any type of evaluation during 2008 and afterwards. See: [http://www.coneval.gob.mx/rw/resource/coneval/EVALUACIONES/Mecanismo\\_2011.pdf](http://www.coneval.gob.mx/rw/resource/coneval/EVALUACIONES/Mecanismo_2011.pdf)

it through the intercession of the CONEVAL as coordinator of the policy evaluation in the social sector. A 2009 performance evaluation set out four recommendations, all of which were accepted except for one relating to the conduct of an impact evaluation of the programme (though no reason was given). The recommendation of the ECR 2011 was not achieved either and refers to update of the diagnostic which stresses the problem that the programme would tackle. The agency did not improve it due to the difficulty of measuring all the variety of subprogrammes of this programme and added that the following administration could probably achieve it.

The EED 2012 recommendations to improve planning and coordination with other organisations were also rejected. The agency argued that changes in the organisation of the programme meant that many of the recommendations could no longer be accomplished. Based on the categories presented at the start of the chapter, most of the recommendations that were achieved for the programmes evaluated above belong to the group called 'design of the programme'. Due to the fact that they are desk job tasks, they are more feasible for agencies to implement. When the recommendations required coordination between agencies (interagency relations), they were generally not achieved. The difficulties of resolving these problems may have motivated the decision to split up many of the programmes (though the agency statement did not mention whether the evaluation had any impact on such decision).

#### *Amended programmes*

This section reviews programmes where policy-makers adopted most of the recommendations suggested by evaluators. As was previously mentioned, there are 15 programmes classified in this category but only four will be analysed here. The EED performed in 2009 of the programme of *provision of healthcare services* operated by the IMSS made five recommendations covering a mix of indicative and operational reforms. While the agency reiterated its position that responses were not required all the recommendations were adopted. For the programme *prevention of addictions* run by the SSA a 2011 design evaluation made five recommendations, three of which were implemented (on changes and improvements to guidelines and data) while two more operational recommendations were not implemented. The lack of action on the latter goals may be due to the recent creation of the Center for the Prevention of Addictions and the time needed to ensure the programme operated more efficiently. Most of the recommendations refer to its design and were achieved thanks to the commitment of operators to improve the programme.

The programme *for protection and integrated development of childhood* operated by two agencies, the SSA and SNDIF has undergone six evaluations. While almost all of the recommendations from the first evaluation EED 2008 were achieved, the agency disagreed with the findings of the evaluation arguing that they did not reflect the current condition of the programme, that the evaluation had been based on old data and that they had not been contacted by the evaluators for better data and coordination. The only recommendation not achieved from the EED 2008 related to the development of a strategic plan, something which the agency argued was not possible given the changing needs of the programme and the proximity of elections. Therefore, the agency suggested that the next administration should boost a programme of medium term to solve these issues.

Subsequent evaluation recommendations, where they were explicit, were achieved (in the case of the EED 2010 and ECR 2011 the agency claimed that no

recommendations were made though it is not possible to verify this as the evaluation reports are not available). The recommendations of an EED evaluation carried out in 2012 refer to the development of diagnostic – only 19 of a total of 31 were developed – per state of the problem that the programme would tackle. Those were not implemented either due to legislative changes which reassigned responsibilities for Children and Young People's rights (law) or to difficulties in measuring potential recipients and coordinating with other ministries and state-level authorities.

A conclusion of the results of this programme is that in spite of the fact that the small number of recommendations not achieved belongs to the category of design of the programme, the majority of the ones attained belong to the same group. It is also worth noting that the agency appeared to have clear reasons for not implementing the recommendations.

In the case of the programme *prevention against obesity* operated by the SSA, the three recommendations made in a 2010 evaluation of design were all implemented. However, the improvements were not exactly as suggested by evaluators. For example, in order to include gender indicators, the agency organized internal meetings to analyse this recommendation but concluded that it was already providing the data. Regarding the need to develop regulation of the programme, in the agency statement it was pointed out that this was complementary to the programme of diabetes and cardiovascular risk and not responsible for solving it, but at the end, the agency developed a specific regulation as the evaluators suggested. As noted, most of the recommendations attained belong to the categories of design of the programme and of the programme's goals. Thus, given that such categories are desk job tasks, an assumption is that all these were achieved due to feasibility to be implemented by operators.

#### *Cancelled programmes*

The third set of programmes to review relates to those classified as *cancellations*. Regarding the programme *National crusade for the quality of health provision*, there is no data from the evaluations of design and ECR performed in 2007. However, in that year, the programme was cancelled and merged into the programme *Sistema Integral de Calidad en Salud* (SICALIDAD). Indeed, the evaluation of design in 2007 of SICALIDAD mentioned that the National crusade was the predecessor of this new programme. Hence, it might be that the former programme was cancelled due to the synergies between them. However, as it was mentioned in the chapter 'The effectiveness of policy evaluation in Mexico', there is no evidence to assume that such changes were due to the evaluation results.

A 2008 evaluation (EED) of the programme for the *Equitable funding of health provision* operated by the SSA suggested a number of recommendations, all of which were adopted though there were no indications of why the recommendations were accepted. The programme itself was subsequently cancelled and merged into another programme *Seguro popular* operated by the same ministry of health though, again, there is no explicit indication that these changes were the result of evaluation.

The *efficient curative care* programme operated by the IMSS has been subject to two evaluations, one on cost effectiveness and the other on process developed. The first of these generated 16 recommendations, only one of which was implemented (on the



training of medical units in the methodology of logical framework).<sup>255</sup> The second evaluation set out a further six recommendations but again only one (requiring the development of indicators related to chronic degenerative diseases) was implemented. While the agency responsible for the programme agreed with the recommendations it indicated that it was not feasible to implement them.

A later update of the agency's statement issued in 2015,<sup>256</sup> reported that the programme merged with another programme called *provision to the reproductive health* to establish a single programme entitled *provision of healthcare services* with the aim of standardising the services provided. In the end, the original programme was cancelled but the agency did not provide a reason for this. In the case of these cancelled programmes, the role of evaluation does not seem to be strategic for actors when decision was made. Perhaps, the programmes' operators decided to cancel them without taking into account the evaluation results even before evaluation was developed such as the former programme *National crusade*.

Table 11 below shows the recommendations of the sample of programmes assessed. The review of the recommendations shows that the main problem related to issues of design and process (44% and 27% respectively). This is in line with the previously mentioned CONEVAL<sup>257</sup> evaluation of social policy 2008 that indicated programme design was a recurring problem. Whether the programmes were amended, unchanged or cancelled, the issue of design was generally one that was reformed, particularly around the setting of indicators. As noted, this is likely to be because making such changes is largely desk-based and there is no need for extra funds, coordination or infrastructure to accomplish them.

However, it is worth noting the recurrent nature of this problem. Even though programmes generally adopted the design reforms recommended in evaluations, subsequent evaluations of the same programmes exhibited similar problems. This phenomenon (recommendations of the category of design) is also observed in those programmes classified as non-social sector. It can be affirmed that programmes assessed – no matter if these were classified as *status quo*, *amendments* or *cancellations* – have generally taken some actions in terms of improvements when recommendations are related to the design category, especially those linked to indicators. As previously noted in the sample of programmes reviewed, problems about design are usually easier to solve for those policy-makers and operators located at the central level.

In contrast to the implementation of design-based reforms, the evaluation recommendations regarding reforms of process were generally not followed up by agencies, even where other reforms were adopted. These process-based reforms would either require the development of a regulatory framework or the harmonization of the operation of mechanisms at local level, and would need time, interagency and/or intragovernmental coordination, specialized bureaucracy, finance, or a combination of these commitments to be attained. Since the extra effort needed would certainly delay the decision-making process, policy makers focused on reforms which were easier to solve. It is also worth noting that other recommendations which

<sup>255</sup> See chapter 'Broadening understandings of policy evaluation and effectiveness' for better reference.

<sup>256</sup> See the agency statement:

<http://www.imss.gob.mx/sites/all/statics/pdf/transparencia/evaluaciones/dpm/ppe002/fepi/FE-2014-2015-PosicionInstitucional.pdf>

<sup>257</sup> See "Informe de evaluación de la política de desarrollo social en México 2008" (evaluation report of the social development policy in Mexico 2008) by the CONEVAL.

were generally not adopted often entailed additional evaluation, greater coordination with other agencies or the development of IT systems.

Table 11

### Categories of recommendations of sample of programmes assessed across countries

| AGENCY/<br>PROGRAMME  | CATEGORIES OF RECOMMENDATIONS |         |                                     |                           |                                   |            |                         |  | Total of<br>recommendations<br>(ACCEPTED) |
|---|-------------------------------|---------|-------------------------------------|---------------------------|-----------------------------------|------------|-------------------------|--|---|
|   | DESIGN                        | PROCESS | STAFF AND<br>PROGRAMME<br>OPERATORS | EVALUATION<br>OR ANALYSIS | COORDINATI<br>ON AND<br>SYNERGIES | IT SYSTEMS | SATISFACTI<br>ON SURVEY | TO<br>PUBLICIZE,<br>PROMOTION<br>AND MEDIA |   |
| MEXICAN HEALTH SECTOR   |                               |         |                                     |                           |                                   |            |                         |  |   |
| Status Quo category   |                               |         |                                     |                           |                                   |            |                         |  |   |
| ISSSTE/EMERG<br>ENCY<br>HEALTHCARE  | 3(2)                          | 1       |                                     | 1                         |                                   |            |                         |  | 5(2)                                      |
| SSA/HEALTH<br>SERVICES<br>PROVIDED<br>ACROSS<br>DIFFERENT<br>MEDICAL SITES  | 6(3)                          | 2       |                                     |                           |                                   | 1          |                         |  | 9(3)                                      |
| SSA/PROGRAM<br>ME FOR<br>VULNERABLE<br>PEOPLE AND<br>FAMILIES   | 6(4)                          | 2       |                                     | 3(2)                      | 4(1)                              |            |                         |  | 15(7)                                     |
| Amended programmes  |                               |         |                                     |                           |                                   |            |                         |  |   |
| IMSS/PROGRA<br>MME OF<br>PROVISION OF<br>HEALTHCARE<br>SERVICES   | 3(3)                          | 2(2)    |                                     |                           |                                   |            |                         |  | 5(5)                                      |
| SSA/PREVENTI<br>ON OF<br>ADDICTIONS   | 2(2)                          | 2       |                                     |                           |                                   | 1(1)       |                         |  | 5(3)                                      |
| SSA-<br>SNDIF/FOR<br>PROTECTION<br>AND<br>INTEGRATED<br>DEVELOPMENT<br>OF<br>CHILDHOOD                              | 6(4)                          | 2(2)    |                                     | 1(1)                      | 2                                 | 2          |                         | 1(1)                                       | 12(8)                                     |
| SSA/PREVENTI<br>ON AGAINST<br>OBESITY   | 3(3)                          |         |                                     |                           |                                   |            |                         |  | 3(3)                                      |
| Cancelled programmes  |                               |         |                                     |                           |                                   |            |                         |  |   |
| SSA/NATIONAL<br>CRUSADE FOR<br>THE QUALITY<br>OF HEALTH<br>PROVISION  |                               |         |                                     |                           |                                   |            |                         |  | NO DATA                                   |
| SSA/EQUITABLE<br>FUNDING OF<br>HEALTH<br>PROVISION  | 3(3)                          | 1(1)    |                                     |                           | 1(1)                              | 1(1)       |                         |  | 6(6)                                      |
| IMSS/<br>EFFICIENT  | 4(1)                          | 10      | 1(1)                                | 3                         | 1                                 | 3          |                         |  | 22(2)                                     |
| CHILEAN HEALTH SECTOR   |                               |         |                                     |                           |                                   |            |                         |  |   |
| Status Quo category   |                               |         |                                     |                           |                                   |            |                         |  |   |
| UNDERSECRET<br>ARY OF THE<br>HEALTHCARE<br>NETWORK/EXP<br>ERIMENTAL<br>HEALTH<br>CENTRES                            | 4                             | 13(2)   | 2                                   | 2                         | 1                                 | 1          | 2                       |  | 25(2)                                     |
| FONASA/PURC<br>HASE FROM<br>PRIVATE<br>SECTOR   | 1                             | 4       |                                     | 4(1)                      |                                   | 2          |                         |  | 11(1)                                     |
| UNDERSECRET<br>ARY OF PUBLIC<br>HEALTH/NATIO<br>NAL<br>PROGRAMME<br>OF<br>SUPPLEMENTA<br>RY FOOD                    | 1                             |         |                                     | 5                         | 1                                 | 2          | 1                       |  | 10  |
| UNDERSECRET<br>ARY OF PUBLIC  | 1                             | 2(1)    | 1                                   | 1                         | 1                                 | 2(2)       |                         | 1  | 9(3)                                      |
| Amended programmes  |                               |         |                                     |                           |                                   |            |                         |  |   |
| FONASA,<br>UNDERSECRET<br>ARY OF PUBLIC<br>HEALTH AND<br>OF<br>HEALTHCARE<br>NETWORKS/OF<br>TREATMENT<br>OF OBESITY | 1(1)                          | 2(1)    |                                     | 1(1)                      | 1(1)                              | 2          |                         |  | 7(4)                                      |
| UNDERSECRET<br>ARY OF PUBLIC<br>HEALTH/COMM<br>ISSION OF<br>PREVENTIVE<br>MEDICINE AND<br>DISABILITY                | 1(1)                          | 13(10)  | 3                                   | 3(2)                      |                                   | 1(1)       |                         |  | 21(14)                                    |
| UNDERSECRET<br>ARY OF PUBLIC<br>HEALTH/OF<br>PREVENTION<br>AND CONTROL<br>OF HIV                                    |                               | 8(8)    |                                     | 1(1)                      | 2(2)                              | 1          |                         |  | 12(11)                                    |
| NATIONAL<br>HEALTH FUND<br>(FONASA)/NATI<br>ONAL HEALTH<br>FUND   | 7(7)                          | 7(3)    | 1(1)                                | 2(1)                      |                                   | 1          |                         |  | 18(12)                                    |
| CENABAST/NA<br>TIONAL<br>SUPPLY<br>CENTER   | 1                             | 8(4)    | 1(1)                                |                           |                                   |            |                         |  | 10(5)                                     |
| Cancelled programmes  |                               |         |                                     |                           |                                   |            |                         |  |   |
| UNDERSECRET<br>ARY OF<br>HEALTHCARE<br>NETWORKS/NE<br>WBOON<br>HEALTH<br>SUPPORT                                    | 1                             | 2(2)    | 1(1)                                | 1(1)                      |                                   |            | 1(1)                    |  | 6(5)                                      |
| TOTAL   | 54(34)                        | 81(36)  | 10(4)                               | 28(10)                    | 14(5)                             | 18(5)      | 4(1)                    | 2(1)                                       | 211(96)                                   |

Source: author's own elaboration.

## Evaluation in the Chilean health sector

The second part of this analysis is of the Chilean health sector. In identifying what factors were achieved in the programmes, it is of interest to explain the success as well as failures of these cases.

### *Status Quo programmes*

Of the five programmes where evaluations led to no or only modest changes, the Comprehensive Spending Review (ECG) for the first of these called *experimental health centres* operated by the Undersecretary of the Healthcare Network performed in 2010, identified 25 recommendations but only two were adopted (to take into account methods used by other agencies and introduce a new system of budgetary control, financial management and programming activities). The remainder were not adopted as of the end of 2014. While the agency responsible for the policy indicated that all the recommendations were being addressed, two years after the evaluation took place in 2012 only two had been resolved. As might be expected, most of the recommendations that were not achieved belong to the group of the process of the programme, which usually entails time, structure and most importantly interagency coordination between the health centres to be achieved.

In the case of the Evaluation of Public Programmes (EPG) in 2013 of the programme *purchase from private sector* run by FONASA, only one of the 11 recommendations was adopted (relating to the inclusion of evaluations from other medical units). Amongst those not yet implemented were a number which were relatively low cost but which required more coordination efforts by policy-makers. There are six other recommendations in the waiting list to be improved and another similar number in terms of implementing a monitoring system.

The 2013 evaluation (EPG) of the *National programme of supplementary food* run by the Undersecretary of public health set out ten recommendations covering issues such as coordination, financial resources and monitoring. However, while the agency responsible for the programme indicated that it agreed with the recommendations suggested by the evaluators and that it planned to implement them, at the time of writing this had yet to happen. These actions include developing coordination with other operators of programmes sharing the same objective to improve the nutrition of the population. Other activities are scheduled in the short term such as the operation of the IT system and a study to analyse the distribution model of services, but there is no evidence that such actions have been executed.

The 2012 evaluation (EPG) of the *National programme of immunizations* operated by the Undersecretary of public health made nine recommendations of which three have so far been adopted (on monitoring for wastage, improving the IT system and reforming budgetary and procurement systems). It is worth noting that these were implemented even though they were relatively complex changes requiring coordination and additional funds. Other recommendations – of varying degrees of complexity – have yet to be implemented.<sup>258</sup>

---

<sup>258</sup> At the time this study took place from 2007 to 2014, the evaluation of this programme identified seven recommendations and only one was improved. However, for the purpose of this section, when this study was updated to 2016, there were identified nine recommendations and three of them adopted. Notwithstanding, this programme is classified in the *status quo* category.

### *Amended programmes*

Four of 13 programmes assessed are located in this category. Four reforms for the 2008 evaluation (EPG) of the *programme of treatment of obesity* – run by FONASA, the Undersecretary of public health and of healthcare networks – related to improving interagency coordination and various data improvements were accepted. As noted before, these actions may have been easier to implement because they did not need a large amount of money to attain them. By contrast, the three recommendations rejected by the operators of this programme dealt with monitoring systems and extending the scope of the programme, tasks which required extra finances to be executed (as the agency responsible for the programme confirmed when it rejected such recommendations due to financial restrictions and the costs to accomplish them).

In the case of the 2008 evaluation (ECG) of the programme of the *commission of preventive medicine and disability* operated by the Undersecretary of public health, 14 of the 21 recommendations were accepted though there was a delay in implementing one which required the development of a new IT system for processing health services online. Others that were accepted related to the conduct of economic impact analyses, monitoring the conduct of staff particularly in relation to the provision of medical leave. Many of these recommendations were implemented thanks to the commitment of the actors and the deployment of multidisciplinary teams to address the problem, in other words, thanks to the interagency coordination. An example of this was the commitment of the Department of Finance and Administration to improve the financial management of transferring funds to the regions timely.

A recommendation achieved regarding the relocation of offices to new facilities required funds and was delayed but finally concluded after a year of works. Three of the recommendations concerning human resources were cancelled due to the reorganisation of the services and major changes such as the creation of the Centralized Medical Comptroller to improve services. In cases where the suggestions were not achieved such as implementing a management model in two regions, the reorganisation of the agency was cited as the reason. However, the fact is the agency did reorganize the services and zones for better recipients' attention. Under these circumstances, the recommendation was considered out of context and they did not improve it. The agency improved the programme, though, which is the important matter but in a different way from that suggested by the evaluators. The remaining recommendations were not adopted though the agency gave no reason to justify the lack of action.

Most of the 12 recommendations relating to the 2010 evaluation (EPG) of the programme of *prevention and control of HIV*, were implemented, according to the operators (Undersecretary of Public Health) mentioned it was thanks to the good organisation and commitment of many agencies (e.g. Undersecretary of Public Health, Undersecretary of Health Networks, National Health Fund, National Supply Center and Institute of Public Health). Such interagency coordination facilitated the implementation of actions though in some cases the process was difficult due to economic restrictions. The only recommendation the agency did not implement was related to the development of an IT system to register statistics and health data (though no explanation was given by the operators for the failure to adopt this recommendation).

The bulk of recommendations (12 out of 18) made by evaluators (ECG) in 2011 for the programme *National Health Fund* (FONASA) operated by the FONASA, have been adopted. They were achieved even though they needed additional resources to be improved suggesting that FONASA had the funds to make the improvements. Of those that were not adopted it may be that these were too costly for FONASA to implement.

A 2010 evaluation (ECG) of the programme *National Supply Center* (CENABAST) operated by the CENABAST has an equal amount of recommendations achieved and not achieved. It is worth noting that many of the recommendations that have been implemented entail process activities, which would arguably have been harder to execute. By contrast those that were not implemented were relatively straightforward entailing deskwork (though a few – such as new storage systems and paying debts to suppliers – would also have required additional resources). For those recommendations that have not yet been achieved there is an office letter issued by the center, which mentions an attempt to solve them in the short term.

#### *Cancelled programmes*

Only one of the evaluated programmes was cancelled. The 2011 evaluation (EPC) of the programme of *newborn health support* operated by the Undersecretary of healthcare networks was cancelled and merged into the programme of *biopsychosocial support* as a result of the evaluation performed. In the course of this change it is worth noting that five of six recommendations suggested by evaluators were agreed. Overall, the ministry not only agreed with the results of the evaluation but was very positive noting that it [report] “was done thoroughly based on how the programme works in both ministries, the evaluation gave us an insight of issues had not realized before...”<sup>259</sup>

Moreover, policy operators agreed to integrate this programme into the programme of biopsychosocial support. Thus, evaluation here was effective in terms of decisions made by policy makers to improve health services through optimizing the operation of both programmes. The only recommendation that was not achieved concerned the way in which programme recipients would be measured where the ministry responded that it would need to conduct further studies to establish the technical and economic feasibility of this proposal.

After the examination of the evaluations and recommendations (see table 11 above), a first assumption is that the problem of lack of action can be attributed in some way to the complexity of the task as mentioned at the beginning of this chapter, particularly in terms of economic resources. Indeed, an assumption of this exercise is that quite a few of these recommendations are not achieved due to financial restrictions. For example, the Chilean programmes of *experimental health centers* and the *national supply center* needed to restructure their medical units, infrastructure and logistic support but could not achieve this suggestion due to the high costs implicit in such improvements.

This is also the case where the programmes need to implement IT and/or monitoring systems. Even in the programmes classified as *amendments*, it appeared that the need for additional resources was a constraint on adopting recommendations. For example, in the *programme of prevention and control of HIV*, the evaluator's

---

<sup>259</sup> See the final report issued by the operators' programme of *newborn health support*.

recommendation for a consolidation of the IT system to register statistics and health data was not achieved probably due to financial restrictions. However, despite the complexity of the task there is a large number of recommendations achieved related to the process compared to other categories. This result proves that tasks like these regarding processes are solved despite complexity and financial restrictions.

It is also worth noting that, based on the responses (agency statement) from some of those responsible for the programmes, interagency and intragovernmental relations were both a problem and a solution to achieve those recommendations suggested by evaluators. In a number of programmes, including those for the *commission of preventive medicine and disability*, the *prevention and control of HIV*, the *national programme of supplementary food* and the *national programme of immunizations*, the operators mentioned the importance of the IGR either as an obstacle or facilitator to solve internal problems. Ultimately, and notwithstanding the many recommendations requiring time, infrastructure, human resources and finances, the role of IGR was a central determinant of effectiveness thanks to the potential ability to negotiate and coordinate actions to the benefit of the programmes.

Regarding the effectiveness of the Chilean policy evaluation, the programme of *newborn health support* is the only programme, which used the evaluation in the decision-making process to improve its operation or at least, the only one that explicitly acknowledged the evaluation in shaping the actions taken. Indeed, its subsequent merger into the programme of *biopsychosocial support* was in response to the evaluation performed in those years.

Table 12

**Recommendations achieved of sample of programmes assessed across countries**

| Focus of recommendations      | Number of recommendations |            | Proportion achieved (%) |                |
|-------------------------------|---------------------------|------------|-------------------------|----------------|
|                               | Mexico                    | Chile      | Mexico                  | Chile          |
| Design                        | 36                        | 18         | 25(69%)                 | 9(50%)         |
| Process                       | 22                        | 59         | 5(23%)                  | 31(53%)        |
| Staff and programme operators | 1                         | 9          | 1(100%)                 | 3(33%)         |
| Evaluation or analysis        | 8                         | 20         | 3(37%)                  | 7(35%)         |
| Coordination and synergies    | 8                         | 6          | 2(25%)                  | 3(50%)         |
| IT systems                    | 6                         | 12         | 2(33%)                  | 3(25%)         |
| Satisfaction survey           | 0                         | 4          | 0                       | 1(25%)         |
| Publicity                     | 1                         | 1          | 1(100%)                 | 0              |
| <b>TOTAL</b>                  | <b>82</b>                 | <b>129</b> | <b>39(48%)</b>          | <b>57(44%)</b> |

Source: author's own elaboration.

In a comparative perspective, while there is a slight difference between recommendations achieved in both countries (see table 12), the contrast is more significant regarding the focus of the recommendations achieved. Whereas in Mexico, most of the recommendations achieved belong to the category of design of the programme, in Chile they belong to the process of the programme. A first conclusion is that, this may be due to the way in which evaluations themselves were conducted. For example, recommendations regarding the design of the programmes are not recurrent over time in Chile as in the Mexican case because in the former there is only one evaluation performed per programme, thus, there were fewer recommendations suggested.

Moreover, as it was mentioned, as these (design category) are desk-job tasks might be easier to solve compared to the rest of the categories, as Greeve (2003) suggests, decision to use evaluations increases when the complexity of the task decreases. Furthermore, the recommendations relating to the programmes' process represent the major number compared to other categories, particularly in Chilean evaluations.

These results are contrary to Greeve's but agree with Johansson's (1995) study, which emphasizes that there is more data use when actors deal with 'hard core' tasks such as these implicit in the process of the programmes. Thus, it seems that complexity of the task is important in the decision-making for policy-makers and operators when they have to decide which recommendations will be attained. However, regarding the other types of recommendations the main area where recommendations of the Chilean programmes were made and not achieved were those of IT reforms, further evaluations and staff training. In many cases these recommendations were characterized by complexity in terms of time, coordination, finances, infrastructure and human resources.

Thus, on the basis of the analysis above, there is a partial perspective when it is assumed that the achievement of recommendations is linked to the complexity of the task. For example, the category of design of the programme is the less complex to attain because is a desk-job task, whereas the rest of the categories of recommendations are difficult to achieve and the less implemented. However, in the Chilean case, recommendations classified as hard tasks and related to the programmes' process were easy to achieve. Moreover, the issue of finance appeared to be sometimes a constraint on adopting changes but there were other cases where such changes were made, suggesting that in some cases there was a strong commitment on the part of policy-makers e.g. category of process of the Chilean programmes.

An assumption based on these results is the interest for solving them (recommendations) arises from the interest of the actors involved to make effective decisions for closing the policy cycle regardless of the complexity of the task or the costs of implementing changes e.g. time, coordination, finances, infrastructure and human resources.

### **The influence of institutional factors on the effectiveness of evaluation**

As can be noted, the results by themselves do not provide a clear view of the variation in the dependent variable to affirm that either the Chilean or Mexican policy evaluation is more effective. The effectiveness varies according to the programme and the complexity of the recommendations suggested. Moreover, to verify whether evaluation is effective it is necessary to go beyond the results presented here to grasp the intricacies of political factors and other country settings. These institutional factors are potential influences in terms of the adoption of evaluation recommendations.

Applying Mill's method to this case study, the method of agreement points out that probable causal factors found in common while measuring the independent variables (annex 1) across countries was the quality of bureaucracy. This is an outcome present in both countries; indeed, this factor places the Chilean evaluation system into a leading position. However, when these outcomes between countries are compared it appeared to be a necessary condition to achieve the effectiveness but not sufficient to assume is the only cause of the effectiveness.

Indeed, while this factor appeared to have a positive impact on the decision-making process appears that other factors are jeopardizing such process. As Ragin (2014: 38) noted “In situations of multiple causation, therefore, the method of agreement is likely to yield incorrect results”. In cases like these, the method of agreement and difference can be used jointly to find factors in common amongst all cases where utilisation of evaluations appears or when the factor is absent and effectiveness does not occurs.

The differences found amongst these countries were the indicators related to the IGR such as the reliability and quality of data, as well as the indicators of the variable of policy evaluation framework measuring perception of operationalization and methodological rigor. All these conditions impacting positively on the effectiveness of evaluation present in the Chilean but absent in the Mexican policy evaluation. The availability of data and distribution of public resources related to the IGR are outcomes present in Mexico, but not in Chile. Mill notes that the method of agreement and difference seeks for the cause of effectiveness in two ways. On the one hand, those few cases where evaluations were utilised to find at least one condition “X” in common, while on the other hand, those programmes in which effectiveness does not occur and have nothing in common save the absence of “X”, then “X” is probably the cause or part of the cause of the effectiveness of policy evaluation.

After the analysis of the results and factors within and across countries, it was found that the dynamics between the different components of government i.e. the IGR were decisive in shaping whether or not the recommendations from evaluations were adopted. Thus, although the quality of bureaucracy, democracy, autonomy and framework of policy evaluation are conditions which influence the effectiveness of policy evaluation, they seem less relevant for the purpose of utilisation compared with the interest of policy operators to lead changes in some way e.g. cancellations, amendments, etc. For instance, in both countries, most of indicators measuring the IGR variable have similar results. That is, the IGR are the condition “X” found in common either present or absent during the decision-making process. In this sense, most of the evidence shows weak forms of coordination, mainly when these refer between actors within the legislature and executive, which are based on legal authority at central level and on opportunistic incentives to enforce cooperation.

Other relevant findings regarding this variable is the relationship between coordinators of the Mexican policy evaluation, which is *sui-generis* in terms of number of actors involved. In addition, an actors’ concern is the absence of information in a timely manner, particularly it is out of time to the tailoring budget process. On the contrary, there are some differences regarding the indicator of reliability and quality of data, which according to Congress people is present in the Chilean policy evaluation but perceived as absent in the Mexican. As noted in the case study chapter, members of the Mexican Congress do not consider evaluations trustworthy, even though these were performed by external evaluators. Conversely, availability of data is an outcome present in the Mexican policy, whereas in Chile is considered an absent outcome.

Lastly, even though in both countries distribution of public resources have ensured the operation of the policy evaluation, there is a reversible tendency from 2013 to 2014 in the Chilean policy evaluation to affirm that the outcome in this country is absent due to an uneven distribution of public resources further explained in the case



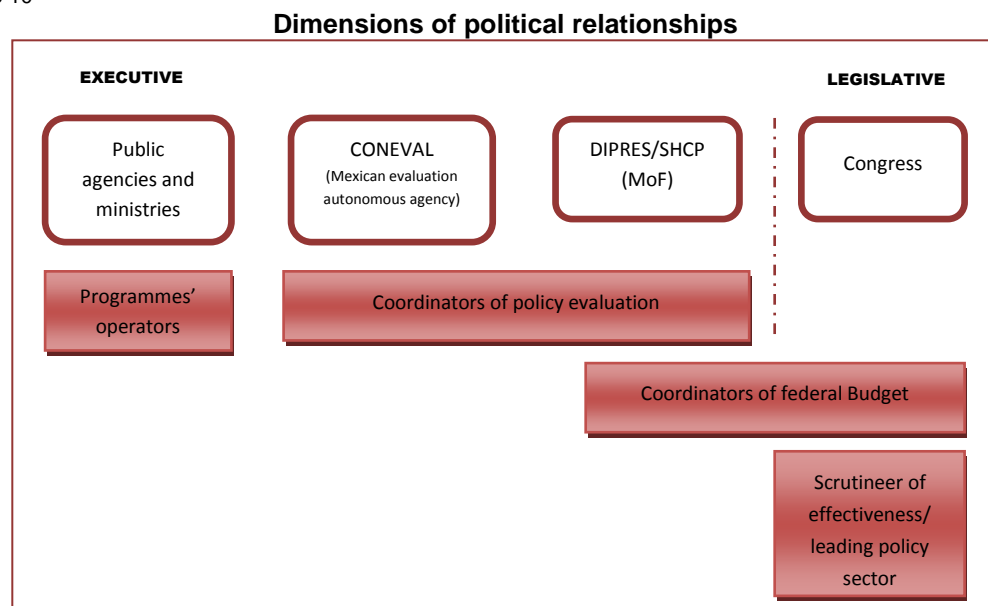
study chapter. This takes the analysis back to the importance of the political factor and the particular interest of these actors involved in the policy evaluation.

This means that no matter how strong the data is and the complexity of the tasks, it does not substitute for the capacity of actors to make political decisions based on evaluations. Therefore, it is possible to establish a relationship between the institutional factors and the decision-making when taking-up recommendations, in particular when these refer to the need of coordination and synergies between areas, programmes, staff, systems, agencies and branches (IGR) as most of the programmes do. Also, in many of the agencies' statement, the IGR plays an important role to solve the main issues mentioned here.

To understand the role of many actors in the effectiveness of evaluation, it is important to point out to different dimensions of political relationships within the variable IGR. After the analysis of these dimensions at the federal (horizontal) level,<sup>260</sup> the following relationships are illustrated in the figure below. In the first place, there is a pattern of interaction between those operators (bureaucracy) of programmes and policies, and coordinators of policy evaluation mostly located in the Ministry of Finance 'MoF' (DIPRES in Chile and SHCP/CONEVAL in Mexico). The former receives evaluations and implement those recommendations suggested by external evaluators while the latter is responsible for the coordination of policy evaluation but most important for planning of the federal Budget.

In the second place, there is a pattern of coordination amongst bureaucracy or programme operators of different agencies to solve problems related to recommendations of evaluations, for which two or more programmes are interlinked in some way. Lastly, it was identified the role of the president through the MoF and Congress. The MoF develops the budget while the Congress authorizes it. Regarding evaluation, the MoF coordinates the policy evaluation while Congress oversees its effectiveness and use evaluations to lead changes in policy sector.

Figure 10



Source: author's own elaboration.

<sup>260</sup>The focus of this study is on the federal level of both countries and explained in the methodology chapter.

The relationship between programmes' operators and coordinators of policy evaluation operates in a way similar to that presented in the agency theory (see Eisenhardt, 1989, Mitnick, 1975 and Ross, 1973), in which the principal (coordinators of policy evaluation) delegates work to the agent, who is the performer of evaluation results (programmes' operators). For instance, during the implementation stage of recommendations, policy makers in both countries operate under supervision of the principal. The traditional approach of most of public agencies regarding evaluation is to fulfil requirements imposed by coordinators of evaluation, particularly by the MoF. However, a problem in the principal-agent relation emerges when these actors do not solve an issue or make different actions from those expected or do not implement actions as result of evaluations.

It must be remembered that most of public programmes should implement some actions after being evaluated. However, most of the Mexican (65%) and Chilean (69%) programmes evaluated have made little changes including cancelations or even worst, there is no data of changes made. These numbers confirms a disruption noted in the agency theory and low commitment to policy evaluation by policy operators. Only few exceptions of agencies working almost independently from these coordinators across government can be found. For instance, in the Mexican case, the Ministry of Social Development 'SEDESOL' arises out of the power of coordinators, leading changes on programmes based on evaluation results. In part, this can be explained thanks to the quality of bureaucracy inside evaluation areas committed to the purpose of evaluation. Indeed, the working team of the CONEVAL (coordinators of policy evaluation of the social sector) comes from the department of evaluation within the SEDESOL who later on founded the council.

Notwithstanding the quality of bureaucracy (actors) as a condition needed to advocate such process, it can be affirmed that whether or not the policy cycle is closed depends upon the extent to which the IGR are committed to the evaluation process and the results of programmes. At this point, the second type of relationship between programmes' operators within public agencies and ministries can be seen as important, the potential to bring about change depends on the extent to which there is a strong form of partnership between these actors to perform joint actions. For example, in the Mexican case, the operators of the *efficient curative care* programme worked in coordination with local health authorities and with other operators of the programme *provision to the reproductive health* to merge both programmes to standardize the services provided by them. However, this action was not in response to an evaluation recommendation but a strategy made by the IMSS to improve healthcare services.

The enormous efforts made by these programme operators required the shared action of the IGR to solve a difficult task since the changes called for extensive coordination at the local level due to the fact that the healthcare services are decentralized and that the local governments operate autonomously, a complicated issue to solve by policy-makers and operators themselves. Therefore, the solution of these problems depends upon the capacity and interest of local governments. This interaction between the federal and local levels is also apparent in the Mexican programmes *Prospera*, *reduction of maternal mortality*, *health caravans*, *integral system of health quality*, and the *inclusive strategy of food welfare*, which needed of intergovernmental coordination to operate.

In the Chilean case, a set of recommendations of the programme of the *commission of preventive medicine and disability* related to audits, medical leaves and manual of health services procedures were implemented thanks to the commitment of the actors operating this programme and the deployment of multidisciplinary teams to address the problems. In other words, thanks to the interagency coordination. Indeed, the operators of the *programme of the commission of preventive medicine and disability*, *programme of prevention and control of HIV*, *the national programme of supplementary food* and *the national programme of immunizations* mentioned the importance of the interagency and intragovernmental relations either as an obstacle or facilitator to improve the programmes.

In other cases, the coordination needed increased the complexity of the task, which might be a reason why the recommended improvements were delayed. An example of the lack of coordination is the Mexican *programme to vulnerable people and families*, in which the evaluators made recommendations relating to cost management, strategic plans and synergies with other programmes. These recommendations were rejected because of complexities in the coordination between different levels of government and agencies.

Earlier was noted that the effectiveness of evaluation depend upon other institutional factors. This means that it is also important to consider the time of revision of the policy evaluation, the reviewing of the historical processes and other country settings to understand the context under which evaluation occurs. Currently, the core of policy evaluation is the Results-based Budgeting, which aims to link governmental planning, design, and implementation of public programmes with the budgetary process through evaluation. Thus, evaluations should be a key element in the allocation of resources within the budgetary process. Also, it should be borne in mind that in full democracies Congress is one of the key decision-makers within the budgetary cycle operating as a counterweight to the executive. It also has the authority<sup>261</sup> to assess public policies, programmes and the reports of results sent by government agencies overseeing the effective use of evaluations.

At this point, the relationship between the executive and legislature seems to be the most important shaping the take up of recommendations with a significant effect over time. This is particularly the case when considering the legislature's commitment to make use of evaluations, e.g. Congress people advocate policy evaluation at first but disengaged later to the systematic evaluation process. For instance, in Mexico, concerning the PbR/SED it was expected that after the 2011 draft budget (see the case study chapter) when evaluations were considered as part of the drafting process, the subsequent years would follow a similar pattern in terms of including more programmes in the analysis. However, there has been no equivalent exercise since 2011 and the Mexican Congress has become a passive actor, even though in 2006 it was a key promoter of the PbR/SED strategy and had competence to reinforce the role of policy evaluation<sup>262</sup> at the time of the strategy's formulation.<sup>263</sup>

Indeed, regarding the question of *whether policy evaluation matters for the public policy and budgetary process*, there are further reasons to affirm that evaluation has

<sup>261</sup> Internal organisation manual of Deputy Chamber and Senators at Congress, official journal, 24/12/2010 and 4/06/2010, respectively.

<sup>262</sup> During Budgeting process (PPEF) 2013, deputies such as those from the 'PRD' party stressed the importance of efficient evaluation mechanisms and strategic expenditure analysis. See daily debates 20/12/2012 at the Mexican Congress.

<sup>263</sup> See the statement of purpose of the draft budget 2006, p. 45. The Mexican Congress.

not played a formal role in the decision-making process. This analysis found that both countries' legislatures did not use evaluation results to improve policies or to reallocate public resources to prioritize sectors such as education, employment, health or public security except for a short period in 2014 by the Chilean Congress (see the case study chapter).

For instance, the table 3 annexed has shown that from 2007 to 2014 all programmes of the Mexican health sector increased their resources every year even in programmes with major recommendations related to the process e.g. to optimize the outpatient services according to the capacity of the hospital of the programme called *health services provided across different medical sites*; except for the *programmes to vulnerable people and families, and childcare centers for supporting working mothers*, which are classified as *status quo* due to low level of achievement of recommendations.

Regarding the Chilean health sector, as table 6 annexed shown all the resources of the programmes evaluated from 2007 to 2014 were increased every year except by two programmes: *bono AUGE* and the *national programme of immunizations* both belonging to the *status quo* category due to low level of achievement of recommendations. However, there are more cases located in this category of *status quo*, which were not considered and affected by politicians during budgetary discussions.<sup>264</sup> Instead, it was found that in this decision-making process the use of informal rules to allocate resources prevails. The overall trend in funding for the health sector has increased over time regardless of bad performance of programmes identified by evaluations.

As mentioned in the studies of evaluation for the Chilean and Mexican health programmes, although the availability, timing, reliability and quality of data and funds – elements to measure the IGR – are settled to support the operation of the policy evaluation, the decision to lead this process rests with those actors who are interested in acknowledging the performance of government. For example, in Chile, during the 2000s when the reliance on evaluation was enhanced by a shared commitment between politicians at Congress and officials inside the executive. However, the strong coordination between actors which underpinned those efforts was fragmented.

Another significant finding from the interviews with politicians is that evaluations are not used for budgeting purposes. In the Mexican case, some of them prefer to use data produced by the Superior Audit Office of the Federation (Auditoría Superior de la Federación 'ASF') and Centre of Public Finances – both belonging to the legislature –, as common users rather than as coordinators and decision makers of the evaluations developed. This shows the limited use and coordination of the policy evaluation by Congress, as well as its limited autonomy as watchdog. In a full democracy context, Congress would have the autonomy to prevent misuse of evaluations giving it freedom from the executive to impose sanctions on policy operators who misused them including the MoF.

By contrast, the flawed democracy prevailing in both countries empowered the role of the president and the executive, who have had a strong influence in this process

<sup>264</sup> It is important to mention that the programmes of the health sector revised in both countries are those that developed an evaluation. Thus, maybe there are more programmes increasing or reducing resources every year but these are not considered in the analysis.

controlling the results of evaluations through the MoF, despite few attempts to remove the total control over the budget held by this ministry such as in the Mexican case.<sup>265</sup> Moreover, Congress has not operated as counterbalance to the excessive power of the executive and has not overseen the achievement of recommendations per programme as it should do. In the end, what prevails in these flawed democracies is the way the IGR are arranged, which exposes the political agenda of the party in power.

In both countries, when the opposition party came to power, the political agenda was focused on developing the Results-based Budgeting strategy. Indeed, both countries' legislatures have been committed to the evaluation process through the approval of the budget presented by the MoF, which responds effectively to the attempts of the executive to operate the strategy. However, despite policy advocacy by Congress people and bureaucracy when evaluations were needed to lead changes in policy sector or specific programmes there was no strong evidence of utilisation by these actors nor opposition parties over time. This understanding explains why many followers including Congress in both countries encouraged evaluation at the beginning but currently reflect on the atomisation of the various actors due to many parties in power with different interests regarding the purpose of evaluation. It might be that these actors have different values, perspectives and above all priorities from those advocating the policy (Hogwood & Gunn, 1984: 208).

Based on these results, perhaps one reason for the limited use of evaluations is that findings disrupt the existing context, which inhabits human, social and political relationships (Abma, 2006: 186) and might cause political instability. For example, the *programme of newborn health support* run by the Chilean Ministry of health, which after the evaluation taken in 2011 was canceled and merged into the *programme of bio psychosocial support*. Although the evaluation report did not mention it, this action might have generated tension and political instability inside the ministry due to a reduction in the budget and a downsizing of the bureaucracy. As Radin (2007: 370) states "depending on where one sits, evaluation can be looked at as a management tool which is necessary for intelligent decision-making or as an unwarranted intrusion on management discretion".

Although, comparatively there is more use of evaluations by the executive than the legislature it is not sufficient to assure that effectiveness of policy evaluation prevails over time. One reason to explain these asymmetries is perhaps that programmes' operators are enforced by the MoF to use evaluations, whereas Congress members are not obliged to oversee the effectiveness because there is not an upper body, agency or even citizenry exerting some pressure over them to do so. Therefore, to respond to an important question for this research, it can be affirm that policy evaluation rarely matters for closing the public policy and budgetary process.

Regarding the conditions required to make policy evaluation effective, it is clear that the *quality of bureaucracy* and skilled personnel operating the policy is closely linked to the IGR. For the purpose of evaluation it is necessary to recruit competitively cadres to lead the evaluation process. Indeed, the results of the indicator of government effectiveness, which measures perceptions of the quality of bureaucracy,

---

<sup>265</sup> Deputy Ricardo Monreal of 'Movimiento Ciudadano' party sent an initiative of law to amend article 23 of the Federal Law of Budget and Fiscal Responsibility 'LFPRH' to establish some boundaries to the SHCP regarding reallocation of public resources once budget is approved, reducing discretionary and quality of budget cycle. However, this initiative was rejected later on. See daily debates 15/11/2012 and 30/04/2013 at the Mexican Congress.

issued by the WGI places both countries above the average (see annex 1). However, according to the case study chapters, there is a significant gap between them. For instance, the Chilean government scores 86 while the Mexican average is 61. In the case of Chile, the quality of bureaucracy seems relevant for that purpose but its influence depends upon the time under review.

The time when Marcel was leading the strategy (see case study chapter) from 2000-06 under the DIPRES (MoF) was when the commitment to evaluation was strongest. It is important to bear in mind that this study does not include that period in the analysis but the DIPRES' study mentioned in the case study chapter noted that during that time, 91% of the recommendations led to some changes in the programmes. This happened thanks to the experience, knowledge and leadership of a group of officials, who were decisive in promoting evaluation. The top Chilean bureaucracy, whose educational background and political capital developed in the private sector helped to put the evaluation agenda in the spotlight. Moreover, they were able to develop a consensus with legislators who advocated this policy (indicating the importance of good intragovernmental relations). In these circumstances, the Chilean evaluation system emerged as one of the strongest and most recognized monitoring and evaluating systems in the world, and became influential to other countries such as Mexico.

The subsequent period from 2007 to 2014, revealed different outcomes. Although this analysis underlines the high standards of the Chilean civil service compared with other Latin American countries (see table 10), the taking-up of recommendations has been influenced "slightly" positively by the bureaucracy expertise. In other words, though the experienced Chilean bureaucracy has acknowledged the importance of taking up the recommendations there is no effect on programmes and many of the recommendations have not been taken yet as shown the results of evaluations performed between 2007 and 2014 (see chart 3, table 6 and 10). Thus, the fact of having experienced cadres such as the Chilean bureaucracy has not guaranteed the effective use of the programmes' recommendations over time.

According to the case study chapters and annex 1, the quality of bureaucracy is an outcome observed across and within countries over time but the effectiveness of policy evaluation is constrained to a specific period of time. This raises the question why there has been no continuity between the quality of officials and the effectiveness of evaluation in the later period? A conclusion is that those executive entrepreneurs led by Marcel from 2000-06 had advocated public agenda including the effectiveness of evaluation – despite misutilisation by Congress –, because the executive was backed up by the president (see case study chapter). Thus, this isolated time period of the Chilean policy evaluation seems to be relevant thanks to the commitment of a few of the high-bureaucracy rather than a systematic process acknowledged by the bureaucracy.

Regarding the difference between the Chilean and Mexican bureaucracy, there is another factor interlinked between these countries called *democracy*. In this sense, it was expected that these emergent democracies had developed a strong civil service over time to fulfil the needs of evaluation but the findings suggest something different. The Chilean *democracy* started earlier in 1990 and the country was aware and interested in the modernization of the State. The reconstruction of democratic institutions needed top executives leading the modernization agenda. A group of qualified officials supported by the president took that challenge and reorganized

the strategic public services including policy evaluation. Conversely, the Mexican *democracy* began in 2000 and despite the government boosted the career civil service since 2003, this did not ensure that personnel become more skilled as noted by Guerrero (2007) and Pardo (2005; 2007). Nor was a leading group of qualified officials identified to implement the reform of the State.

Moreover, comparing this variable with others, the quality of bureaucracy matters to the extent that other actors inside Congress and executive work in partnership (IGR) to ensure evaluation utilisation not only to improve public policy but to close budgetary cycle. However, despite the years of experience in evaluation topics of the people interviewed, neither in the executive nor legislative branches was it possible to establish the experienced cadres necessary to assure the effectiveness of policy. Comparing both countries, the Chilean bureaucracy is supposed to be more effective than the Mexican regarding coordinating and the effectiveness of evaluations. Indeed, this factor is a key question formulated at the beginning of this research. However, it seems there is no difference between having professional cadres leading the evaluation, except during the period ruled by Marcel from 2000-06.

Given these results, it is difficult to conclude that the quality of bureaucracy prevails over other factors in which evaluations take place, except in the 2000s in the Chilean administration due to the reasons already mentioned. One reason for this might be that in developing economies a sort of “elite” and “oppressive” bureaucracy could emerge (Goodnow, LaPalombara, cited on Heady, 2000: 498-9), whose power could hinder the development of democratic policies (i.e. evaluation) despite the commitment of the qualified bureaucracy. As examined here, the strong influence by the presidents – in both countries – and the centralization of government transform the decision-making process into political negotiations by those involved in the policy and those operating the programmes. As Riggs (cited on Heady, 2000: 497) affirms, “the bureaucrats tend to appropriate of the political role”.<sup>266</sup> For instance, the evolution of evaluation in both countries shows that the control and power of bureaucracies has enhanced it from the centre of the executive.

This occurs mainly because the executive is where the resources are provided (DIPRES, SHCP) and political decisions are protected either by this “elite” bureaucracy, the chief of the executive or both. In the case that those responsible for decision-making recognize evaluation as necessary, the process become more effective or, if the contrary, they could block any effort when considering unnecessary or harmful. Here, it seems that such bureaucracy is only interested in the development of evaluations and the political discourse of accountability but not in the effectiveness of such policy over time. In his study of the bureaucracies and political systems “Public Administration. A comparative perspective”, Heady (2000) agrees with Riggs and his fieldwork in Thailand arguing that “transitional” societies – similar to the Chilean and Mexican societies – seldom find a balance between political institutions and the structure of bureaucracy.

Finding the balance between the political and bureaucratic power becomes a difficult task given the excessive authority exerted by the president through this elite bureaucracy e.g. MoF, which is even more powerful than Congress. It might be as Heady (2000: 518-9) suggests that the lack of balance is the result of the

<sup>266</sup> In Almond-Coleman and Banks-Textor's (cited on Heady, 2000: 512) study, they noted that 19 of the 27 developing country bureaucracies analysed had had an excessive participation in political roles, whereas in developed ones, only 6 of the 30 analysed did so.

bureaucracy heritage since independence, and the pressure for solving public problems forces the bureaucracy to appropriate a political role. But it may also be due to the fact of being emergent democracies, which have yet to differentiate between the roles of the executive and legislature. Particularly, the role of the legislature which should operate as a counterweight to the executive acting independently as a society's watchdog of the State's activity, and concerning evaluation, encouraging its utilisation (effectiveness).

Another assumption may be that there is not a merit-based civil service ensuring the officials' commitment to the policy. This is mainly due to the centralized regime, which affects the implementation of an impartial civil service and the non-separation of the legislative branch from the presidency. These issues are located at the micro level (Jordan et al., 2008: 24) where obstacles are related to bureaucracy and how inexperience could delay the effective implementation of policy-making and utilisation of results.

These issues also transform decision-making into political decisions regardless the quality of bureaucracy. Therefore, any attempt to improve public policy has to bear in mind the human factor. It might be as Arellano (2000: 137) suggests the civil service is more linked to the establishment of rigid or semi-rigid criteria rather than the assumption of improvement of organisations' efficiency. Hence, the quality of bureaucracy is one factor in the decision-making process on the use of evaluation but it is not the decisive one.

This raises the question of the influence of democracy on effectiveness. The case study chapters characterised both countries as flawed democracies. The transition to a full democracy requires the strengthening of transparency and accountability and a mechanism of checks and balances. The shift to a more democratic system of governance needs an effective mechanism of evaluation essential to the process of change. However, as a result of these flawed democracies (see annex 1) what has been found is the introduction of policies such as the results-based budgeting mixed with the traditional approach of taking decisions (i.e. decisions enforced by the President) and retaining power through the control of public resources. This process has taken place in an uneven way and is contingent upon institutional factors such as those mentioned here.

It could be argued that, while a democratic state has not been a pre-requisite for the establishment of a system of evaluation, it has been for its effectiveness. To illustrate this, it is worth noting that in Mexico there were several efforts to introduce evaluation mechanisms in the past – even before the democratic elections of 2000 –, whereas in Chile, there were a few attempts during the military regime to establish them. However, only in a democratic context, were the results of evaluation considered for improving the way of governing through better policies and programmes. In such democratic contexts, those responsible for programmes are aware of the consequences when not taking action based on the recommendations. Besides that, an effective counterbalance system by the legislature should operate along with the oversight provided by the NGOs, academics and citizenry. The latter should represent the demand-driven approach noted by Weisner (2011), whose interest claim for effectiveness of evaluation.

In spite of the fact that both countries are located in the same category of flawed democracies, there are several differences between them. For instance, regarding the



functioning of government, the democratic structures in Chile are stronger than those in the Mexican case. The Chilean government scores 7.8, whereas Mexico obtained 6.68 (see the case study chapters). This means that even within the same category, the effectiveness of democracy varies across countries and political systems. However, despite the interest placed in creating conditions to operate evaluation as a system similar to any democratic State, the asymmetries and some gaps in the operation of the policy are visible. The failure to make use of evaluations could be the result of the weaknesses of the government in a number of respects. These could be in the way that policy is conducted, in the system of checks and balances or in the operation of the legislature.

One reason for the failure to use evaluations is that there is no mechanism to encourage these practices nor to impose sanctions where evaluations are misused or government funds are misappropriated. In spite of the efforts to bring evaluation onto the government agenda and to achieve effectiveness, the shadow of presidentialism prevails over public action. This might appear to be contradictory because in both cases, the emphasis on evaluation was driven by the government's agenda rather than by pressure from citizens. Indeed, it could be said that the success of policy evaluation depends upon the president's interest. For example, in Mexico the Presidential Goals System carried out during the presidency of Fox (2000-06) was successful because it was directly monitored by him as was the implementation of the PbR/SED.

The fact is that such governments in transition to democracy tend to favour presidential power over the others branches of government without any opposition by the legislature. Regarding the question *Under what circumstances decision-makers are likely to use evaluation results?* it seems that evaluation is effective, at least in these countries, to the extent that it is important for the president. An example of this is the Chilean case where the president was able to negotiate with the legislature and to make evaluation a priority on the government's administrative reform agenda. Yet the incipient democracy has failed to lay the foundations of a comprehensive policy evaluation. What is found instead is a weak *policy evaluation framework*, which is not harmonized with other evaluation and monitoring systems. Also, the prevalence of presidentialism in these fragile democracies has not provided *autonomy* to the *bureaucracy* coordinating the policy; indeed, they need the president's approval through the MoF when major changes in programmes are necessary.

The role of Congress is also weak and ambiguous when it should be a counterbalance of the executive promoting the effectiveness. Certainly, Congress members should review the achievement of these recommendations across time to verify whether or not the evaluations have improved programmes. Whereas there are some examples of the effectiveness of evaluation in some of the programmes reviewed, being a "flawed" democracy would be a constant disruption of such effectiveness every presidential term, in which those advocating the policy depend upon the president's interest on it.

Regarding whether the *framework of policy evaluation* has an influence on its effectiveness, it is important to bear in mind the vision of an integral system. Ideally, a coherent framework requires a feasible and flexible system to be implemented by operators and users. In these countries there appears to be an imbalance between the design and operationalization of the policy evaluation. Concerning the indicator

that investigates connection between policy evaluation outputs with outcomes, although in Chile there is a balance of the policy evaluation at the macro, meso and micro levels (see case study chapter), there is a gap to gather other evaluation systems developed by other agencies at the horizontal level.

Furthermore, in both countries, the indicators measuring overlaps in the policy evaluation framework with similar regulations, and measuring whether policy coordinators hold hearings to find consensus, evidence that such outcomes are absent (see annex 1). For instance, as mentioned in the case study chapters, there are around 12 evaluation and monitoring systems including the PbR/SED operating across the Mexican central government under the responsibility of different actors and agencies, while in the Chilean case there are three evaluation systems operatively disconnected from the main system SECG (see annex 4 and 7). Thus, a myriad of monitoring and evaluation systems and practices result in the marginalization of policy evaluation and their results. This situation clearly reflects the difficulty in linking these systems into a comprehensive model to be used by policy-makers.

In some instances, however, policy operators follow up recommendations and pay more attention to evaluations of those who exert more political control. An example is the Mexican *programme Prospera*, which is the most evaluated in the course of time and controlled by political actors due to the large amount of economic resources invested in it. In such a case, policy-makers are aware of the importance of taking up the recommendations because many actors inside and outside the government oversee the programme. In the case of Mexico, it was also found that the mechanism to follow-up evaluation results issued in 2008 was not clear in terms of which agencies should report the recommendations attained. In 2011, the coordinators amended this regulation, which is compulsory to all public agencies and state enterprises operating budgetary programmes and those who have performed any type of evaluation during 2008 and afterwards.

In the meantime, there were some programmes which recommendations were not attained due to an ambiguous framework. For example, after the evaluation of the programme *Health services provided across different medical sites* took place, the agency replied to policy evaluation coordinators (CONEVAL, SFP) that this programme was not subject to the mechanism to follow up evaluation results and did not implement any suggestions made by evaluators. Conversely, the programme of *provision of healthcare services* implemented the recommendations even though the agency replied that the programme was not subject to the mechanism to follow up evaluation results.

In other cases, there are no evaluation reports available to verify the achievement of recommendations e.g. the programme *for protection and integrated development of childhood*, or there are differences between what evaluators suggested and actions taken by the agencies such as the programmes *of prevention against obesity* and the *national supply centre*. In these cases, the coordinators of the policy did not follow up the agencies' results to verify whether or not the recommendations were achieved, nor did they assess the coherence of the recommendations suggested and attained, or the time needed to implement these. Regarding the latter, in Chile, the *experimental health centres* are examples of such gaps in terms of time, i.e. the time when evaluators formulate recommendations and when policy-makers achieved these. In this particular case, two years had passed since the evaluation took place in 2012

until 2014 when the agency issued a report of the results of the recommendations attained.

Another consequence of the fragmented framework is the emergence of internal conflicts due to the diversity of authorities conducting evaluations. Here, the policy coordinators need to devise strategies to ensure a full convergence of all the evaluation systems spread out across public agencies such as the school teacher's performance evaluation, SIMCE and SEA in the Chilean case, as well as the myriad developed by the Mexican government. For this, only top-level officials and politicians inside both branches are able to ensure cohesion and legal harmonization with the policy evaluation.

This process should also clarify the roles and responsibilities of those involved in the evaluation process both internally and externally. On the external side, the inclusion of interest groups such as the NGOs and academics, provides a better understanding of this modernization and democratic governance trend, contributing to the on-going process of representing citizenry and widespread social accountability. Similarly, a better co-governance between branches might internally simplify the operationalization of evaluation. An option for this simplification is through a strong and single evaluation system with some important implications to the policy. First, it would eliminate the overlapping responsibilities between different authorities. Secondly, the definition of one authority regarding evaluation would reduce internal conflicts between policy-makers and coordinators. Thirdly, a feasible system would change the puzzling image of the policy framework design. In the end, these features seek to potentialise the effect of evaluation inside (micro, meso and macro level) and outside (interest groups) of the government.

Regarding the structure of M&E system, in both countries, the indicators of technical operationalization of the system and methodological rigor indicate that these outcomes have not yet been achieved and still need to attain a higher technical level of sophistication to gather the large amount of data needed by many actors across agencies. In the Mexican case, the range of systems brings practical difficulties for technical operationalization and monitoring purposes, whereas in Chile, the IT platform needs further improvements to facilitate its operation. However, the perception of operationalization by actors is more positive in the Chilean case compared to the Mexican. As mentioned in the case study chapter, officials and politicians surveyed point out to a difficult, rigid and unfriendly system.

Regarding the rigor of the system, there are problems in both countries in terms of the order in which evaluations should be performed (see annex 1). An example is when policy coordinators determine to execute an evaluation of process without considering the need to perform an evaluation of programme' design in the first place. As mentioned earlier, this is the result of the lax oversight of the policy by such coordinators as well as the lack of sanctions to those who misused evaluations. These discrepancies within the framework hinder the operationalization of the system and reduce the possibility to make effective use of evaluations.

The results attributable to the evaluation system's operationalization provide arguments to affirm that the *framework of policy evaluation* is still ambiguous and complex, which makes the evaluation less effective. Nevertheless, the relevance of this factor does not seem to be a determinant of the effectiveness itself. The fact is that the design of the policy as the basis for a coherent evaluation system should be

connected to the efforts of the *bureaucracy* to improve the coordination (IGR) which would ensure that evaluation functioned as a single process, and contributed to the implementation of recommendations.

Lastly, the variable of *autonomy* is understood as the ability of an agency in terms of its competences, powers and functions to operate independently from a hierarchical supervisory body (executive); as Christensen (1999) notes, exempted from full political supervision by the central head of authority. Nevertheless, it is also a notion that varies relying on the approach of those involved. On the one hand, for those inside the executive, autonomy is perceived as the “freedom” to decide whether to conduct evaluation or to use external evaluators. On the other hand, from an external view, autonomy relates to the independence of the agency running the policy evaluation from those who will be assessed, that is to say from the executive (DIPRES and SHCP/CONEVAL in Chile and Mexico, respectively). Nevertheless, in these cases, the autonomy of evaluation is restricted to the selection of the evaluator in performing such evaluations but not to the agency coordinating them.

In both countries, either the evaluation agency or agencies coordinating the policy evaluation have low capacity regarding policy, financial and legal autonomy toward pursuing policy evaluation goals (see annex 1). That is, the results show that there is limited autonomy granted to these agencies to decide themselves what matters are important (Verhoest et al., 2004) exempted from central supervision (Christensen, 1999). In spite of positive improvements such as the emergence of “independent” agencies in Mexico such as the CONEVAL or INEE to build credibility in its institutions, these are not sufficient for the purpose of autonomy and still operate under supervision of the Ministry of Finance ‘SHCP’ (legal), partly funded by central government (financial) and do not decide on individual applications of general rules (policy).

In the Chilean case, the autonomy of agencies is currently constrained to the hiring process of external evaluators by the Ministry of Finance ‘DIPRES’. While the executive’s aim is to create the National Evaluation Office to coordinate policy evaluation independently of central government, the top official responsible for this office will be appointed by the executive threatening its autonomy because the executive plays a double role of judge and jury (see case study chapter). As noted in the methodology chapter, the credibility of evaluation results depends upon the extent to which the agency that coordinates evaluation is operationally independent from the executive branch to prevent the coordinating agency from playing the roles of judge and jury. The state of the art of autonomy is one in which policy and decision-makers, users and actors involved (stakeholders) are independently willing to use and exchange data (evaluations) for further decisions.

However, regarding one of the questions established in this research *The policy evaluation is effective to the extent that the coordinating agency keeps some control through political power*, despite this seems contradictory and the excessive control clashes with the vision of a democratic government and autonomy of public agencies, who wants to implement an effective system of check and balances. The historical analysis and hierarchical bureaucracies seem to confirm that the authority pursued by the government (executive) becomes necessary to exert some control and political power over the rest to influence the taking-up of the recommendations and making evaluations effective. For Hojlund (2014: 39) this is a coercive adoption mode, in

which organisational rationality is not a priori and agencies are forced upon by one or more organisations in order to implement evaluation results.

The autonomy of the evaluating agency raises questions of the degree of political control and the role of the executive. As implied before, at least at the beginning of the strategy, the adhesion of most of the public agencies to the PbR/SED (Mexico) and the SECG (Chile) was because the budget takes priority and the Ministry of Finance (SHCP and DIPRES) led the strategy. Maybe the reason for this was due to the historical subordination of agencies and bureaucracy to the president and MoF, which prevails to some extent in transitional democracies. However, this subordination did not last long to strengthen the use of evaluations in the course of time. Thus, the autonomy of coordinators of the policy is needed to develop evaluations in terms of quality and objectivity amongst others, but also to guarantee the effectiveness of evaluation.

However, a potential risk threatening the autonomy of the agency coordinating policy evaluation is when this is located within the executive, which powers and functions are restricted for demanding the effectiveness of evaluation from the judicial and legislative branches. Therefore, the use of evaluations is bounded to public agencies of the executive. At the end, actions needed should be similar to the statement noted in this research that *the more autonomous those who are responsible for policy evaluation, the more credible and effective the evaluation*.

Moreover, the autonomy is more likely to be shaped by other factors such as a full *democracy*, which are closely related and make a difference if these were entirely operated as they should. In this sense, a full democracy agrees those coordinators of policy to be autonomous from the decisions of the executive improving the relationship (IGR) between branches elucidating the roles and responsibilities of those involved in the evaluation process both internally and externally. For example, if these coordinators were from an autonomous Congress could oversee the achievement of recommendations even to impose sanctions to those who misuse evaluations.

Studying results across countries aims drawing some conclusions to address the central question of *whether policy evaluation is effective across these countries?* As noted in the methodology chapter, the Chilean policy evaluation is known worldwide as one of the strongest and most recognized M&E systems. Thus, it was expected that the former was more effective than the Mexican and the outcomes were constant over time. However, after comparing institutional factors to establish a causal relationship among them, a first conclusion is that policy evaluation is not effective across and within these countries in the course of time. If all country settings such as the democratic context and institutional factors are settled to operate, why is policy evaluation not effective? Presumably all these factors are not effectively working together.

This leads the discussion to the political factor, in which the IGR are the constant factor either to accomplish or not the effectiveness of evaluations performed, as many of the programmes revised here, which despite the complexity of the tasks (recommendations) and financial restrictions were solved thanks to the commitment of programmes' operators e.g. *national programme of immunizations* and *of treatment of obesity*. Moreover, when it is referred that politics matters is because a significant number of recommendations attained respond to a political factor which

inhabits in the decision made by actors, who have the power either to achieve or not the effectiveness of evaluation.

Even though other factors such as democracy, autonomy of evaluators, policy framework and quality of the bureaucracy *ceteris paribus* or influence the evaluation process in some way, the IGR are the factor on which they differ and occurs as a constant circumstance in many programmes. This seems to be the likely cause or at least part of the cause that remains as decisive and immovable to attain effectiveness. Therefore, the power to impact any programme through evaluation is on the extent those actors are committed to use evaluations to bring about changes in policies. Moreover, a real interest to achieve the effectiveness of evaluation is evident to the extent in which it would be a priority to look at.

## CHAPTER 8. Conclusions

The study has set out to explore the effectiveness of policy evaluation in Chile and Mexico, drawing upon the theoretical debates on the nature and purpose of evaluation, particularly with regard to its contribution in bringing about changes in policies and programmes. Central to the research is to identify the institutional factors which are likely to influence whether evaluation is effective and the circumstances under which this occurs. It also attempts to explain the differences in effectiveness between these countries, recapitulating the main empirical findings presented in the case study chapters, and the institutional factors examined in the comparative chapter.

In principle, the evaluation of policy encourages the notion that findings will determine some specific actions to avoid their misuse or symbolic use.<sup>267</sup> This means, avoiding the *status quo* and finding out whether or not the programme has addressed its objectives. Evaluation will also lead to actors' involvement in any changes in programmes and policies. However, through the evidence obtained, it can be affirmed that *evaluation is of limited value in supporting decision-making*. This slight effectiveness of evaluation evidences a gap between the expectation that evaluation should create public value in the form of better policies and the extent of change in policies and programmes once they have been assessed. The study of the health sector programmes assessed from 2007 to 2014 shows that the effectiveness of evaluation was achieved in the Mexican programmes in 33%, whereas in the Chilean programmes was of 31%.

Thus, most of the Mexican (67%) and Chilean (69%) programmes evaluated have made little changes and classified as *status quo*, or were cancelled without evidence that evaluation was considered in such decisions. Therefore, *policy evaluation rarely matters for closing the public policy and budgetary process*. When utilisation of evaluations was further explored in the comparative chapter to elucidate which country has been more effective and whether the achievement is linked either to the complexity of the task or the costs of implementing changes, the findings are the following. Comparatively, there is a slight difference between recommendations achieved in Mexico (see table 12), where most changes related to the design of the programme, while in the Chilean health sector, the majority of the recommendations achieved related to process.

As noted, recommendations concerning design are largely desk-based and less complex tasks which might be easier to solve compared to those which are more costly to finance. It cannot be asserted that this is the pattern followed by both countries given results of the Chilean programmes (see table 12), in which policy makers solved most of the recommendations classified as hard tasks due to the costs of implementing changes e.g. time, coordination, finances, infrastructure and human resources. But compared to Mexico, Chile has achieved slight less number of recommendations. Therefore, regardless of the complexity of the task and the issue of finance, the interest in solving recommendations arises from the interest of the actors involved in those programmes evaluated to make effective decisions for closing the policy and budgetary cycle.

---

<sup>267</sup> The symbolic use means that evaluation could be used only to justify a pre-existing position of an agency (Mark & Henry, 2004: 36).

To verify whether evaluation is effective it is necessary to grasp the intricacies of political factors and other country settings. In this sense, this chapter argues that evaluation often fails to translate into action on the part of the policy makers; instead there is a noticeable gap in the decision-making process even if, as Hill (2009: 8, 278) notes the idea that “the policy process should involve the explicit identification of objectives, translated as effectively as possible into action and thus susceptible to evaluation afterwards dies hard”. This gap between evaluation and change is seen in the analysis of the health sector process where there has been a limited take-up of findings over time. Although, stakeholders acquired evaluations political factors minimized their effectiveness. This is in line with Mayntz’s (1994: 80-4) argument that the political factor is implicit in government decisions including policy design, even dominating over technical or social aspects of the policy.

This sceptical view about the extent to which evaluation is effective and the narrow action by policy-makers were noted by scholars in the review of the literature (Dahler-Larsen, 2005, Dye, 2011 and Hill, 2009) which highlight the limited connection between findings and policy improvement. As mentioned in previous chapters, evaluation could be a two edged sword. On the one hand, it promotes efficiency, transparency and accountability but on the other hand, it exposes corruption, inefficiency and “bureaucratic” government procedures to external scrutiny. This implies political risks that not everyone wants to assume especially when it refers to programmes, which represent votes, even when findings suggest major changes or cancellation.

These political risks are also related to the most pressing issues for the academic literature such as the subtle nuances of ethics, intended and unintended influence, beliefs, values, rational and subjective, as well as decisions and nondecisions (Askim, 2007, Bachrach & Baratz, 1963, Christie, 2009, Dahler-Larsen, 2000, Fleischer & Ledermann, 2011, Kirkhart, 2000, Teirlinck et al., 2013 and Weiss, 1979), which regularly transform the decision-making into a political process.

Regarding those factors fostering the effectiveness, although it is difficult to have control of all factors, it should bear in mind which elements are the most important. After the review of all the variables mentioned in the chapter ‘Broadening understandings of policy evaluation and effectiveness’ such as the evaluation capacity building, the role of evaluator, stakeholders’ involvement, and intentionality (Fleischer & Christie, 2009), or the organisational context, the political context and the intentions of the evaluation demand argued by Teirlinck et al. (2013: 374), amongst others, institutional factors such as the quality of bureaucracy, level of democracy, policy evaluation framework, the degree of autonomy and intragovernmental relations were selected as the most important helping to explain the differences in the effectiveness of policy evaluation across Chile and Mexico.

As earlier mentioned, these factors embrace those relatively under-explored variables addressing the question of *whether are the most influential institutional (political) factors shaping the effectiveness of evaluation?* Comparing these factors between countries and based on Mill’s method, the only outcome present in both countries is the quality of bureaucracy. Particularly, in Chile during the 90s, the quality of the bureaucracy operating the policy evaluation was the determinant factor shaping effectiveness, as it was concluded in the chapter ‘The effectiveness of policy evaluation in Chile’. In this case, the presence of skilled personnel in the decision-making process would help to achieve recommendations more effectively, because



these people would understand the role of evaluation in improving policy. However, based on the rate of effectiveness achieved in programmes or policies, this variable is not sufficient to assume it is the only cause factor shaping the impact of evaluation on policy change.

In exploring the influence of these factors, the Intra Governmental and Interagency Relations (IGR) are the condition either present or absent during the decision-making process and for the purpose of utilisation. The element necessary to transform the decision-making is the action of and relationship between those actors implicit in this process. This variable is highly useful in the evaluation process as influences the final decision whether or not to take up recommendations and make further changes to programmes and policies. Given this, the IGR are understood as the set of decisions and nondecisions made by actors, who should be seeking for better coordination and commitment in favour of evaluation. Indeed, in many agencies' statement the relevance of the IGR is highlighted as a factor in addressing issues raised by evaluations. In particular, these refer to the need of coordination and synergies between areas, programmes, staff, systems, agencies and branches.

As mentioned in the comparative chapter, most of the evidence shows weak forms of coordination, mainly between actors within the executive and legislature in which, the former make little use of evaluations and fulfil requirements based on legal authority at central level by the MoF while members of Congress have not utilised evaluations to bring about changes in policy or to plan the federal Budget. An exception was the FY 2011 in Mexico, when both branches worked together for the first time in the draft budget under the approach of the results-based budgeting. It can therefore be stated that some discrepancies in the styles of ruling relations (IGR) between these branches remain. More commitment by the MoF and Congress is needed for taking-up recommendations as established in the RbB when tailoring and approving the budget, and by public agencies for better decision-making of sector policies.

The significance of the IGR factor depends on the extent to which actors inside branches could work harmoniously (co-governance). However, this is a reminder of the real world of policy, where decision-making, negotiation and chaos occur in political arenas. Although the effectiveness of policy evaluation is perceived valuable by those involved in the decision-making process, as Weiss (1993: 94) noted, policies and programmes are the result of political decisions, have a political stance and enter into political arena to feed decision-making. In this context, politics is the art of government whose power has the ability to achieve a desired outcome (Karlsson & Conner, 2006: 227). Whereas for Dahl (1984: 9-10 cited on Karlsson & Conner, 2006: 226) defines it as a pattern of control, power or authority influencing relationships.

The mere act of engaging in evaluation constitutes a set of political decisions. It is not a neutral political fact aside from power, even though evaluators try to avoid it (Monnier, 1996: 173). In the end, evaluation is the judgement of data, to value premises, intimately linked to ethical issues in a political context. In practice, data could be biased in favour of particular interests of policy-makers. Hence, there is a negative connotation of politics and how evaluation becomes adulterated when both are mixed and the evaluator adopts a standpoint – neutral, sensitive or critical – (Karlsson & Conner, 2006: 225, 231-7). This works in a similar way to Simons' (2006: 243) approach on "ethical dilemmas" where the decision emerges from the

evaluation, and the choice taken between alternatives or solutions depends upon the ethical view of policy makers.

As much as the results provided by evaluations could be interesting and relevant, the decision by which evaluation leads to changes in policy, programme, plan, project or even personnel often relies on political arenas. Thus, after testing variables one conclusion is that *the political factor in the IGR is decisive for the effectiveness of policy evaluation*, including the RbB strategy in both countries. After reviewing historical processes of evaluation in emerging economies, this is similar to Weiss's arguments (1976, 1987 cited in Pawson & Tilley, 1997: 12), when refers to political decisions as gradual accretions of choices that sometimes appear as a political U-turn.

The best explanation for this assumption is before and after the fiscal budget 2011,<sup>268</sup> where the debate around budget at the Mexican Congress clearly lacked of any criteria based on results to allocate (decrease or increase) public resources. As result, the budget was tailored based on political factors, on the ability of deputies of every commission to defend the budget items of specific sectors, programmes or projects. These arrangements appear to be a form of political "libertinaje" (libertine) with Congress misusing its political power to distribute finances to programmes subjectively. These actions maximise the benefits to particular political party (ies) instead of national priorities. Besides that, policy evaluation coordinators need to call for support from Congress members every administrative period (4 and 3 years in Chile and Mexico, respectively) because they are easily removed from their positions and new actors came in the legislature.

In the end, the strategy called results (based, for, of) budgeting should operate as an incentive that every country – depending upon the context, regime and institutional capabilities – implement or boost through its government agencies to ensure better results. Whether policy evaluation matters not only for the budgetary process but also for the public policy, hence, effectiveness is widely acknowledged for closing this process helping to create public value. However, in the Mexican case, the majority of programmes assessed (67%) have not achieved yet the evaluation results for improving them and closing the policy cycle, these figures include cancelation of programmes. Regarding the budgetary cycle, the review from 2007 to 2014 shows that only in 2011, the draft budget was formulated based on results (see case study chapter).

Similarly, the Chilean policy evaluation has almost had the same results as concluded in the case study chapter. Many of the sampled programmes assessed (69%) have not implemented the recommendations suggested by external evaluators yet, neither, evidence of the use of evaluations of the health sector for the budgetary process during the same period analysed. As concluded in previous chapters, there is a trend to increase public resources every year with no implications of evaluations as a tool to support decision-making. In this sense, an idea that merits consideration is that policy evaluation will be more efficient when using the budget as incentive (positive or negative) to guarantee evaluations results are taken into account in the policy process.

For scholars as Wiesner (2002: 138), there are reasons to assume that the key to develop evaluation capacity (demand-led) is to assure a linkage between

<sup>268</sup> As mentioned, only in the 2011 budgetary process (purpose statement), the SHCP sent to Congress the budgeting proposal based on objective data from evaluations in order to allocate public resources.

evaluation results and budget allocations. Given the extent to which evaluation results do not affect the budget allocation process, it will be difficult to develop real evaluation which is demand-led and reliable. Evaluation can hardly be boosted itself, it needs incentives such as the budget (RbB and PbR/SED) to be adopted homogenously across public agencies, as Wiesner (2002: 146) argues,

Given importance of institutional factors, evaluation will be boosted on the extent on which has a solid regulatory backup, and closer to budgetary process. The key for developing evaluation capacity is when it comes from the demand [citizenry]...

Therefore, budgeting should not be seen as part of the evaluation process; it is only an outcome of the policy process and should be considered part of an incentives structure and not a purpose itself. However, it is important to bear in mind a concern in the Mexican case. If evaluation takes the PbR/SED as incentive, it is likely to contradict objectives enacted by the Constitution, at least those related to the social sector. According to the first article "all human rights and guarantees for their protection noted in this regulation and in international trades cannot be restricted or suspended, except when established by the Constitution". Additionally, article 18 of the General law of social development claims that "programmes and public resources allocated to social development cannot be reduced, except when established by Congress when approving the Federal Budget".

Moreover, this law enacted that the amount of finances cannot be less than the previous year (article 20). Hence, there is a potential risk of threatening the Constitution and basic constituencies' rights when incentives are negative as result of evaluation and finances need to be cut back. However, fewer attempts have been made to introduce negative incentives when results are not achieved or evaluations are not developed. There are no consequences to programmes that have not developed the government-ordered evaluations such as the following programmes of the Mexican health sector: *childcare centers for supporting working mothers*, *reduction of diseases*, and *improvement of medical facilities*, amongst others. In the Chilean case, examples of this are the programmes of *experimental health centers* and *the national supply center*. This raises the question of *what coordinators should do when policy implementers fail to arrange for evaluations to be conducted?*

Despite these problems, the evaluation strategy has continued over time, mainly due to the leadership of the MoF. Thus, the reason why evaluation works might be because the ministry responsible for the strategy's coordination is also responsible for the budget management. As mentioned in the key questions, *evaluation is effective on the extent to which the coordinating agency keeps some control through political power and the monopoly of finances*. Certainly, such political control explains the effectiveness of evaluation in some programmes across countries.

From the discussion thus far, it is clear that evaluation is an essential activity influenced by politics, and takes place within a political context. Despite bias being inevitable sometimes, the scenario where it occurs should be acknowledged for closing the policy cycle. In the first place, Knill & Tosun (2012: 175) note a number of factors that complicate the process such as the identification of policy goals, appropriate definition and measurement of performance indicators, isolation of a policy's effects from other factors, and the political context. These should be identified in order to make evaluation more effective.

Secondly, there should be recognition of the stakeholders, bureaucrats, politicians, and non-governmental members of policy subsystems, who deal with government

issues and the policy [evaluation] in question (Howlett & Ramesh, 1995: 168). Vedung (cited in Bundi, 2016: 2) noted that “scholars have completely left the parliament as a demander of evaluations out of discussion, although parliamentarians are an important stakeholder of evaluations”. In line with this, intervention by constituency, researchers, academics, private consultants (consultancy firms), lobbyists, think tanks, non-governmental organisations (NGOs) and professional associations of evaluators<sup>269</sup> as promoters of use of evaluations that advocate collective demands and exert power over policies and their policy-makers should be considered. This is labelled as the most difficult issue to deal with, due to human factors.

Rationality of political decision-making, evaluation process and personnel are interrelated spheres, which depend upon them. Indeed, the findings in Nicholson-Crotty & Miller’s (2012: 365-6) study of bureaucratic effectiveness influencing legislative process shows a positive relation to enhance the policy process. For Knoepfel et al. (2007: 243-9 cited in Hill, 2009: 282) the process represents a particular form of advice, where information and recommendations are used strategically either to support the *status quo*, to attack policy developments or for further changes by governments, and need participation of stakeholders and opponents to achieve them.

In this line, the scope of democratic governments is to take off public issues from the *status quo*, which traditionally remains private, as Cochran & Malone (2005: 45-6) argue “the political system has a bias in favor of the *status quo* and will resist the addition of new issues to the policy agenda”. By contrast, the role played by interest groups mentioned above expands the scope of the programmes into the public arena. These groups are “intimately linked to the extent to which democracy itself consolidates” (Hague & Harrop, 2004: 177), closing gaps between society and State, gaining territory for the transparency of public action, and acting independently from the State. Thus, if governments want to be democratic should assume these groups as a force interacting with the State.

Thirdly, the predominant use of the implementation theory rather than programmatic theory (Weiss, cited in Rogers, 2007: 64) could also be a limitation. The former is based on activities, intermediate – also immediate – outcomes, practitioners’ assumptions and logic models, whereas the latter adopts a theory-based evaluation, outcomes, looking at the quality of the theory, alternative causal explanations and perspectives, and mechanisms of change. This implementation theory might not provide a scientific basis for policy making, which result in bias of the outcomes achieved. An example of this is when the results-based budgeting (RbB) strategy is considered as the core of policy evaluation.

That is to say, when the RbB (Chile) and PbR/SED (Mexico) are considered the core feature of policy evaluation as other countries do, it is clearly assumed the failure (effectiveness) of this policy. This is due to the strategy is limited to the intermediate outcomes and activities related to the budgetary process as Weiss suggested in the implementation theory. Conversely, the programmatic theory or theory-based evaluation should be a comprehensive model including all paths taken by evaluation results and systems operated by other agencies. This should be a model looking for alternative causal explanations of effectiveness based on the State capacity-building

---

<sup>269</sup>For example, the National Academy of Mexican Evaluators (ACEVAL), American Evaluation Association (AEA), and EvalPartners, amongst others.

and administrative reforms; a strategy from a perspective of results-based management with the aim of consolidating a comprehensive policy evaluation.

### **Theoretical contributions and policy implications of the thesis**

This dissertation makes a number of contributions to the debate on policy evaluation. The first of these relates to the political nature of evaluation. The key for better understanding and achieving a successful evaluation process is to bear in mind that politics along with ethical issues could affect any attempt to implement an objective and scientific method. A key finding of the research is that *the political factor in the IGR is decisive for the effectiveness of policy evaluation*. Recognising political constraints is a precondition for usable evaluation research, as Weiss (1993: 94) noted,

The policies and programmes with which evaluation deals are the creatures of political decisions...

Because evaluation is undertaken in order to feed into decision-making, its reports enter the political arena...

And perhaps least recognized, evaluation itself has a political stance...

To understand the effectiveness [or not] of policy evaluation takes the discussion to the public sector reforms in emerging economies and the gap of the State's capability to adopt and implement such policy. As mentioned in the chapter 'Broadening understandings of policy evaluation and effectiveness', the institutional capability is explained in terms of the factors presented here. Significant transformation of the government requires changes in the services and programmes delivered, career civil service, accountability systems, monitoring and evaluation practices, deregulation, transparency and efficiency of public resources. However, evaluation is always disruptive and could modify the stability of the programme and the status *quo* of the public agencies in some way. This is a strong reason why bureaucracy is reluctant to feedback from evaluations. They find evaluation results as something with negative repercussions, which might affect organisational routines.

In developed economies, evaluation did not emerge itself, it needed political support given by administrative reforms. Moreover, the "innovative" ideas of the New Public Management 'NPM' (results-based budgeting, performance evaluation, indicators, incentives) have operated in countries widely transparent, with a solid regulatory framework and civil service (Arellano, 2002: 14). However, the countries analysed here should not be seen as adopting administrative reforms in the same way as Dussauge (2009: 449) and Pollit & Bouckaert (2004: 8-9) claim. The stance of these scholars is similar to Llerenas & Huerta's (2004: 233) opinion, who suggest that this reform was an intermediate *managerialist* strategy,

...it naively assumed that it was possible to arrive at economic, efficient and effective government without changing the nature of elite power relations within the state... This is because the transformation of administration also requires the transformation of politics.

For instance, in the Mexican State, the transformation of the government has been difficult to achieve due to presidential control over the governmental apparatus (Arellano & Guerrero (2003: 2-3). This characteristic prevails in hybrid regimes due to serious weaknesses in the functioning of government, political culture and participation (EIU democracy index). The approach adopted in Mexico is about building capacity providing institutional remedies to do old things in different ways and discover new things that need to be done but cannot eliminate retrograde tendencies (Jones & Kettl, 2003: 12) within the system. Although according to this index both countries have moved to one regime more likely to a flawed democracy characterized by a diverse composition at Congress, competitive elections, less loyalty

to the president and transformation of the government apparatus, some political control still prevails, which is considered as an institutional factor inhibiting the evaluation process.

To balance such political control, the legislative branch should operate as a system of opposition and counterbalance of the executive and as a democratic entity representing constituency.<sup>270</sup> The findings in Bundi's research (2016) of the Swiss parliament confirm the importance of demand for evaluation by the legislature. Indeed, those parliamentarians who request evaluations do so because they want to fulfil their oversight function to achieve an accountable government. However, the reality is this does not work in practice as full democracies such as Switzerland. In these emergent economies, other actors are needed demanding more evaluations' use such as those who belong to the oversight committee inside Congress, the more experienced or in a leading position and those having a positive attitude toward effectiveness. A reason for this is without opposition parties and leaders, society is less able to demand accountability and transparency from the executive.

In this sense, what government in these developing countries has been doing is moving through a transformation, which could fully implement the policy evaluation and promote government's efficiency, a "neutral" reform, as Arellano & Guerrero (2003: 11) claim 'that technically improves government efficiency through administrative techniques'. For others like Cejudo (2008: 123), referring to the NPM experience, the reforms are:

Timid organisational reforms that may alter the managerial jargon and some practices, but do not affect the centralized and hierarchical structure of the public sector, do not grant more autonomy to public agencies, [and] do not base budgetary decisions on performance assessments...

The evidence shows the RbB does not work as it was thought, so the budget is an incentive but not the main objective. Once evaluation arrives at Congress, they (deputies) have the decision-making to close the policy cycle but also could take different pathways. On the one hand, the evaluation results could be taken for better budget decision. On the other hand, they could get rid of the problem or those high-risk<sup>271</sup> programmes through eliminating or reducing their finances. At the end, these are choices made by individuals, whose actions are based on subjective judgments in order to obtain particular benefits. An exercise of comparison between the budgetary process in Chile and Mexico, shows that they oscillate in the middle of the interest group determinism and process of viewing politics of budget (Rubin, 2010: 354-5). In the "determinism" vision, the interests dominating the budget are represented by the weaker or dominant actors.

Another contribution is that in countries with "presidentialism" regime systems such as these, the power is on the side of the dominant party, and these dominant actors are usually close to the president. Indeed, there will be a clash in the budgetary goals if the dominant party is the same of the president representing broad coalitions, leaving aside priorities of the narrowed coalitions who hardly try to favour their budget goals. This is the "process" viewing centered on budget politics, achieving funding to particular programmes through the ability of political actors doing tradeoffs and negotiations of the narrow amount of budget negotiable. Following

---

<sup>270</sup> See the results of the Commission of Citizen Participation at Congress (Cardozo, 2006: 206).

<sup>271</sup> Particularly risky those programmes of the social sector, which might influence in a negative way the political campaign and ballots (votes).

the discussion of the budgetary process, unlike what is commonly known, the executive plays a key role in the RbB strategy and allocation of finances. The MoF is responsible for the budget tailoring making initial arrangements of allocation of resources of each programme.

In this context, the study of Ugalde (2014: 10) is important for the Mexican case. This scholar asserts to say that the available amount of finances or the so-called “bolsa negociable” at Congress for negotiations during the budgetary process is around 6 to 10% of the total budget once the MoF elaborates the draft budget. Thus, most of the negotiations have taken place inside the executive without any transparency of such modifications (increase, cut) in the budget. As this author says “The fact that 90% of the budget is allocated does not mean this is ‘unbudgeable’ but Congress turns the look over the 10% in order to attain their party’s goals and votes”. So, the differential power between actors – budget office, head of ministries, lobbyists, legislators, governors and local authorities – has a different level of power depending upon the party in the power.

Despite the fragmentation of the budget process and political actors, this has not persuaded Congress to use performance data systematically once they receive evaluations by the executive. The empirical findings of this study show that evaluation use has slightly increased at the agency level (executive) compared to the legislative. To some extent, progress can be observed in terms of the strategic planning process e.g. programmes are better aligned to national objectives. Indeed, in both countries most of the recommendations implemented and recurrent over time are related to the design of the programme as indicated in the comparative chapter. Furthermore, some agencies are using evaluation and performance data to make internal management decisions. This is the reason to refer to this strategy as policy evaluation instead of results-based budgeting. Moreover, based on empirical findings, depending on the extent to which evaluation is constrained to the budget process and decisions are limited to the allocation of government grants, the strategy has failed.

Despite a low rate of effectiveness, in a broad perspective, evaluation has been slightly effective because results have been used – marginally – as a deliberate effort to improve programmes. The theoretical arguments of this thesis suggest that the policy evaluation should be linked to incentives such as RbB, considering what Wiesner (2002: 138) suggests, to develop a real demand-led evaluation system.

### **Limitations and recommendations**

As a consequence of the methodology, the scale of this research has been limited in terms of the comparative perspective of the countries reviewed while the focus was also limited to the health sector programmes operating at the central level addressing the question of whether a cross-national comparison of the evaluations of the health sector might be used to infer or generalize findings to other sectors. However, the results offered here only could provide possible explanatory reasons about why some policy evaluations perform better than others. Additionally, the study encountered a limitation regarding the interviews of the Chilean bureaucracy, politicians and academics due to restrictions of resources. To compensate for these limitations, a documentary analysis of evaluation was conducted based on a large database, government material, indexes and international organisations reports with the aim of looking towards a patterned variation in the occurrence of the effectiveness of policy evaluation in countries with similar conditions.

One of the main problems in operationalising “effectiveness” concerned the analysis of the recommendations outlined in the so-called agency statements, which are the only resource publicly available to provide evidence of the extent to which agents have responded to evaluations. However, agency statements are not easily accessible, even in some cases evaluations, in practice, these documents are uploaded in pdf format in a data repository not in a database system as specified in regulations. One of the consequences resulting from this rigid format is that these documents are easily interchangeable with one another with the risk to reformulate recommendations from those reported by evaluators and deadlines previously committed as noted in the case study chapters.

Moreover, the review of results to these working reports and agency statements in pdf format could jeopardize the following-up process. In some way, these issues are the reason to code the subvariable of technical operationalisation of M&E system for both countries as an outcome absent due to these gaps identified. Therefore, what would be needed to have done this better is to improve the M&E system, particularly in terms of the IT platform. This would have facilitated the technical operationalisation of the variable, addressing the problem of missing data and recommendations. It also needs to manage the vast amount of information produced by diverse agencies such as the Mexican M&E system in an only database.

Another remaining gap which could bias the operationalising effectiveness is the acceptance of results by those programmes evaluated. To fill such gap each response by decision-makers was considered in order to classify each recommendation with the criteria for judging the effectiveness. Last but not least important was the decision to operationalise the concept of effectiveness, for which an effective and feasible approach was taken instead of a complex conceptualisation offered by scholars as explained in the methodology chapter, in order to determine whether programme operators use evaluations to lead changes on programmes.

What are the implications of this study for the central problem investigated here *whether policy evaluation is effective across these countries?* The evidence of this study calls into question claims that policy evaluation is effective, highlighting instead a low rate of evaluations utilised in the policy making process. Nonetheless, it is important to look around the context of policy evaluation. From another perspective, there is a slightly better result of effectiveness in the Mexican than the Chilean policy evaluation, which has been considered a model system to be adopted by other emerging economies such as Mexico. Moreover, significant improvements have taken place in government, enhancing planning, use of evaluations and boosting evaluation culture, including changing behaviour of actors.

Indeed, regarding exclusion of the judicial and legislative branches to perform any evaluation (e.g. object under evaluation), the Mexican Congress has taken an important step forward on this. At the beginning of year 2017, it was enacted an agreement to implement the Performance Evaluation System within the Chamber of Deputies,<sup>272</sup> which represent a relevant effort by Congress to be accountable of their performance. This effort is coherent with initiatives launched by international organisations and professional associations of evaluators such as Eval Partners, which support the parliamentarians’ movement for evaluation<sup>273</sup> through the key principles

---

<sup>272</sup> Official Journal, 06/03/2017.

<sup>273</sup> <https://www.evalpartners.org/globalparliamentariansforum/background>



of Sustainable Development Goals (SDGs) “No one left behind” and the Global Evaluation Agenda 2016-2020. See also “Ten reasons for parliamentarians to engage in evaluation policy process”<sup>274</sup>

Future studies should include a number of large-scale countries’ evaluations and wider policy sectors e.g. education, energy, social, to discuss the extent to which these sectors follow the same pattern of effectiveness as the health sector analysed here. Such empirical findings will contribute to confirm whether effectiveness of policy evaluation is extensive to all sectors for further generalizations. Most importantly, this thesis highlights the relevance of looking at other factors (see chapter ‘Broadening understandings of policy evaluation and effectiveness’) for understanding how and whether these factors could change the results, particularly, the need to bring questions of leadership, power and control into the analysis to ensure a comprehensive perspective.

The understanding of major similarities and differences of independent variables across countries has potential for prediction once generalizations are validated. Similarly to deductive reasoning and what Hague & Harrop (2004: 70) pointed out ‘The ability to predict is not only a sign of systematic knowledge but it also gives some base for drawing lessons across countries’. These government experiences will help to draw inferences about causes of policy evaluation stagnation or possible explanatory reasons about why some policies are more effective than others.

### **Final statements**

This thesis has contributed to the theoretical field of evaluation in discussing whether policy evaluation is effective across countries, raising the question *what is evaluation if nobody wants to take it for further decision-making in the policy process?* A central strand might be to construct better programme theories (Weiss, 2007: 78), as the backbone for public policy process and interventions. For this, the research questions have been answered with a clear statement for each question. It is therefore important to reinforce some statements for the acknowledgement of the significance of the study. For instance, evaluation has become a commonplace across governments but still faces resistance by decision-makers due to several factors already discussed here. These factors jeopardize the effectiveness of policy evaluation. Therefore, the wider significance of this research is on the extent it can enter into political arenas as ammunition for debate.

With the aim to impact political stances and influencing the way policy evaluation is conducted by actors and to maximise its use for decision-making, legitimise and give direction to political decisions, and to acknowledge the topic, as Weiss (2004: 161) affirms “Utility is what evaluation is all about”, the government reforms should move forward to develop a wider policy evaluation, using performance information and ensuring the effectiveness (modifications, funding, cancellation or learning process) of evaluation in the policy process.<sup>275</sup>

Moreover, to enhance the use of evidence, Walter et al. (2005: 352-4) suggest having collaborative and interactive communication between researchers and policy-makers to ‘test out’ findings within local contexts, to offer technical, financial,

---

<sup>274</sup> <http://www.pfde.net/images/pdf/101.pdf>

<sup>275</sup> In the beginning, Medina (2007: 26-7) points out that ‘external evaluation was performed for accountability and transparency purposes, not for linking results with the budgetary decision making, planning or learning process’. However, many political issues influenced this transformation.

organisational and emotional support as well as to include incentives for promoting research use and enabling debate about implications. However, these scholars' suggestions are also those that are most resources intensive such as partnership development – to strengthen the IGR –, well-funded facilitative programmes and multifaceted interventions (Walter et al., 2005: 355).

Moreover and based on the assumption that evaluation emerges more effectively in a context of financial crisis, therefore, it is fundamental to know about those programmes performing well for better and efficient decision-making concerning the allocation of finances. In fact, governments are facing spending constraints and shrinking financial incomes, for instance, México depends upon oil income, but this is currently a downturn due to the need of financial resources for exploration of new oil fields and drilling in deepwater fields, amongst others. The lack of funds constitutes a strong device for strengthening effectiveness in the public agenda. In the experiences studied here, what makes the difference between governments is the institutional capabilities such as leadership, technical and practical skills, knowledge, experience, institutional efficiency and the modernization of the public management, amongst others.

In a wider perspective, these governments also need to understand that the strategic planning process is the best way for economic growth, transparency and accountability. Conversely, a poor process would have a negative impact on the budget. In this sense, evaluation should impact values, judgements and priorities of policy-makers more accurately, ensuring that evaluative facts and better knowledge influence the trade-offs in decisions. Evaluation will be effective when it is based on a consensus among political and non-political actors and when the government machinery is centred upon outcomes facing resistances from those supporting the *status quo*.

### **List of acronyms**

APF: Federal Public Administration (Administración Pública Federal).  
ASF: Federal Superior Audit (Auditoría Superior de la Federación).  
CONACYT: National Council of Science and Technology (Consejo Nacional de Ciencia y Tecnología).  
CONEVAL: National Evaluation Council of the Social Development Policy (Consejo Nacional de Evaluación de la Política de Desarrollo Social).  
COSEFF: Monitoring and Control Financial System of the Federal Public Expenditure.  
EIU: The Economist Intelligence Unit.  
FY: Fiscal Year  
IGR: Intra Governmental and Interagency Relationships.  
LOAPF: Federal Government Organic Law (Ley Orgánica de la Administración Pública Federal).  
OIC: Internal Audit and Control Units (Órganos Internos de Control).  
MoF: Ministry of Finance.  
PbR/SED: Results-based Budgeting and Performance Evaluation System.  
PEF: Federal Expenditure Budget (Presupuesto de Egresos de la Federación).  
PND: National Development Plan (Plan Nacional de Desarrollo).  
RbB: Results-based Budgeting.  
SECODAM: Ministry of the Controllorship and Administrative Development (Secretaría de Contraloría y Desarrollo Administrativo).  
SECOGEF: Ministry of the General Controller of the Federation (Secretaría de la Contraloría General de la Federación).  
SFP: Ministry of Public Administration (Secretaría de la Función Pública).  
SICEG: System for Government Monitoring and Evaluation (Sistema de Control y Evaluación Gubernamental).  
SIG: Management Indicator System (Sistema de Indicadores de Gestión).  
SIMEP: Presidential Goals' System (Sistema de Metas Presidenciales).  
SHCP: Ministry of Finance and Public Credit (Secretaría de Hacienda y Crédito Público).  
SNPD: National Democratic Planning System (Sistema Nacional de Planeación Democrática).  
SPP: Ministry of Planning and Budget (Secretaría de Programación y Presupuesto).

## **List of interviews**

### **MEXICO**

- IM1 Senior Official, ASF, February 20, 2014
- IM2 Middle-Level Official, SHCP, February 24, 2014
- IM3 Senior Official, SSA, February 24, 2014
- IM4 Middle-Level Official, SHCP, February 25, 2014
- IM5 Senior Official, SSA, February 26, 2014
- IM6 Middle-Level Official, SSA, February 26, 2014
- IM7 Middle-Level Official, SNDIF, February 27, 2014
- IM8 Former Deputy, March 9, 2014
- IM9 Deputy, Budget Commission, March 13, 2014
- IM10 Middle-Level Official, CEFPE, March 13, 2014
- IM11 Deputy, March 13, 2014
- IM12 Middle-Level Official, STPS, March 18, 2014
- IM13 Deputy, Oversight Commission, March 25, 2014
- IM14 Deputy, Budget Commission, March 25, 2014
- IM15 Deputy, March 26, 2014
- IM16 Deputy advisor, March 27, 2014
- IM17 Middle-Level Official, SFP, March 14, 2014
- IM18 Middle-Level Official, SFP, March 15, 2014
- IM19 Middle-Level Official, SFP, March 27, 2014

### **CHILE**

- IC1 Deputy, Health Commission, June 13, 2014
- IC2 Deputy, Budget Commission, August 7, 2014
- IC3 Senior Official, Ministry of Health, August 19, 2014
- IC4 Academic expert, University of Chile, September 9, 2014

## References

- Abma, T. (2006) *The social relations of evaluations* in "The SAGE handbook of evaluation". London: SAGE Publications Ltd., eds. Ian F. Shaw, Jennifer C. Greene and Melvin M. Mark, ch. 8, pp. 184-199.
- Aberbach, J. and Rockman, B. (2002) *Conducting and coding elite interviews*. Washington, D.C.: American Political Science Association, PS: Political Science and Politics, vol. 35, no. 4, pp. 673-676.
- Agranoff, R. and Radin, B. (2015) *Deil Wright's overlapping model of Intergovernmental Relations: The basis for contemporary Intergovernmental Relationships*. Publius: The Journal of Federalism, 45, issue 1, pp. 139-159.
- Aguilar, L. (1996) *Reformas y retos de la administración pública mexicana (1988-1994)*. México: Foro Internacional, Jan.-June 1996.
- Aguilar, L. (1996a) *La silenciosa, heterodoxa reforma de la administración pública*. Mexico: Instituto Nacional de Administración Pública, Revista de Administración Pública (Ensayos sobre la administración pública), núm. 91. Available at: <http://www.juridicas.unam.mx/publica/rev/indice.htm?r=rap&n=91>
- Aguilar, L. (2013) *El gobierno del gobierno*. Mexico: Instituto Nacional de Administración Pública 'INAP'.
- Alkin, M. (1985) *A guide for evaluation decision makers*. Beverly Hills: Sage publications Ltd.
- Alkin, M. (ed. 2004) *Evaluation roots: tracing theorists' views and influences*. London: Sage Publications Ltd.
- Alkin, M. and Christie, C. (2004) *An evaluation theory tree* in "Evaluation roots: tracing theorists' views and influences". London: Sage Publications Ltd.
- Anderson, W. (1960) *Intergovernmental Relations in review*. Minneapolis: University of Minnesota Press.
- Andrews, M. (2008) *Are One-Best-Way Models of Effective Government Suitable for Developing Countries?* Massachusetts: Harvard University, John F. Kennedy School of Government, Faculty Research Working Papers Series, March 2008.
- Arellano, D. (2000) *Aspectos organizacionales de los sistemas de servicio civil: dimensiones y puntos críticos para la experiencia mexicana [Organisational factors of the career civil services: critical factors and dimensions of the Mexican experience]*. Mexico: Universidad Autónoma de México, Revista UAM, Iztapalapa, no. 48, pp. 135-150.
- Arellano, D. (2002) *Nueva Gestión Pública: ¿el meteorito que mató al dinosaurio? Lecciones para la reforma administrativa en países como México*. Venezuela: Revista del CLAD Reforma y democracia, no. 23, June 2002.
- Arellano, D. and Guerrero, P. (2003) ed. Schneider, B. and Heredia, B. *Stalled administrative reforms of the Mexican State* in "Reinventing Leviathan. The Politics of Administrative Reform in Developing Countries". Miami: North South Center Press, University of Miami.
- Arellano, D. and Ramírez, J. (2000). *Presupuesto dirigido a resultados: los dilemas de la evaluación por desempeño en contextos latinoamericanos. Lecciones desde México [Results based Budgeting: dilemmas of performance evaluation in Latin American contexts. Lessons from México]*. Santo Domingo: CLAD, V Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública. Available at: <http://www.orfis.gob.mx/InformacionPublica/Asofis/presupuestoresultados.pdf>
- Arenas, A. and Berner, H. (2010) *Presupuesto por Resultados y la consolidación del sistema de evaluación y control del gobierno central*. Santiago de Chile: Ministerio de Hacienda, Dirección de Presupuestos DIPRES. Available at:

- [http://www.dipres.gob.cl/594/articles-60578\\_doc\\_pdf.pdf](http://www.dipres.gob.cl/594/articles-60578_doc_pdf.pdf)
- Armijo, M. (2002) *Reforma y modernización del Estado. Experiencias y desafíos*, Armijo, M. and Tomassini, L. (eds.). Santiago de Chile: LOM ediciones and Instituto de Asuntos Públicos, Universidad de Chile.
- Askim, J. (2007) *How do politicians use performance information? An analysis of the Norwegian local government experience*. London: SAGE publications, pp. 453-72.
- Auditoría Superior de la Federación (2007) *Results Report of the Account of Public Treasury 2006*. Mexico: Auditoría Superior de la Federación, sector Finance and Public Credit, issue III, vol. I. Available at: <http://www.asf.gob.mx/Section/58 Informes de auditoria>
- Auditoría Superior de la Federación (2012) *Evaluation of the Performance Evaluation System [Sistema de Evaluación del Desempeño]*. Mexico: Superior Audit Office of the Federation, evaluation no. 117. Available at: [http://www.asf.gob.mx/Trans/Informes/IR2012i/Documentos/Auditorias/2012\\_0117\\_a.pdf](http://www.asf.gob.mx/Trans/Informes/IR2012i/Documentos/Auditorias/2012_0117_a.pdf)
- Auditoría Superior de la Federación (2013) *Informe general de la Cuenta Pública 2013 [Account of Public Treasury 2013]*. Mexico: Superior Audit Office of the Federation. Available at: <http://www.asf.gob.mx/Trans/Informes/IR2013i/Paginas/Master Ejecutivo.htm>
- Bachrach, P. and Baratz, M. (1962) *Two faces of power*. The American Political Science Review, vol. 56, issue 4, pp. 947-52.
- Bachrach, P. and Baratz, M. (1963) *Decisions and nondecisions: an analytical framework*. The American Political Science Review, vol. 57, issue 3, pp. 632-42.
- Bailey, J. (1980) *La Secretaría de Programación y Presupuesto*. Mexico: Revista de Administración Pública. Instituto de Investigaciones Jurídicas, no. 44. Oct.-Dec. Available at: <http://www.juridicas.unam.mx/publica/librev/rev/rap/cont/44/pr/pr2.pdf>
- Bailey, J. (1984) *Public Budgeting in Mexico, 1970-1982*. Public Budgeting and Finance, volume 4, issue 1, pp. 76-90.
- Bamberger, M. (1991) *The politics of evaluation in developing countries*. Washington, D.C.: The World Bank, The Economic Development Institute, 'Evaluation and Program Planning', vol. 14, pp. 325-330.
- Bangura, Y. and Larbi, G. (eds.) (2006) *Public Sector Reform in Developing Countries. Capacity Challenges to Improve Services*. Hampshire: Palgrave Macmillan and United Nations Research Institute for Social Development (UNRISD).
- Bardon, Á., Carrasco, C and Vial, Á. (1985) *Una década de cambios económicos. La experiencia Chilena 1973-1983*. Santiago de Chile: editorial Andrés Bello.
- Barzelay, M. (2001) *The New Public Management: Improving Research and Policy Dialogue*. Berkeley: University of California Press.
- Barzelay, M. (2003) *The New Public Management: Invitation to Cosmopolitan Dialogue "La Nueva Gerencia Pública: Invitación a un diálogo cosmopolita"*. Mexico: Centro de Investigación y Docencia Económicas (CIDE), Revista Gestión y Política Pública, II semestre, vol. 12, no. 2, pp. 241-51.
- Beamer, G. (2002) *Elite interviews and state politics research*. London: SAGE Publications Ltd., State politics and policy quarterly, spring 2002.
- Becerril-Montekio, V., Reyes, J. and Manuel, A. (2011) *Sistema de salud de Chile [The health system of Chile]*. Mexico: Instituto Nacional de Salud Pública, vol. 53, issue 2. Available at: [www.scielo.org.mx/pdf/spm/v53s2/09.pdf](http://www.scielo.org.mx/pdf/spm/v53s2/09.pdf)
- Bergoeing, R., Kehoe, P., Kehoe, T., Soto, R. (2002) *A decade lost and found: Mexico and Chile in the 1980s*. Review of Economic Dynamics, vol. 5, issue 1, pp. 166-205
- Bernard, H. R. (2013) *Social research methods. Qualitative and quantitative approaches*. London: SAGE Publications Ltd., 2<sup>nd</sup>. ed.



- Bernstein, B. (2000) *Pedagogy, symbolic control and identity: theory, research, critique*. Maryland: Rowman & Littlefield publishers, inc., 2<sup>nd</sup>. ed., ch. 1, pp. 3-24.
- Birkland, T. (2011) *An introduction to the policy process: theories, concepts, and models of public policy making*. New York: M. E. Sharpe, Inc., 3<sup>rd</sup>. ed.
- Bovens, M., 't Hart, P. and Peters, B. G. (2001) *Success and failure in public governance. A comparative analysis*. Cheltenham: Edward Elgar Publishing Ltd., New horizons in public policy, series editor: Wayne Parsons.
- Bovens, M., 't Hart, P. and Kuipers, S. (2006) *The politics of policy evaluation* in "The Oxford handbook of public policy". Oxford: Oxford University Press, ed. Michael Moran, Martin Rein and Robert E. Goodin, chapter 15.
- Bromage, A. (1943) *Federal-State-Local Relations*. The American Political Science Review, vol. 37, no. 1, Feb. 1943, pp. 35-48.
- Bundi, P. (2016) *What do we know about the demand for evaluation? Insights from the parliamentary arena*. Thousand Oaks: Sage publications, American Journal of Evaluation, pp. 1-20.
- Burnham, P., Gilland, K., Grant, W. and Layton-Henry, Z. (2004) *Research Methods in Politics*. New York: Palgrave Macmillan.
- Cabrero, E. (2000) *Practices and Traditions in Public Policies Making. Limits of Policy Sciences in political and cultural different contexts (Usos y Costumbres en la Hechura de las Políticas Públicas en México. Límites de las Policy Sciences en contextos cultural y políticamente diferentes)*. Mexico: Centro de Investigación y Docencia Económicas, Revista Gestión y Política Pública, II semestre, year/vol. IX, no. 002, pp. 180-229.
- Cabrero, E. (2003) *Gerencia pública municipal. Conceptos básicos y estudios de caso*. Mexico: Miguel Ángel Porrúa-CIDE, pp. 19-90.
- Cabrero E., Arellano D. and Amaya M. (2003) *Cambio en organizaciones gubernamentales: innovación y complejidad [Public agencies changing: complexity and innovation]* in Cabrero E., Arellano D. and Del Castillo A. (Coord) *"Reformando al gobierno: Una visión organizacional del cambio gubernamental [Improving government: an organisational visión of governmental changing]"*. Mexico: Miguel Ángel Porrúa-CIDE, 2003, pp. 137-168.
- Cámara de Diputados (2017). *Acuerdo por el que se autoriza el Sistema de Evaluación del Desempeño de la Cámara de Diputados y sus Lineamientos [Agreement to approve the Performance Evaluation System of the Chamber of Deputies]*. Mexico: Official Journal, 06/03/2017.
- Cámara de Diputados (2003) *Ley del Servicio Público del Ejecutivo [Law of Senior Public Executive no. 19.882]*. Santiago de Chile: National Congress Library [03/03/2014 last reform].
- Cámara de Diputados (2004) *Ley General de Desarrollo Social [The General Law of Social Development]*. Mexico: Official Journal, 20/01/2004.
- Cámara de Diputados (2005) *Constitución Política de la República de Chile*. Santiago de Chile: Chamber of Deputies [03/05/2014 last reform].
- Cámara de Diputados (2009) *Ley Orgánica de la Administración Pública Federal [Federal Government Organic Law]*. Mexico: Official Journal, 17/06/2009.
- Cámara de Diputados (2011) *Ley de Planeación [Planning Law]*. Mexico: Official Journal, 20/06/2011.
- Cámara de Diputados (2011) *Memoria anual 2010-11*. Valparaíso: Chamber of Deputies. Available at: [http://www.camara.cl/camara/media/especiales/cuenta\\_2011/memoria\\_10-11.pdf](http://www.camara.cl/camara/media/especiales/cuenta_2011/memoria_10-11.pdf)
- Cámara de Diputados (2012) *Ley Federal de Presupuesto y Responsabilidad Hacendaria [Federal Law of Budget and Fiscal Responsibility]*. Mexico: Official Journal, 09/04/2012.

- Cámara de Diputados (2012) *Mexican Political Constitution [Constitución Política de los Estados Unidos Mexicanos]*.
- Cámara de Diputados (n.d.) *Plan Nacional de Desarrollo [National Development Plan]*. Mexico.
- Cámara de Diputados (2014) Federal Budget (PEF 2014). Mexico: Chamber of Deputies, 2014.
- Campbell, D. and Stanley, J. (1966) *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Cardozo, M. (2006) *The evaluation of programmes and public programmes. The case of social development programmes in México [La evaluación de políticas y programas públicos. El caso de los programas de desarrollo social en México]*. Mexico: Miguel Ángel Porrúa and Mexican Congress. Available at: [biblioteca.diputados.gob.mx/janium/bv/ce/.../eval\\_pol\\_prog\\_pub.pdf](http://biblioteca.diputados.gob.mx/janium/bv/ce/.../eval_pol_prog_pub.pdf)
- Carpenter, D. (2010) *Reputation and power. Organizational image and pharmaceutical regulation at the FDA*. Princeton, NJ: Princeton University Press.
- Carrillo, A. and García, S. (1981) *Las empresas públicas en México*. Mexico: Editorial Miguel Ángel Porrúa.
- Caughey, D., Chatfield, S. and Cohon, A. (2009) *Defining, mapping, and measuring bureaucratic autonomy*. Berkeley: University of California.
- Cejudo, G. (2008) *Explaining change in the Mexican public sector: the limits of New Public Management*. London: Sage Publications.
- Centro de Estudios de las Finanzas Públicas (2008) *El Sistema de Evaluación del Desempeño y Programa Anual de Evaluación [Performance Evaluation System and Annual Programme Evaluation]*. Mexico: Chamber of Deputies. Mexican Congress. Available at: <http://www.cefp.gob.mx/intr/edocumentos/pdf/cefp/2008/cefp1152008.pdf>
- Centro de Estudios de las Finanzas Públicas (2008b). *Evolución del gasto en salud [Evolution of health expenditure]*. Mexico: Chamber of Deputies. Mexican Congress, 106/2008. Available at: <http://www.cefp.gob.mx/intr/edocumentos/pdf/cefp/2008/cefp1062008.pdf>
- Centro de Estudios de las Finanzas Públicas (2009). *Proyecto de Presupuesto de Egresos de la Federación 2010, recursos asignados a la función salud y a la Secretaría de Salud [Draft of Federal Budget 2010, public resources to health sector and Ministry of Health]*. Mexico: Chamber of Deputies. Mexican Congress, 104/2009. Available at: <http://www.cefp.gob.mx/intr/edocumentos/pdf/cefp/2009/cefp1042009.pdf>
- Centro de Estudios de las Finanzas Públicas (2013). *Comparativo del decreto de Presupuesto de Egresos de la Federación 2012 - 2013 [Comparative of decree of Federal Budget 2012-2013]*. Mexico: Chamber of Deputies. Mexican Congress, 001/2013. Available at: <http://www.cefp.gob.mx/publicaciones/documento/2013/enero/cefp0012013.pdf>
- Centro de Estudios de las Finanzas Públicas (2013a). *Infografías del Presupuesto de Egresos de la Federación 2013 [Infographics of Federal Budget 2013]*. Mexico: Chamber of Deputies. Mexican Congress, 12/2013. Available at: <http://www.cefp.gob.mx/publicaciones/infografias/2013/dep/infdep0122013.pdf>
- Centro de Estudios de las Finanzas Públicas (2013b). *Iniciativas de decreto por el que se expiden la Ley de Pensión Universal y la ley del Seguro de Desempleo [Initiative law of retirement pension and law of unemployment benefit]*. Mexico: Chamber of Deputies. Mexican Congress, 9/sep/2013. Available at: <http://www.cefp.gob.mx/publicaciones/nota/2013/septiembre/notacefp0692013.pdf>



- Centro de Estudios de las Finanzas Públicas (2014). *Comparativo del proyecto y decreto de Presupuesto de Egresos de la Federación 2014 [Comparative of draft and decree of Federal Budget 2014]*. Mexico: Chamber of Deputies. Mexican Congress, 001/2014. Available at: <http://www.cefp.gob.mx/publicaciones/documento/2014/enero/cefp0012014.pdf>
- Chapoy, D. (2003) *Planeación, programación y presupuestación [Planning, Programming and Budgeting]*. Mexico: Universidad Nacional Autónoma de México, Instituto de Investigaciones Jurídicas, Serie Doctrina Jurídica, no. 145. Available at: <http://biblio.juridicas.unam.mx/libros/libro.htm?l=1009>
- Chase, S. (2005) *Narrative inquiry: multiple lenses, approaches, voices* (Norman K. Denzin, N. and Lincon, Y. eds.) in *"The SAGE handbook of qualitative research"*. London: Sage.
- Chen, H. (2004) *The roots of theory-driven evaluation. Current views and origins* in *"Evaluation roots: tracing theorists' views and influences"*. London: Sage Publications Ltd. ed. Alkin, M., chapter 8.
- Chen, H. and Rossi, P. (1981) *The multi-goal theory-driven approach to evaluation: a model linking basic and applied social science* in H. Freeman and M. Solomon (eds.) *"Evaluation Studies Review Annual"*, vol. 6, Beverly Hills: SAGE, pp. 38-54.
- Christensen J. (1999) *Bureaucratic autonomy as a political asset*. Denmark: Paper presented on the ECPR Joint Workshops, Mannheim, 26-31, March, 1999.
- CLAD (2001) *La evaluación como herramienta para una gestión pública orientada a resultados*. Caracas: Centro Latinoamericano de Administración para el Desarrollo (CLAD). Available at: <http://siare.clad.org/siare/innotend/evaluacion/evaluacio1.html>
- Clarke, A. and Dawson, R. (1999) *Evaluation Research. An introduction to principles, methods and practices*. London: SAGE Publications Ltd.
- Cochran, Ch. and Malone, E. (2010) *Methods and Models for Policy Analysis in "Public Policy perspectives and choices"*. London: Lynne Rienner Publishers, Inc., 4<sup>th</sup> ed., pp. 29-58.
- Cochran, Ch. and Malone, E. (2010) *Public Policy perspectives and choices*. London: Lynne Rienner Publishers, Inc., 4<sup>th</sup> ed.
- Collier, D. And Mahoney, J. (1996) *Insights and Pitfalls: Selection Bias in Qualitative Research*. Cambridge: Cambridge University Press, World Politics, vol. 49, no. 1, oct. 1996, pp. 56-91.
- Consejo Nacional de Evaluación de la Política de Desarrollo Social (2008) *Informe de Evaluación de la Política de Desarrollo Social en México 2008*. Mexico: CONEVAL.
- Consejo Nacional de Evaluación de la Política de Desarrollo Social (2011) *Informe de Evaluación de la Política de Desarrollo Social en México. En materia de acceso a los servicios de salud 2011*. Mexico: CONEVAL.
- Contandriopoulos, D., Lemire, M. Denis, J. and Tremblay, E. (2010) *Knowledge exchange process in organizations and policy arenas: a narrative systematic review of the literature*. Milbank Quarterly, vol. 88, no. 4, pp. 444-83.
- Contandriopoulos, D. and Brousselle, A. (2012) *Evaluation models and evaluation use*. London: SAGE publications, Evaluation, special issue: European Evaluation Society Conference, vol. 18, no. 1, January 2012, pp. 61-77.
- Coote, A., Allen, J., and Woodhead, D. (2004) *Finding Out What Works. Building knowledge about complex, community-based initiatives (Summary)*. London: Kings Fund Publications.
- Cousins, B. and Leithwood, K. (1986) *Current Empirical Research on Evaluation Utilization*. London: Sage Journals, Review of Educational Research vol. 56, no. 3, pp. 331-364.

- Christensen J. (1999) *Bureaucratic autonomy as a political asset*. Denmark: Paper presented on the ECPR Joint Workshops, Mannheim, 26–31 March, 1999.
- Christensen, T. (2006) *Smart policy?* in *The Oxford handbook of public policy*. Oxford: Oxford University Press, ed. Michael Moran, Martin Rein and Robert E. Goodin, chap. 21.
- Curristine, T. (2005) *Government Performance: Lessons and Challenges*. Paris: Organisation for Economic Co-operation and Development, OECD Journal on Budgeting, vol. 5/1. Available at: [http://www.oecd-ilibrary.org/governance/government-performance\\_budget-v5-art6-en](http://www.oecd-ilibrary.org/governance/government-performance_budget-v5-art6-en)
- Curristine, T. and Bas, M. (2007) *Budgeting in Latin America: Results of the 2006 OECD Survey*. Paris: Organisation for Economic Co-operation and Development, OECD Journal on Budgeting, vol. 7/1. Available at: <http://dx.doi.org/10.1787/budget-v7-art4-en>
- Curristine, T., Lonti, Z. and Joumard, I. (2007) *Improving Public Sector Efficiency: Challenges and Opportunities*. Paris: Organisation for Economic Co-operation and Development, OECD Journal on Budgeting, vol. 7/1. Available at: <http://www.oecd.org/gov/budgeting/43412680.pdf>
- Dahler-Larsen, P. (2000) *Surviving the Routinization of Evaluation: The Administrative Use of Evaluations in Danish Municipalities*. London: SAGE publications, Administration & Society, pp. 70–92.
- Dahler-Larsen, P. (2005) *Evaluation and Public Management* in “*The Oxford handbook of public management*”. Oxford: Oxford University Press, ed. Ewan Ferlie, Laurence E. Lynn Jr. and Christopher Pollitt.
- Denzin, N. and Lincoln, Y. (eds.) (2005) *Handbook of Qualitative Research*. London: SAGE Publications Ltd.
- Dexter, L. (2006). *Elite and specialized interviewing*. Colchester: The European Consortium for Political Research, ECPR Press, theUniversity of Essex.
- Diario Oficial (1930) *Ley sobre Planeación General de la República*. Mexico: Diario Oficial de la Federación 12/07/1930.
- Diario Oficial (1989) *Acuerdo por el que se establecen las bases para la ejecución, coordinación, y evaluación del Programa General de Simplificación de la Administración Pública Federal*. Mexico: Diario Oficial de la Federación 09/02/1989.
- Diario Oficial (1991) *Reglamento interior de la Secretaría de la Contraloría General de la Federación*. Mexico: Diario Oficial de la Federación (26/02/1991, 16/01/1989).
- Diario Oficial (1994) *Decree regarding to revision of the Account of Public Treasury of 1994*. Mexico: the Treasury Accounting Office, Official Journal 11/12/1995.
- Diario Oficial (2001) *Reglamento interior de la Secretaría de la Contraloría y Desarrollo Administrativo*. Mexico: Diario Oficial de la Federación (16/07/2001, 12/04/1995).
- Diario Oficial (2004) *Ley General de Desarrollo Social [General Law of Social Development]*. Mexico: Diario Oficial de la Federación [07/11/2013 last reform].
- Diario Oficial (2006) *Ley General de Presupuesto y Responsabilidad Hacendaria*. Mexico: Chamber of Deputies [11/08/2014 last reform].
- Diario Oficial (2006) *Manual de Organización General de la Secretaría de la Función Pública*. Mexico: Diario Oficial de la Federación 10/05/2006.
- Diario Oficial (2006) *Proyecto de Presupuesto de Egresos de la Federación para el ejercicio fiscal 2006 (PPEF) [The statement of purpose of the draft budget 2006]*. Mexico: Diario Oficial de la Federación.
- Diario Oficial (2008) *Presupuesto de Egresos de la Federación para el ejercicio fiscal 2008 (PEF) [Federal budget 2008]*. Mexico: Diario Oficial de la Federación.

- Diario Oficial (2009) *Ley Orgánica de la Administración Pública Federal*. Mexico: Diario Oficial de la Federación (17/06/2009, 28/12/1994, 21/02/1992, 25/05/1992, 29/12/1976).
- Diario Oficial (2011) *Ley de Planeación*. Mexico: Diario Oficial de la Federación 20/06/2011.
- Diario Oficial (2011) *Reglamento interior de la Secretaría de la Función Pública*. Diario Oficial de la Federación (03/08/2011, 24/08/2009, 15/04/2009, 21/06/2005, 12/12/2003).
- Diario Oficial (2012) *Diario de los Debates* [Daily debates]. Mexico: Chamber of Deputies, LXI Legislatura, versiones pdf, segundo periodo ordinario del tercer año de ejercicio, Feb.-April 2012.
- Diario Oficial (2012) *Diario de los Debates* [Daily debates]. Mexico: Chamber of Deputies, LXII Legislatura, versiones pdf, primer periodo ordinario del primer año de ejercicio, Aug.-Dec. 2012.
- Diario Oficial (2014) *Presupuesto de Egresos de la Federación para el ejercicio fiscal 2014 (PEF) [Federal budget 2014]*. Mexico: Diario Oficial de la Federación.
- Diario Oficial (2015) *Constitución Política de los Estados Unidos Mexicanos*. Mexico: Diario Oficial de la Federación [10/07/2015 last reform].
- Diener, M. (2007) *Results-based budgeting: The presidential goal's system in México in "Towards institutionalization of monitoring and evaluation systems in Latin America and the Caribbean"* eds. García, R., Laguzzi, H. and Guerrero, W. [Presupuesto para resultados: El sistema de metas presidenciales en México en "Hacia la institucionalización de los sistemas de monitoreo y evaluación en América Latina y el Caribe"]. Washington, D.C.: Inter-American Development Bank, II Conferencia Regional de M&E.
- DIPRES (n/d) *Guía para la construcción de indicadores de desempeño en los servicios públicos*. Santiago de Chile: Ministerio de Hacienda. Available at: [http://www.subdere.cl/sites/default/files/documentos/articles-84009\\_archivo\\_fuente.pdf](http://www.subdere.cl/sites/default/files/documentos/articles-84009_archivo_fuente.pdf)
- DIPRES (2001) *Evaluación de programas e indicadores de desempeño. Transparencia y mejoramiento de los procedimientos para la elaboración y discusión presupuestaria*. Santiago de Chile: Ministerio de Hacienda. Available at: [http://www.dipres.gob.cl/594/articles-89670\\_doc\\_pdf.pdf](http://www.dipres.gob.cl/594/articles-89670_doc_pdf.pdf)
- DIPRES (2001a) *Programas de Mejoramiento de la Gestión 'PMG'*. Santiago de Chile: Ministerio de Hacienda. Available at: [http://www.dipres.gob.cl/594/articles-90104\\_doc\\_pdf.pdf](http://www.dipres.gob.cl/594/articles-90104_doc_pdf.pdf)
- DIPRES (2005) *El proceso presupuestario en Chile, aspectos generales: marco legal, actores institucionales, principales aspectos de modernización*. Santiago de Chile: Ministerio de Hacienda. Available at: <http://bibliotecadigital.dipres.gob.cl/bitstream/handle/123456789/8323/El%20Proceso%20Presupuestario.pdf?sequence=1>
- DIPRES (2008) *System of Evaluation and Management Control Results-based Budgeting. The Chilean experience*. Santiago de Chile: DIPRES, presentation of Heidi Berner, Head of Management Control Division, Budget Office, Ministry of Finance.
- DIPRES (2012) *Requisitos técnicos y medios de verificación Programa de Mejoramiento de la Gestión (PMG). Año 2013*. Santiago de Chile: Ministerio de Hacienda. Available at: [http://www.dipres.gob.cl/594/articles-105244\\_doc\\_pdf.pdf](http://www.dipres.gob.cl/594/articles-105244_doc_pdf.pdf)
- DIPRES (2014) *Programas de Mejoramiento de la Gestión PMG: Evolución 2011-2014 y desafíos futuros*. Santiago de Chile: Ministerio de Hacienda. Available at: [http://www.dipres.gob.cl/594/articles-114720\\_doc\\_pdf.pdf](http://www.dipres.gob.cl/594/articles-114720_doc_pdf.pdf)

- Dumas, V., Lafuente, M. and Parrado, S. (2013) *El fortalecimiento del centro de gobierno para resultados en Chile. La experiencia del Ministerio de la Presidencia y su Unidad Presidencial de Gestión del Cumplimiento (2010-13)*. Washington, D.C.: Inter-American Development Bank, technical note no. 563.
- Dunn, W. N. (2012) *Public policy analysis*. Boston: Pearson Education, Inc., 5th ed.
- Dussauge, M. (2009) *Comparative literature about administrative reforms. Developments, constraints and opportunities* [La literatura comparada sobre reformas administrativas. Desarrollos, limitaciones y posibilidades]. México: CIDE, Revista Gestión y Política Pública, vol. XVIII, no. 2, II semestre 2009.
- Dussauge, M. (2012a) *Chile's M&E system* in G. Lopez-Acevedo, P. Krause and K. Mackay (eds.) "Building Better Policies. The nuts and bolts of monitoring and evaluation systems". Washington D.C.: International Bank for Reconstruction and Development, The World Bank, pp. 183-193
- Dussauge, M. (2012b) *Transferencia de políticas y modernización administrativa: historia (y moralejas) de cuando el Programa de Mejoramiento de la Gestión Chileno "viajó" a México*. Mexico: Revista de Gestión Pública, vol. I, no. 2, Jul.-Dec. 2012, pp. 237-272.
- Dussauge, M. (2013) *Cross-national policy learning and administrative reforms. The making of 'Management for Results' policies in Chile and Mexico (1990-2010)*. PhD thesis. London: London School of Economics and Political Science.
- Duverger, M. (1992) *Instituciones políticas y derecho constitucional*. Barcelona: editorial Ariel.
- Dye, T. (2011) *Understanding Public Policy*. Boston: Pearson Education, Inc., 13<sup>th</sup> ed.
- Einsele, S. (1962) *Los sistemas políticos de los imperios*. Madrid: Revista de Occidente.
- Eisenhardt, K. (1989) *Agency theory: an assessment and review*. The Academy of Management Review, vol. 14, no. 1 (January 1989), pp. 57-74.
- Elliott, J. (2005) *Listening to people's stories; the use of narrative in qualitative interviews, Using Narrative in Social Research*. London: SAGE Publications Ltd., chapter 2.
- Enríquez, I. (2010) *Evolution of Thought on Development in the International Monetary Fund and the World Bank. A Comparative Analysis in Latin America in "Trayectorias"*. Mexico: Revista de Ciencias Sociales de la Universidad Autónoma de Nuevo León, 12, no. 31, Jul.-Dec. 2010.
- Essama-Nssah, B. (2013) *What Questions Do Evaluations Answer?* Washington, D.C.: the World Bank, PREMnotes no. 25, special series on "The Nuts & Bolts of M&E Systems".
- Fazio, H. (1996) *El programa abandonado: balance económico social del gobierno de Aylwin*. Santiago de Chile: LOM ediciones.
- Figueroa, V., Olavarría, M., and Navarrete, B. (2011) *Política de modernización de la gestión pública en Chile 1990-2006: evidencias a partir de un modelo de análisis*. Mexico: Universidad Autónoma del Estado de México, Convergencia. Revista de ciencias sociales, vol. 18, no. 57, sept.-dec. 2011, pp. 61-99.
- Fitzpatrick, J., Sanders, J. and Worthen, B. (2004) *Program evaluation: alternative, approaches and practical guidelines*. Boston: Pearson Education, Inc., 3<sup>rd</sup> edition.
- Fleischer, D. and Christie, Ch. (2009) *Evaluation Use. Results from a survey of U.S. American Evaluation Association Members*. Thousand Oaks: SAGE journals. American Journal of Evaluation, vol. 30, no. 2, june 2009, pp. 158-75.
- Foucault, M. (1980) *Power/knowledge: selected interviews and other writings, 1972-1977* (Gordon, C. ed.). New York, N.Y.: Pantheon Books.
- Franco, A. (2006) *The performance evaluation context: The public sector reform and its theoretical references* [El contexto de la evaluación del desempeño: La



- reforma del sector público y sus referentes teóricos*] in “La evaluación del desempeño en la construcción de una nueva gestión pública en México”. El Modelo Integral de Desempeño de Órganos de Vigilancia y Control”. Mexico: Secretaría de la Función Pública.
- Franz, H. (1985) *Interorganizational arrangements and coordination at the policy level in Guidance, control, and evaluation in the public sector. The Bielefeld Interdisciplinary Project*. Berlin: Walter de Gruyter, eds. Franz-Xavier Kaufmann, Giandomenico Majone and Vincent Ostrom.
- Galvez, R. (1985) *Los instrumentos del sistema de control y evaluación gubernamental y su operación en la Administración Pública Federal*, México: Instituto Nacional de Administración Pública (INAP).
- George, A. and Bennett, A. (2005) *Case studies and theory development in the social sciences*. Cambridge: MIT Press, BCSIA studies in international security.
- Gerring, J. (2008) *Case selection for case-study analysis: qualitative and quantitative techniques* in “*The Oxford handbook of political methodology*” ed. Box-Steffensmeier, J., Brady, H. and Collier, D. Oxford: Oxford University Press Inc., ch. 28, pp. 645-684.
- Gómez-Dantés, O., Sesma, S., Becerril, V., Knaul, F., Arreola-Ornelas, H. and Frenk, J. (2011) *Sistema de salud de México [The health system of México]*. Mexico: Instituto Nacional de Salud Pública, vol. 53, issue 2.
- Grimes, C. E. and Simmons, Ch. (1969) *Bureaucracy and Political Control* in “*Mexico: Towards an Assessment*”. Public Administration Review, Vol. 29, No. 1 (Jan-Feb, 1969), pp. 72-79.
- Guba, E. and Lincoln, Y. (1981) *Effective Evaluation. Improving the Usefulness of Evaluation Results Through Responsive and Naturalistic Approaches*. San Francisco: Jossey-Bass Inc., Publishers.
- Guerrero, G. (2002) *Contaduría Mayor de Hacienda de la Cámara de Diputados en El gobierno y la gestión municipal en México, I parte*. Mexico: Instituto Nacional de Administración Pública, A.C. (INAP), Gaceta Mexicana de Administración Pública Estatal y Municipal, núm. 64. Available at: <http://www.juridicas.unam.mx/publica/rev/indice.htm?r=gac&n=64>
- Guerrero, O. (1989) *El Estado y la Administración Pública en México. Una investigación sobre la actividad del Estado Mexicano en retrospectiva y prospectiva*. Mexico: Instituto Nacional de Administración Pública, A.C.
- Guerrero, O. (1995) *El análisis de implementación de políticas públicas [Analysis of public policies' implementation]* in Salazar, C. (ed.) “*Las políticas públicas*. Pontificia Universidad Javeriana”. Facultad de Ciencias Jurídicas, vol. 19 profesores.
- Guerrero, O. (2007) *The professional career civil service in México and academic training problems [El servicio profesional de carrera en México y sus problemas de formación académica]*. Mexico: Revista Buen Gobierno, no. 2, pp. 44-60.
- Gutiérrez, E. et al. (2007) *México: Democracia, participación social y proyecto de nación*. Mexico: Editorial Siglo XXI. Universidad Autónoma de Nuevo León, pp.134-177, 211-237, 251-256.
- Guzmán, M. (2005) *Sistema de control de gestión y presupuesto por resultado: la experiencia Chilena*. Caracas: Centro Latinoamericano de Administración para el Desarrollo (CLAD), X Congreso Internacional del CLAD sobre la reforma del Estado y de la administración pública, Santiago, Chile, 8-21 oct. 2005.
- Hague, R. and Harrop, M. (2004) *Comparative Government and Politics. An Introduction*. New York: Palgrave Macmillan, 6th ed.
- Hall, P. and Taylor, R. (1996) *Political Science and the Three New Institutionalisms*. Cambridge: Harvard University, MPIFG discussion paper 96/6.

- Hantrais, L. and Mangen, S. (eds.) (1996) *Cross-national research methods in the social sciences*. New York: Pinter.
- Hay, C. (1997) *State of the art divided by a common language: political theory and the concept of power*. Oxford: Blackwell publishers, vol. 17, issue 1, pp. 45-52.
- Hay, C. and Wincott, D. (1998) *Structure, Agency and Historical Institutionalism*. Oxford: Political Studies Association, vol. XLVI, issue 5, pp. 951-7.
- Hawkesworth, I., Huerta, O. and Robinson, M. (2013) *Selected budgeting issues in Chile: Performance budgeting, medium-term budgeting, budget flexibility*. Paris: OECD Journal on Budgeting, vol. 2012, issue 3. Available at: [http://www.keepeek.com/Digital-Asset-Management/oecd/governance/selected-budgeting-issues-in-chile\\_budget-12-5k44tpx5bl48#page9](http://www.keepeek.com/Digital-Asset-Management/oecd/governance/selected-budgeting-issues-in-chile_budget-12-5k44tpx5bl48#page9)
- Heady, F. (2000) *Administración pública. Una perspectiva comparada*. Mexico: Fondo de Cultura Económica (1966), pp. 36-63.
- Herbert, J. (2014) *Researching evaluation influence: A review of the literature*. London: Sage publications, Evaluation Review, vol. 38, no. 5, pp. 388-419.
- Hill, M. (2009) *The Public Policy Process*. Essex: Pearson Education Ltd., 5<sup>th</sup> edition.
- Hintze, J. (2001) *Control y evaluación de gestión y resultados*. Buenos Aires: Centro de Desarrollo y Asistencia Técnica en Tecnología para la Organización Pública "TOP". Available at: <http://www.top.org.ar/Documentos/HINTZE,%20Jorge%20-%20Control%20y%20evaluacion.pdf>
- Hintze, J. (2001) *Evaluación del valor público (efectividad versus eficiencia)*. Buenos Aires: Biblioteca virtual TOP sobre Gestión Pública.
- Hintze, J. (2003) *Instrumentos de evaluación de la gestión del valor público*. Panamá: Ponencia presentada para el VIII Congreso Internacional del CLAD sobre la Reforma del Estado y la Administración Pública, 28-31 Oct. 2003.
- Hill, M. (2009) *The Public Policy Process*, 5<sup>th</sup> ed. Essex: Pearson Education Ltd.
- HM Treasury (2011) *The Magenta book. Guidance for evaluation*. London: The National Archives.
- Hogwood, B. and Gunn, L. (1984) *Policy analysis for the real world*. Oxford: Oxford Press.
- Hojlund, S. (2014) *Evaluation use in the organizational context – changing focus to improve theory*. London: Sage publications, vol. 20, no. 1, pp. 26-43.
- Hood, C. (1995) *The "New Public Management" in the 1980s: variation on a theme*. Elsevier Science Ltd., Journal Accounting, Organizations and Society, vol. 20, no. 2&3, pp. 93/109.
- Hood, C., James, O., Jones G., Scott, C. and Travers, T. (1999) *Regulation Inside Government*. Oxford: Oxford University Press.
- Hood, C. and Peters, G. (2004) *The Middle Aging of New Public Management: Into the Age of Paradox?* Oxford: Journal of Public Administration Research and Theory, vol. 14, no. 3, pp. 267-282.
- Horton, D. and Mackay, R. (1999) *Evaluation in Developing Countries: An Introduction*. The Hague: Knowledge, Technology & Policy, winter 1999, vol. 11, no. 4, pp. 5-12.
- House, E. (1983) *Assumptions underlying evaluation models in "Evaluation Models. Viewpoints on Educational and Human Services Evaluation"*, Madaus, George F., Scriven, Michael S. & Stufflebeam, Daniel L. (eds.). Hingham: Kluwer Academic Publishers, chapter 3.
- Howlett, M. and Ramesh, M. (1995) *Studying public policy: policy cycles and policy subsystems*. Oxford: Oxford University Press.
- Howlett, M., Ramesh, M. and Perl, A. (2009) *Studying public policy: policy cycles and policy subsystems*. Oxford: Oxford University Press, 3<sup>rd</sup>. ed..

- Hughes, O. (2003) *Public management and administration. An introduction*. New York: Palgrave Macmillan, 3<sup>rd</sup>. ed.
- INAP (1983) *Exposición de motivos de la Ley de Planeación [Purpose statement of the Planning Law]* in “*El Sistema Nacional de Planeación Democrática*”. Mexico: Revista de Administración Pública, no. 55/56. Available at: [http://www.inap.org.mx/portal/images/REVISTA\\_A\\_P/rap\\_55-56\\_1983.pdf](http://www.inap.org.mx/portal/images/REVISTA_A_P/rap_55-56_1983.pdf)
- Inter-American Development Bank (1997) *Evaluation: a management tool for improving project performance*. Washington, D.C.: Evaluation Office (EVO).
- Jennings, E. and Hall, J. (2012) *Evidence-Based Practice and the use of information in State agency decision making*. Oxford: Journal of Public Administration Research & Theory, 22 (2).
- Johnson, K., Greenesid, L., Toal, S., King, J., Lawrenz, F. And Volkov, B. (2009) *Research on evaluation use. A review of the empirical literature from 1986 to 2005*. Thousand Oaks: Sage publications, American Journal of Evaluation, vol. 30, no. 3, pp. 377-410.
- Jones, L. (2002) *The New Public Management: Improving Research and Policy Dialogue by Michael Barzelay*. The Netherlands: Journal of Comparative Policy Analysis, Book Review, vol. 3, no. 2, pp. 81-89.
- Jones, L.R. and Kettl, D.F. (2003) *Assessing public management reform in an international context*. Philadelphia: Taylor & Francis, Inc., International Public Management Journal 4 (2003), pp. 1-19.
- Jordan, G.B., Hage, J. and Mote, J. (2008) *A theories-based systemic framework for evaluating diverse portfolios of scientific work, part one: micro and meso indicators in “Co-Constructing a Contextually Responsive Evaluation Framework”*. New Directions for Evaluation, vol. 2008, issue 118, pp. 7-24.
- Karlsson, O. and Conner, R. (2006) *The relationship between evaluation and politics in “The SAGE handbook of evaluation”* (Ian F. Shaw, Jennifer C. Greene and Melvin M. Mark eds.). London: SAGE Publications Ltd.
- Kaufmann, F. (1985) *The Relationship between Guidance, Control and Evaluation in “Guidance, control, and evaluation in the public sector. The Bielefeld Interdisciplinary Project”*. Berlin: Walter de Gruyter, eds. Franz-Xavier Kaufmann, Giandomenico Majone and Vincent Ostrom.
- Kaufmann, F., Majone, G. and Ostrom, V. (1985) *Experience, Theory and Design in “Guidance, control, and evaluation in the public sector. The Bielefeld Interdisciplinary Project”*. Berlin: Walter de Gruyter, eds. Franz-Xavier Kaufmann, Giandomenico Majone and Vincent Ostrom.
- Kaufmann, D., Kraay, A. and Mastruzzi, M. (2010) *The Worldwide Governance Indicators: Methodology and analytical issues*. Washington, D.C.: The World Bank, Development Research Group, Macroeconomics and growth team, policy research working paper no. 5430.
- Khan, M. A. (1998) *Evaluation Capacity Building. An Overview of Current Status, Issues and Options*. London: SAGE Publications Ltd., vol. 4, issue 3, pp. 310-328.
- Kirkhart, K. (2000) *Reconceptualizing evaluation use: an integrated theory of influence* in “The expanding scope of evaluation use”. San Francisco: Wiley Periodicals, Inc. and the American Evaluation Association, special issue winter 2000, vol. 2000, issue 88, pp. 5-23.
- Knill, Ch. and Tosun, J. (2012) *Public policy: a new introduction*. Hampshire: Palgrave Macmillan.
- Lanz, J. (1987) *La contraloría y el control interno en México. Antecedentes históricos y legislativos*. Mexico: Secretaría de la Contraloría General de la Federación y Fondo de Cultura Económica.

- Ledermann, S. (2011) *Exploring the necessary conditions for evaluation use in program change*. Thousand Oaks: SAGE journals. American Journal of Evaluation, vol. 34, no. 4, June 2011, pp. 1-20.
- Lehoucq, F. (2007) *Why is Structural Reform Stagnating in Mexico? Policy Reform Episodes from Salinas to Fox*. Mexico: Centro de Investigación y Docencias Económicas (CIDE), working paper no. 195, May 2007.
- Lehoucq, F., Negretto, G., Aparicio, F., Nacif, B. and Benton, A. (2005) *Political Institutions, Policymaking Processes, and Policy Outcomes in Mexico*. Washington, DC: Inter-American Development, Latin American Research Network, Research network working paper num. R-512.
- Leviton, L., and Hughes, E. (1981) *Research on the utilization of evaluations: A review and synthesis*. London: Sage Publications, Evaluation Review, vol. 5, no. 4, pp. 525-548.
- Lijphart, A. (2012) *Patterns of democracy: government forms and performance in thirty-six countries*. New Haven: Yale University Press, 2<sup>nd</sup> ed.
- Llerenas, V. and Huerta, O. (2004) *Country profile. Administrative Reform in Mexico*. London: Routledge, Policy Studies, vol. 25, no. 3. Available at: <http://dx.doi.org/10.1080/0144287012000277507>
- Lonti, Z. and Woods, M. (2008) *Towards Government at a Glance: Identification of Core Data and Issues related to Public Sector Efficiency*. Paris: OECD Working Papers on Public Governance, No. 7, OECD Publishing.
- López, X. (2007) *Law (initiative) of Government Management Evaluation*. Mexico: Chamber of Deputies. Mexican Congress, Gaceta Parlamentaria, year X, no. 2238, annex IV, 24/04/2007.
- Lopez-Acevedo, G., Krause, P. and Mackay, K. (eds.) (2012) *Building Better Policies. The nuts and bolts of monitoring and evaluation systems*. Washington D.C.: International Bank for Reconstruction and Development, The World Bank.
- Lowndes, V. and Roberts, M. (2013) *Why institutions matter: the new institutionalism in political science*. Hampshire: Palgrave Macmillan.
- Lukes, S. (2005) *Power a radical view*. New York: Palgrave Macmillan, 2<sup>nd</sup> ed. Available at: <http://lib.myilibrary.com/Open.aspx?id=85996>
- Manning, N. (2001) *The legacy of the New Public Management in developing countries*. London: International Review of Administrative Sciences, SAGE Publications, vol. 67, pp. 297-312.
- Marcel, M. (2002) *Las opciones para la Reforma del Estado en Chile*, in S. Valdés (ed.), "Reforma de Estado", vol. II: Dirección Pública y Compras Públicas. Santiago de Chile: Centro de Estudios Públicos, pp. 209-282.
- Marcel, M. (2006) *La dirección de presupuestos avances 2000-2005 y desafíos futuros. Olmué 2006*. Santiago de Chile: DIPRES' executive workshops, ppt presentation, January 20<sup>th</sup>, 2006. Available at: <http://bibliotecadigital.dipres.gob.cl/handle/123456789/3683>
- Mark, M. and Henry, G. (2004) *The Mechanisms and Outcomes of Evaluation Influence*. London: Sage Publications Ltd., vol. 10, issue 1, pp. 35-57.
- Mark et al. (2006) *The evaluation of policies, programs, and practices* in "Handbook of Evaluation. Policies, Programs and Practices". London: SAGE Publications Ltd., ed. Ian F. Shaw, Jennifer C. Greene, Melvin M. Mark, introduction.
- Mark, K. and Pfeiffer, J. (2011) *Monitoring and Evaluation in the United States Government: An Overview*. Washington, D.C.: the World Bank Group, Independent Evaluation Group, ECD working paper series, no. 26.
- Mayntz, R. (1994) *Administración y Política* in "Sociología de la Administración Pública". Madrid: Alianza Universidad, pp. 71-92.



- Medina, A. (2007) *El Sistema Nacional de Monitoreo y Evaluación de la Gestión Pública en México*. Washington, D.C.: World Bank and Centro Latinoamericano de Administración para el Desarrollo, Documentos Estudios de Casos del CLAD, no. 3, Fortalecimiento de los sistemas de monitoreo y evaluación (M&E) en América Latina. Available at:  
[http://www.clad.org/siare\\_isis/innotend/evaluacion/mexico.pdf](http://www.clad.org/siare_isis/innotend/evaluacion/mexico.pdf)
- Medina, L. (2000) *Ruptura y cambio en Hacia el nuevo Estado, México 1920-1994*. Mexico: Fondo de Cultura Económica, ch. VIII.
- Mejía, J. (2005) *La evaluación como herramienta para una gestión pública orientada a resultados: la práctica de la evaluación en el ámbito público mexicano [Evaluation as a tool for public management based in results: evaluation practice in the Mexican government]*. CLAD.
- Meyer, L. (2000) *El Presidencialismo Mexicano: En busca del justo medio*. Mexico: Centro de Investigación y Docencia Económicas, A.C. CIDE, serie Istor, dossier, no. 3, winter 2000, pp. 41-57. Available at:  
[http://www.istor.cide.edu/archivos/num\\_3/dossier3.pdf](http://www.istor.cide.edu/archivos/num_3/dossier3.pdf)
- Meyers, M. and Vorsanger, S. (2007) *Street-Level bureaucrats and the Implementation of Public Policy in "Handbook of Public Administration"*. London: SAGE Publications Ltd., ed. Peters, B. Guy and Pierre, Jon, pp. 153-163.
- Mill, J. S. (1843) *System of logic, ratiocinative and inductive. Being a connected view of the principles of evidence, and the methods of scientific investigation*. London: John W. Parker, West strand, vol. I.
- Ministerio de Hacienda (2003) *Systems of Management Control and Results-Based Budgeting. The Chilean Experience*. Chile: Management Control Division, National Budget Office, Jan. 2003.
- Mitnick, B. (1975) *The theory of agency: the policing "paradox" and regulatory behavior*. Public Choice, vol. 24, winter 1975, pp. 27-42
- Moctezuma, E. (1989) *La Secretaría de Programación y Presupuesto a través de sus reglamentos interiores*. Mexico: Instituto Nacional de Administración Pública, Revista de Administración Pública, no. 74, May-Aug.
- Monnier, E. (1996) *Objectives and targets of evaluations [Objetivos y destinatarios de las evaluaciones]*. Mexico: Revista de Administración Pública, Ensayos sobre la administración pública, no. 91, pp. 173-96.
- Muñoz, M. (2005) *Evaluación de resultados en Chile: el Sistema de Seguimiento de la Programación Gubernamental*. Caracas: Centro Latinoamericano de Administración para el Desarrollo (CLAD), case study no. 2 "La evaluación como herramienta para una gestión pública orientada a resultados".
- Muñoz, Ó. and Stefoni, C. (2002 coord.) *El periodo del Presidente Frei Ruiz-Tagle*. Santiago de Chile: Editorial universitaria-FLACSO-Chile.
- Naím, M. (1994) *Economic Reform and Democracy-Latin America: The Second Stage of Reform*. Journal of Democracy, Vol. 5, October.
- Naím, M. (2000) *Washington Consensus or Washington Confusion?* Washington: Washingtonpost Newsweek Interactive, LLC, Foreign Policy, no. 118, spring, pp. 86-103.
- Naím, M. (2000b) *Fads and fashion in economic reforms: Washington Consensus or Washington Confusion?* Third World Quarterly, vol. 21, no. 3, pp. 505-528.
- Navarrete, J. (2008) *Sistema Político Mexicano: Desarrollo y reacomodo del poder. Ensayo*. México: Ibero fórum, Revista Electrónica del Departamento de Ciencias Sociales y Políticas de la Universidad Iberoamericana, A.C., III, No. 6, July- Dec. 2008, pp. 131-148.
- Navia, P. and Velasco, A. (2003) *The Politics of Second-Generation Reforms* in P. P. Kuczynski e J. Williamson, "After the Washington Consensus: Restarting Growth

- and Reform in Latin America*". Washington, D.C.: Institute for International Economics, ch. 10.
- Nef, J. (2007) *Public Administration and Public Sector Reform in Latin America in "Handbook of Public Administration"*. London: SAGE Publications Ltd., ed. Peters, B. Guy and Pierre, Jon, pp. 323-335.
- Nicholson-Crotty, J. and Miller, S. (2012) *Bureaucratic Effectiveness and Influence in the Legislature*. Oxford: Oxford University Press, *Journal of Public Administration Research and Theory*, vol. 22, issue 2, April 2012, pp. 347-71.
- OECD (1998) *OECD Economic surveys 1997-1998. Special features health care reform. México*. Paris: Organisation for Economic Co-operation and Development.
- OECD (2009) *Managing for Development Results*. Policy Brief. Paris: Organisation for Economic Co-Operation and Development. Available at:  
<http://www.oecd.org/dataoecd/18/53/42447575.pdf>
- OECD (2009b) *Review of Budgeting in Mexico*. Paris: OECD, *Journal on Budgeting*, vol. 2009, supplement 1.
- OECD (2011) *Health at a Glance 2011: OECD Indicators*. Paris: OECD Publishing. Available at: <http://www.oecd.org/health/healthpoliciesanddata/49105858.pdf>
- OECD (2011) *Why is health spending in the United States so high?* In "*Health at a Glance 2011: OECD indicators*". Paris: Organisation for Economic Co-Operation and Development. Available at:  
<http://www.oecd.org/health/health-systems/health-at-a-glance.htm>
- OECD (2013) *Health at a glance 2013 OECD indicators*. Paris: Organisation for Economic Co-Operation and Development. Available at:  
<http://www.oecd.org/els/health-systems/Health-at-a-Glance-2013.pdf>
- Ormond, D. and Löffle, E. (1999) *New Public Management: what to take and what to leave? "Nueva Gerencia Pública: ¿qué tomar y qué dejar?"*. Caracas: Centro Latinoamericano de Administración para el Desarrollo, *Revista del CLAD Reforma y Democracia*, no. 13, Feb. 1999.
- Oropeza, A. (2005) *La evaluación de la función pública en México [Evaluation of public management in México]*. Mexico: Plaza y Valdés, S.A. de C.V.
- Osborne, D. & Gaebler, T. (1992) *Reinventing government: how the entrepreneurial spirit is transforming the public sector*. Reading, Mass.: Addison-Wesley.
- Ospina, S., Cunill, N., and Zaltsman, A. (2004) *Performance evaluation, public management improvement and democratic accountability*. London: Routledge Taylor & Francis Ltd., *Public Management Review*, vol. 6, issue 2, pp. 229-51.
- Ostrom, V. and Ostrom, E. (1965) *A behavioral approach to the study of intergovernmental relations*. Thousand Oaks: SAGE Publications, *The ANNALS of the American Academy of Political and Social Science*, pp. 137-146.
- O'Toole, L. (2007) ed. Peters, B.G. and Pierre, J. *Interorganizational Relations in Implementation in "Handbook of Public Administration"*. London: SAGE Publications Ltd., pp. 143-150.
- Page, E. (2012) *Policy without Politicians. Bureaucratic Influence in Comparative Perspective*. Oxford: Oxford University Press.
- Palumbo, D. and Wright, P. (1981) *Decision Making and Evaluation Research in "Evaluating and Optimizing Public Policy"*. Massachusetts: Lexington Books D.C. Heath and Company, edited by Palumbo, Dennis J. Fawcett, Stephen B., and Wright Paula.
- Pardo, M. (1999) *Viabilidad política del Servicio Civil de Carrera en México in "El Servicio de Carrera en la Administración Pública Mexicana" [Political feasibility of Civil Career Service in Mexico in "Civil Career in Mexican Public Administration"]*. Mexico: Instituto Nacional de Administración Pública, *Revista de Administración Pública*, no. 103. Available at:

- [http://www.inap.mx/portal/images/REVISTA\\_A\\_P/rap\\_103\\_1999.pdf](http://www.inap.mx/portal/images/REVISTA_A_P/rap_103_1999.pdf)
- Pardo, M. (2005) *El Servicio Profesional de Carrera en México: de la tradición al cambio*. Mexico: El Colegio de México, Foro internacional, vol. XLV, no. 4, oct.-dec. 2005, pp. 599-634. Available at: <http://www.jstor.org/stable/27738732>
- Pardo, M. (2007) *La gerencialización de la administración pública*. Mexico: El Colegio de México, Foro internacional, vol. XLVII, no. 4, oct.-dec. 2007, pp. 895-925. Available at: <http://redalyc.uaemex.mx/pdf/599/59911153007.pdf>
- Parsons, W. (1995) *Public policy: An introduction to the theory and practice of policy analysis*. Cheltenham: Edward Elgar Publishing Ltd.
- Patton, M. (2004) *The roots of utilization-focused evaluation* in "Evaluation roots: tracing theorists' views and influences". London: Sage Publications Ltd. ed. Alkin, M., chapter 18.
- Patton, M. (2012) *Essentials of utilization-focused evaluation*. London: Sage publications Ltd.
- Pawson, R. and Tilley, N. (1994) *What works in evaluation research?* Oxford: British Journal of Criminology, vol. 34, no.3, summer 1994, pp. 291-306.
- Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*. London: SAGE Publications, Ltd.
- Pennings, P., Keman, H. and Kleinnijenhuis, J. (2006) *Doing research in Political Science. An introduction to comparative methods and statistics*. London: SAGE Publications Ltd., 2<sup>nd</sup>. ed.
- Peters, G. (1996) *The Future of Governing: Four Emerging Models*. Kansas: University Press of Kansas.
- Peters, G. (1999) *La política de la Burocracia*. Mexico: Fondo de Cultura Económica, pp. 7-61.
- Peters, G. (2006) ed. Peters, G. and Pierre, J. *Concepts and Theories of Horizontal Policy Management* in "Handbook of Public Policy". London: SAGE Publications Ltd., pp. 115-138
- Peters, G. (2012) *Institutional theory in political science: the new institutionalism*. New York: The Continuum International Publishing Group, 3<sup>rd</sup>. ed.
- Pichardo, I. (1972) *10 años de planificación y administración pública en México (ensayos)*. Mexico: Instituto de Administración Pública.
- Pimienta, C. (2002) *Sistemas de evaluación de la gestión pública* in Armijo, M. and Tomassini, L. (eds.) "Reforma y modernización del Estado. Experiencias y desafíos". Santiago de Chile: LOM ediciones and Instituto de Asuntos Públicos, Universidad de Chile.
- Pollitt, Ch. (2006a) *Performance management in practice: a comparative study of executive agencies*. Oxford: Oxford University Press, J Public Adm Res Theory, January 2006, 16 (1), pp. 25-44.
- Pollitt, Ch. (2006b) *Who are we, what are we doing, where are we going? A perspective on the performance management community*. Leuven: Katholieke Universiteit Leuven, keynote speech to the second transatlantic dialogue "A performing public sector", 1 June, 2006.
- Prats, J. (1997) *La Reforma del Estado y de la Administración Pública*. Venezuela: CLAD, presentation in the II International Congress of the CLAD.
- Presidencia de la República (1977) *Base legal y lineamientos para el Funcionamiento Institucional y Sectorial de la Administración Pública Federal*. Mexico: Colección Lineamientos. Coordinación General de Estudios Administrativos.
- Presidencia de la República (2001) *Modelo Estratégico para la Innovación y la Calidad Gubernamental*. Mexico: Oficina de la Presidencia para la Innovación Gubernamental.

- Presidencia de la República (2002) *Agenda of Good Government [Agenda de Buen Gobierno]*. Mexico: Oficina de la Presidencia para la Innovación Gubernamental, Nov. 2002. Available at: <http://innova.fox.presidencia.gob.mx/funcionarios/formacion/archivos/teleconferencia01/apbg.pdf>
- Presidencia de la República (2006) *Informe de Rendición de Cuentas. Administración 2000-2006*. Mexico: Oficina de la Presidencia para la Innovación Gubernamental. Available at: <http://fox.presidencia.gob.mx/archivos/7/0/3/7/2/files/archivos/sip-3801.pdf>
- Presidencia de la República (2006a) *Informe de Ejecución del Plan Nacional de Desarrollo*. Mexico: Oficina de la Presidencia.
- Presidencia de la República (2006b) *Sexto Informe de Gobierno*. México: Oficina de la Presidencia.
- Presidencia de la República (2013) *Programa de Gobierno 2014-2018*. Santiago de Chile: Gobierno de Chile.
- Radin, B. A. (2000) *Intergovernmental relationships and the federal performance movement*. Oxford: Publius, The Journal of Federalism 30: 1-2, winter/spring 2000.
- Radin, B. (2007) *The Instruments of Intergovernmental Management in "Handbook of Public Administration"*. London: SAGE Publications Ltd., ed. Peters, B. Guy and Pierre, Jon, pp. 365-376.
- Ragin, Ch. (2014) *The comparative method. Moving beyond qualitative and quantitative strategies*. Oakland: University of California Press.
- Rebollo, Á. (1993) *La política pública de la desincorporación de entidades paraestatales [Public policy of disincorporation of State enterprises]*. Mexico: INAP, Revista de Administración Pública, no. 84, Jan.-June.
- Repetto, F. (2003) *Institutional capability: a way to faced challenges of social policy "Capacidad Institucional: un camino para enfrentar los desafíos de la política social"*. Buenos Aires: Tecnología para la Organización Pública, A.C., no. 5.
- Reyna, J. (2006) *México: una democracia incipiente. En publicación: En los intersticios de la democracia y el autoritarismo. Algunos casos de Asia, África y América Latina*. Cornejo, Romer. Buenos Aires: CLACSO (Consejo Latinoamericano de Ciencias Sociales), Oct. 2006, pp. 131-171.
- Rhodes, R.A.W. (1991) *The New Public Management*. Special issue of Public Administration, vol. 69:1, spring 1991.
- Rhodes, R.A.W. (1996) *The New Governance: Governing without Government*. Oxford: Political Studies Association, XLIV, pp. 652-667.
- Rist, R., Boily, M., and Martin, F. (coord.) (2011) *Influencing change. Building evaluation capacity to strengthen governance*. Washington, D.C.: The International Bank for Reconstruction and Development and The World Bank.
- Rogers, P. (2007) *Theory-Based Evaluation: Reflections Ten Years On*. San Francisco: Jossey-Bass, New Directions for Evaluation, no. 114, published in Wiley InterScience.
- Rojas, G. (2006) *Cambios recientes en el empleo público en México, la situación del trabajo en México 2006*. México: Uam-let-Afl/Cio-Plaza y Valdés.
- Rojas, F., Mackay, K., Matsuda, Y., Shepherd, G., Del Villar, A., Zaltsman, A. and Krause, P. (2005) *Chile: análisis del programa de evaluación del gasto público*. Washington, D.C.: International Bank for Reconstruction and Development, The World Bank, estudios del Banco Mundial sobre América Latina y el Caribe.
- Rose, N., O'Malley, P. and Valverde, M. (2006) *Governmentality*. Annual Review of Law and Social Science, vol. 2, pp. 83-104, Sydney Law School research paper no. 09/94.
- Ross, S. (1973) *The economic theory of agency: The principal's problem*. Nashville: American Economic Association, American Economic Review, Papers and



- Proceedings of the Eighty-fifth Annual Meeting of the American Economic Association, vol. 63, no. 2, pp. 134-139.
- Rossi, P., Lipsey, M. and Freeman, H. (2004) *Evaluation: a systematic approach*. London: Sage Publications Ltd., 7<sup>th</sup> edition.
- Rubin, I. (2010) *Public budgeting: The concept of budgeting as political choice* in "Public administration. Concepts and cases". Boston: WADSWORTH CENGAGE Learning, pp. 345-59.
- Rubin, I. and Kelly, J. (2005) ed. Ferlie, E., Lynn Jr., L. and Pollitt, Ch. *Budget and accounting reforms* in "The Oxford handbook of public management". Oxford: Oxford University Press.
- Rubio, G. (2012) *The Mexican government's M&E system* in G. Lopez-Acevedo, P. Krause and K. Mackay (eds.) "Building Better Policies. The nuts and bolts of monitoring and evaluation systems". Washington D.C.: International Bank for Reconstruction and Development, The World Bank, ch. 12, pp. 167-182
- Sáenz, K., Casillo, J. and Guillen, A. (2007) *El sistema electoral mexicano, panorama de las reformas electorales desde 1990*. Dialnet, Letras Jurídicas, revista electrónica de derecho, no. 5. Available at:  
<http://letrasjuridicas.cuci.udg.mx/numeros/articulos5/Dra%20%20Karla%20Saenz%20%20Articulo.doc>
- Sánchez, G. (1993) *Breves consideraciones sobre aspectos jurídicos de la planeación nacional y la planeación urbana en México*. Mexico: Universidad Nacional Autónoma de México and Instituto de Investigaciones Jurídicas, Boletín Mexicano de Derecho Comparado, No. 78, Sep.-Dec. 1993.
- Sánchez, G. (1996) *Evolución legislativa de la planeación del desarrollo y la planeación urbana en México*. Mexico: Universidad Nacional Autónoma de México and Instituto de Investigaciones Jurídicas, Boletín Mexicano de Derecho Comparado, No. 86, May-Aug. 1996.
- Sanderson, I. (2000) *Evaluation in complex policy systems*. London: SAGE Publications, vol. 6 (4): 433-454.
- Sanderson, S. (1983) *Presidential Succession and Political Rationality in Mexico*. Cambridge: World Politics, Vol. 35, No. 3, pp. 315-334.
- Sartori, G. (2004) *Where political science goes? [Hacia dónde va la ciencia política?]*. Mexico: Centro de Investigación y Docencia Económicas (CIDE), Revista política y gobierno, vol. XI, no. 2, II semestre 2004, pp. 349-354.
- Schedler, K. (2003) *... and politics? Public management development in the light of two rationalities* in "Public Management Review". London: Routledge, Vol. 5, issue 4.
- Schwandt, T. and Burgon, H. (2006) *Evaluation and the study of lived experience* in "The SAGE handbook of evaluation" Ian F. Shaw, Jennifer C. Greene and Melvin M. Mark (eds.). London: SAGE Publications Ltd.
- Scriven, M. (1991) *Evaluation thesaurus*. London: SAGE Publications Ltd., 4<sup>th</sup> edition.
- Secretaría de Contraloría y Desarrollo Administrativo (1999) *Planeación Estratégica (Guía Técnica). Metodología de Modernización y Reforma al Sistema Presupuestario en la Administración Pública*. Mexico: Unidad de Desarrollo Administrativo.
- Secretaría de Contraloría y Desarrollo Administrativo (1999) *Sistema de Medición a través de Indicadores. Metodología de Modernización y Reforma al Sistema Presupuestario en la Administración Pública*. Mexico: Unidad de Desarrollo Administrativo.
- Secretaría de Hacienda y Crédito Público (2007) *Sistema de Evaluación del Desempeño (SED)*. Mexico: SHCP. Available at:  
[http://www.shcp.gob.mx/EGRESOS/pbr/docs\\_consultta/directricessed.pdf](http://www.shcp.gob.mx/EGRESOS/pbr/docs_consultta/directricessed.pdf)

- Secretaría de Hacienda y Crédito Público (2011) *Proyecto de Presupuesto de Egresos de la Federación para el ejercicio fiscal 2011, Exposición de motivos 2011 – Avances para mejorar la eficiencia y la calidad del gasto público* -. Mexico: SHCP.
- Secretaría de la Contraloría General de la Federación (1985) *Informe de labores al H. Congreso de la Unión*. Report of activities. Mexico: Talleres Gráficos de la Nación.
- Secretaría de la Contraloría General de la Federación (1992-1993) *Informe de labores al H. Congreso de la Unión*. Report of activities. Mexico: Talleres Gráficos de la Nación.
- Secretaría de la Función Pública (2004) *The Integral Performance Model of Control and Oversight Bodies MIDO, Manual of Operation, Internal Control Bodies*. Mexico: Coordinación General de Órganos de Vigilancia y Control, dec. 2004. Available at: [www.mido.gob.mx/regulacion/MO%20MIDO%20OICs%202005.pdf](http://www.mido.gob.mx/regulacion/MO%20MIDO%20OICs%202005.pdf)
- Secretaría de la Presidencia (1967) *Informe sobre la reforma de la administración pública mexicana*. Mexico: Comisión de Administración Pública, Coordinación General de Estudios Administrativos, Colección: Fuentes para el estudio de la Administración Pública Mexicana, serie a) documental, no. 1, edición facsimilar del original.
- Secretaría de Programación y Presupuesto (1981) *Primer seminario de evaluación en la administración pública federal*. Mexico: Memora, Morelos, June 1981.
- Secretaría de Programación y Presupuesto (1981a) *El Sistema Nacional de Planeación y la Instrumentación del Plan Global de Desarrollo*. Mexico, 2nd. ed.
- Secretaría de Programación y Presupuesto (1988) *Planeación democrática*. Mexico: Fondo de Cultura Económica, Tercer título de la colección, Cuadernos de Renovación Nacional.
- Sharpe, L. (1985) *Intergovernmental Policy-Making: The Limits of Subnational Autonomy* in “Guidance, control, and evaluation in the public sector. The Bielefeld Interdisciplinary Project”. Berlin: Walter de Gruyter, ed. Franz-Xavier Kaufmann, Giandomenico Majone and Vincent Ostrom.
- Simmons, H. (2006) *Ethics in evaluation* in “Handbook of Evaluation. Policies, Programs and Practices”. London: SAGE Publications Ltd., ed. Ian F. Shaw, Jennifer C. Greene, Melvin M. Mark, ch. 11.
- Sosa, J. (2011) coord. Ramos, J., Sosa, J. and Acosta, F. *Bases conceptuales para el abordaje de la política federal de evaluación en México: una propuesta analítica [Conceptual basis for approaching ofevaluation federal policy in Mexico: an analytical proposal]* in “La evaluación de políticas públicas en México”. Mexico: Instituto Nacional de Administración Pública, A.C. and El Colegio de la Frontera Norte, pp. 101-120. Available at: <http://www.inap.org.mx/portal/images/RAP/evaluacion%20politicas%20publicas.pdf>
- Stake, R. (2004) *Standards-Based and Responsive Evaluation*. Thousand Oaks, CA: Sage.
- Stufflebeam, D. (1983) *The CIPP model for program evaluation* in “Evaluation Models. Viewpoints on Educational and Human Services Evaluation”, Madaus, George F., Scriven, Michael S. & Stufflebeam, Daniel L. (eds.). Hingham: Kluwer Academic Publishers, ch. 7.
- Stufflebeam, D. (2004) *The 21<sup>st</sup>-century CIPP model. Origins, development, and use* in “Evaluation roots: tracing theorists’ views and influences”. London: Sage Publications Ltd. ed. Alkin, M., chapter 16.
- Stufflebeam, D. and Shinkfield, A. (2007) *Evaluation theory, models, and applications*. San Francisco: Jossey-Bass.
- Stufflebeam, D. and Webster, W. (1983) *An analysis of alternative approaches to evaluation* in “Evaluation Models. Viewpoints on Educational and Human

- Services Evaluation*", Madaus, George F., Scriven, Michael S. & Stufflebeam, Daniel L. (eds.). Hingham: Kluwer Academic Publishers, chapter 2.
- Suchman, E. (1967) *Evaluative research: principles and practice in public service and social action programs*. New York: Russell Sage.
- Sundell, K. (2014) *How do government agencies use evidence?* Stockholm: Socialstyrelsen: The National Board of Health and Welfare, Report for National Board of Health and Welfare, Stockholm, Sweden.
- Talbot, C. (2010) *Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom*. Washington, D.C.: Independent Evaluation Group, the World Bank Group, ECD working paper series, no. 24.
- Tamayo, J. (1992) *Las entidades paraestatales en México. Origen, evolución y perspectivas [State enterprises in Mexico. Origin, evolution and perspectives]*. Mexico: Instituto Nacional de Administración Pública, Revista de Administración Pública.
- Tandon, A., Evans, D., Murray, Ch., Lauer, J. (n.d.) *Measuring overall health system performance for 191 countries*. Geneva: World Health Organisation. Available at: <http://www.who.int/healthinfo/paper30.pdf>
- Teirlinck, P., Delanghe, H., Padilla, P., and Verbeek, A. (2013) *Closing the policy cycle: Increasing the utilization of evaluation findings in research, technological development and innovation policy design*. Oxford: Oxford University Press, Science and Public Policy no. 40, pp. 366-77.
- Tello, F. (2011) *La política de reforma y modernización de la gestión pública en Chile. Actores y procesos*. Santiago de Chile: Universidad de Talca, UNIVERSUM, no. 26, vol. 2, pp. 245-65.
- Teutli, G. (1984) *La evaluación de la gestión pública: El papel de los comisarios y delegados de la SECOGEF in El Sistema Nacional de Control y Evaluación*. Mexico: Instituto Nacional de Administración Pública, Revista de Administración Pública, no. 57-58, Jan.-June. Available at: <http://www.juridicas.unam.mx/publica/rev/indice.htm?r=rap&n=57>
- The Democracy Ranking Association (2015) *The democracy ranking of the quality of democracy*. Vienna: the DRA. Available at: <http://democracyranking.org/wordpress/>
- The Economist Intelligence Unit (2014) *Democracy index 2014. Democracy under stress*. London: the EIU. Available at: <http://www.eiu.com/>
- The Inter-American Development Bank. *The Governance Indicators database*. Available at: <http://www.iadb.org/DataGob/>
- The World Bank (2005) *Institutionalizing M&E systems in Latin American and Caribbean Countries*. Washington, D.C., USA, PREMnotes, number 102.
- The World Bank (2005b) *Towards the institutionalization of monitoring and evaluation systems in Latin America and the Caribbean: Proceedings of a World Bank/Inter-American Development Bank Conference*. Washington, D.C.: The World Bank and Inter-American Development Bank, May, E., Shand, D., Mackay, K., Rojas, F. and Saavedra, J. eds. Available at: [www.worldbank.org/ieg/ecd/docs/proceedings\\_la\\_eng.pdf](http://www.worldbank.org/ieg/ecd/docs/proceedings_la_eng.pdf)
- The World Bank. *The Worldwide Governance Indicators*. Available at: [http://info.worldbank.org/governance/wgi/mc\\_chart.asp#](http://info.worldbank.org/governance/wgi/mc_chart.asp#)
- The World Economic Forum (2012) *The Global Competitiveness Report 2012-2013*. Geneva: World Economic Forum, Schwab, Klaus (ed.), Sala-i-Martin, X. and Breda, B. Available at: [www3.weforum.org/docs/WEF\\_GlobalCompetitivenessReport\\_2012-13.pdf](http://www3.weforum.org/docs/WEF_GlobalCompetitivenessReport_2012-13.pdf)

- Ugalde, L. (2014) *La negociación política del presupuesto en México 1997-2012. El impacto sobre la composición y ejecución del gasto público*. Mexico: Información y Tecnología para Asuntos Públicos, S.C., INTEGRALIA, Editores e Impresores Profesionales, EDIMPRO/SITESA and Senate (Upper Chamber).
- United Nations Development Group (2011) *Results-Based Management handbook. Harmonizing RBM concepts and approaches for improved development results at country level*. New York: United Nations Development Group. Available at: <http://www.undg.org/docs/12316/UNDG-RBM%20Handbook-2012.pdf>
- Uvalle, Ricardo (1994). *La Gerencia Pública: una opción para mejorar la acción del buen gobierno*. México, FCPyS-UNAM, Revista Estudios Políticos, no. 5, pp. 95-112.
- Vázquez, D. (2006) *El Gobierno Dividido en México. 2000-2006*. Article published on the web. Available at: <http://expectativasracionales.wordpress.com/2006/10/01/%e2%80%9cel-gobierno-dividido-en-mexico-2000-2006%e2%80%9d/>
- Vásquez, P. (2010) *La planeación, el presupuesto y las políticas públicas in "Los grandes problemas de Mexico" [Planning, Budget and Public Policies in 'The major problems of Mexico']*. Mexico: El Colegio de Mexico, coord. José Luis Méndez, v. XIII, Public Policies. Available at: <http://2010.colmex.mx/16tomos/XIII.pdf>
- Vedung, E. (2006) *Evaluation Research in "The Oxford handbook of public policy"*. Oxford: Oxford University Press, ed. Michael Moran, Martin Rein and Robert E. Goodin, chapter 23.
- Verhoest, K., Peters, B. G., Bouckaert, G. and Verschuere, B. (2004) *The study of organisational autonomy: a conceptual review*. Public Administration and Development, vol. 24, issue 2, pp. 101-118, published in Wiley InterScience.
- Waissbluth, M. (2005) *Quince años de profesionalización de la gerencia pública de Chile*, paper delivered at the 'X Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública'. Santiago de Chile: CLAD, october, 2005.
- Walter, I., Nutley, S. and Davied, H. (2005) *What works to promote evidence-based practice? A cross-sector review*. Evidence & Policy: a journal of research, debate and practice, vol. 1, no. 3, sept. 2005, pp. 335-363.
- Weiss, C. (1972) *Evaluation research: methods of assessing program effectiveness*. Englewood Cliffs, N.J.: Prentice Hall, Inc.
- Weiss, C. (1979) *The many meanings of research utilization*. Public Administration Review, vol. 39, 426-431.
- Weiss, C. (1993) *Where Politics and Evaluation Research Meet*. Thousand Oaks: Sage Publications Ltd., American Journal of Evaluation, vol. 14, no. 1, pp. 93-106.
- Weiss, C. (1997) *Theory-Based Evaluation: Past, Present, and Future*. San Francisco: Jossey-Bass Publishers, New Directions for Evaluation, no. 76.
- Weiss, C. (1998) *Have we learned anything new about the use of evaluation?* Thousand Oaks: Sage Publications Ltd., American Journal of Evaluation, vol. 19, march 1998, no. 1, pp. 21-33.
- Weiss, C. (2004) *Rooting for evaluation. A cliff notes version of my work in "Evaluation roots: tracing theorists'views and influences"*. London: Sage Publications Ltd. ed. Alkin, M., chapter 9.
- Weiss, C. (2007) *Theory-Based Evaluation: Past, Present, and Future*. San Francisco: Jossey-Bass, New Directions for Evaluation, no. 114, published in Wiley InterScience.
- Weiss, R. & Rein, M. (1983) *The evaluation of broad-aim programs: experimental design, its difficulties, and an alternative in "Evaluation Models. Viewpoints on Educational and Human Services Evaluation"*, Madaus, George F., Scriven, Michael S. & Stufflebeam, Daniel L. (eds.). Hingham: Kluwer Academic Publishers, chapter 8.



- Wiesner, E. (2002) *La evaluación de resultados en la modernización del Estado en América Latina [The assessment of results in State modernization in Latin America]*. Bogotá: Revista de Economía Institucional, Universidad Externado de Colombia, primer semestre, año/vol. 4, núm. 006, pp. 132-152.
- Wiesner, E. (2011) *The Evaluation of Macroeconomic Institutional Arrangements in Latin America* in "Influencing change. Building evaluation capacity to strengthen governance". Washington, D.C.: The International Bank for Reconstruction and Development and The World Bank, coord. Rist, Ray C., Boily, Marie-Helene, and Martin, Frederic.
- Williamson, J. (2002) *Did the Washington Consensus Fail?* Washington, D.C.: Peterson Institute for International Economics, speech at the Center for Strategic & International Studies, Nov. 6, 2002.
- Williamson, J. (2003) *From Reform Agenda. A short history of the Washington Consensus and suggestions for what to do next*. Finance and Development, Sept. 2003.
- Wollmann, H. (2003) *Evaluation in public-sector reform: Toward a 'third wave' of evaluation?* in "Evaluation in Sector Reform. Concepts and Practice in International Perspective". Cheltenham: Edgar Elgar Publishing, Inc.
- World Bank (2004) *Monitoring & Evaluation: Some Tools, Methods & Approaches*. Washington, D.C., USA. Available at:  
<http://web.worldbank.org/WBSITE/EXTERNAL/EXTOED/EXTEVACAPDEV/O,,contentMDK:22293257~menuPK:4585748~pagePK:64829573~piPK:64829550~theSitePK:4585673,00.html>
- World Health Organisation (2011) *Monitoring, evaluation and review of national health strategies. A country-led platform for information and accountability*. Geneva: WHO.
- Wright, D. (1974) *Intergovernmental relations: an analytical overview*. Thousand Oaks: SAGE Publications, the ANNALS of the American Academy of Political and Social Science, vol 416, issue 1, pp. 1-16.
- Wu, X., Ramesh, M., Howlett, M. and Fritzen, S. (2010) *The public policy primer: managing the policy process*. Oxon: Routledge.
- Zabaleta, D. (2008) coord. González, A. *Una mirada institucional a la política federal de evaluación del desempeño en México: algunas implicaciones a la luz del paradigma gerencialista [An institutional view of the federal policy of performance evaluation in Mexico: some implications related with managerialist paradigm]* in "Gobernar por resultados? Implicaciones de la política de evaluación del desempeño del gobierno mexicano". Mexico: Gestión Social y Cooperación, A.C., pp. 42-73. Available at:  
[http://www.iapqroo.org.mx/website/biblioteca/LibroGobernarporResultados%20\(1\).pdf](http://www.iapqroo.org.mx/website/biblioteca/LibroGobernarporResultados%20(1).pdf)
- Zaltsman, A. (2009) *The effects of performance information on public resource allocations: A study of Chile's performance based budgeting system*. London: International Public Management Journal, vol. 12, no. 4, pp. 450-83.

**Website:**

<http://bordepolitico.com/HOME/>

<http://curul501.org/>

<http://fundar.org.mx/index.html/>

<http://www.presupuestoabierto.mx/>

<http://www.transparenciapresupuestaria.gob.mx/es/PTP/SED#Evaluaciones>

<http://visionlegislativa.com/>

[http://www.shcp.gob.mx/EGRESOS/PEF/programacion/programacion\\_16/lineamientos\\_programacion\\_presupuestacion\\_2016.pdf](http://www.shcp.gob.mx/EGRESOS/PEF/programacion/programacion_16/lineamientos_programacion_presupuestacion_2016.pdf)

<http://www.imss.gob.mx/transparencia/evaluaciones-externas>

<http://www.imss.gob.mx/imss-prospera/evaluacion>

[http://portal.salud.gob.mx/codigos/columnas/evaluacion\\_programas/evaluacion.html](http://portal.salud.gob.mx/codigos/columnas/evaluacion_programas/evaluacion.html)

<http://www2.issste.gob.mx:8080/index.php/informes-y-programas>

<http://www.coneval.gob.mx/Evaluacion/Paginas/Evaluaciones-y-resultados-de-programas.aspx>

<http://sn.dif.gob.mx/transparencia/transparencia-focalizada/planeacion-institucional-2/>

<http://www.inmujeres.gob.mx/index.php/trans/transparencia-focalizada/evaluacionesexternas>

<http://www.dipres.gob.cl/594/w3-propertyvalue-15697.html>

Annex 1

**Measurement of independent variables**

| Variable      | Intra-governmental relations |                      |        |                         |                               | Quality of bureaucracy        | Level of democracy                                |                | Policy evaluation framework                        |                          |                                     |  |  |                      | Degree of autonomy |                    |                |
|---------------|------------------------------|----------------------|--------|-------------------------|-------------------------------|-------------------------------|---|----------------|--|--------------------------|-------------------------------------|--|--|----------------------|--------------------|--------------------|----------------|
| Indicator     | Forms of coordination        | Information          |        |                         | Economic resources (treasure) | Government effectiveness (WB) | Authoritarian regimes/ hybrid or flawed democracy | Full democracy | Coherence in the design                            |                          |                                     | Structure of monitoring and evaluation system  |  |                      | Policy autonomy    | Financial autonomy | Legal autonomy |
| Sub-indicator | Strong/weak -minimum         | Availability of data | Timing | Reliability and quality | Even/Uneven                   |                               |   |                | Matching policy evaluation objectives and outcomes | Overlap with regulations | Holding hearings of interest groups | Technical operationalization on EASY/DIFFICULT | Perception of operationalization on EASY/DIFFICULT | Methodological rigor | High/Low           | High/Low           | High/Low       |
| México        | 0                            | 1                    | 0      | 0                       | 1                             | 1                             | 0   | 0              | 0  | 0                        | 0                                   | 0  | 0  | 0                    | 0                  | 0                  | 0              |
| Chile         | 0                            | 0                    | 0      | 1                       | 0                             | 1                             | 0   | 0              | 0  | 0                        | 0                                   | 0  | 1  | 1                    | 0                  | 0                  | 0              |

Note: 0 = condition low/outcome absent; 1 = condition high/outcome present.  
Source: author's own elaboration.

Annex 2

**THE EFFECTIVENESS OF POLICY EVALUATION**

This questionnaire is part of the research strategy to find out the differences in the effectiveness of policy evaluation across countries. Therefore, your point of view is highly valuable for understanding of the way in which evaluation is managed across countries. Thanks for your response.

1. Gender:    F \_\_\_\_ M \_\_\_\_
2. What is your scholar degree? BA \_\_\_\_    MA \_\_\_\_    PhD \_\_\_\_    Another \_\_\_\_
3. Job position \_\_\_\_\_
4. How many years of experience do you have?  
    In government \_\_\_\_\_  
    In the evaluation area \_\_\_\_\_
5. What is your role regarding programmes' evaluation?  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
6. Can you describe the framework of policy evaluation (Annual Evaluation Programme, mechanism to follow recommendations, time of execution, coordination between the executive and legislative)?  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
7. Can you describe your programme/institution's approach to developing and implementing evaluation?  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
8. Do you utilise evaluation results?                      YES      NO      \_\_\_\_      \_\_\_\_
9. For what purpose do you use them? \_\_\_\_\_  
\_\_\_\_\_

---

---

*Questionnaire: The effectiveness of policy evaluation.*

10. At what stage of the policy process? \_\_\_\_\_  
\_\_\_\_\_
11. Is the data resulting from the evaluation?  
Available on time \_\_\_\_\_  
Reliable \_\_\_\_\_  
Good quality \_\_\_\_\_
12. Please describe whether and in what ways the following factors have facilitated or hindered the development of evaluation in your programme/institution?  
Economic resources \_\_\_\_\_  
Human resources \_\_\_\_\_  
Time \_\_\_\_\_  
Evaluation system \_\_\_\_\_
13. How does the relationship between the strategy coordinators and your institution influence the effectiveness of evaluation?  
\_\_\_\_\_
14. How does the relationship between Congress and your institution influence the effectiveness of evaluation?  
\_\_\_\_\_
15. Can you describe the form of coordination between the executive and the legislative regarding evaluation?  
\_\_\_\_\_  
\_\_\_\_\_
16. Considering that autonomy is when the public agency can decide for itself which matters it considers most important and that decisions concerning policy-making are independent and not subject to influence from by factors and actors, can you describe the level of autonomy do you have regarding evaluation?  
\_\_\_\_\_  
\_\_\_\_\_

Annex 3

## Office letter 1761. DIPRES in Chile



ORD. N° **1761** /

**ANT.** Solicitud de acceso a la información pública N° AE002W-405, de fecha 03 de noviembre de 2014.

**MAT.** Da respuesta que indica.

**SANTIAGO, 24 NOV. 2014**

**DE : DIRECTOR DE PRESUPUESTOS**

**A : BLANCA LÓPEZ**  
**bl84@sussex.ac.uk**

En el marco de la Ley N° 20.285 sobre Acceso a la Información Pública, esta Dirección recibió su solicitud de acceso identificada en el antecedente, referida a obtener:

"... le agradezco mucho a la unidad su respuesta a mi solicitud anterior. En dicha respuesta se menciona que existe una evaluación comprehensiva del gasto del programa "Red de urgencia", pero no la encuentro ni en el sistema DIPRES, ni en la página del MINSAL, me pueden hacer el favor de informarme como localizarla?"

Por último, la ev. comprehensiva del gasto es la misma que la ev. del gasto institucional?" [sic]

Tal como se señaló anteriormente en respuesta a la solicitud de información N° AE002W-394, se reitera que la evaluación del gasto institucional del programa "Red de Urgencia" se encuentra actualmente en desarrollo, por lo que no se cuenta con un informe final para publicar.

Respecto a la segunda consulta, la evaluación comprehensiva del gasto (ECG) fue reemplazada por la evaluación del gasto institucional (EGI) a partir del año 2011, con la finalidad de obtener información más pertinente para el sector, que cumpla con ser precisa, útil y oportuna.

Saluda atentamente a Ud.,

  
**SERGIO GRANADOS AGUILAR**  
**DIRECTOR DE PRESUPUESTOS**

  
PDA/AAM

Incluye: lo señalado

**DISTRIBUCIÓN:**

- Destinataria
- Subdirección de Presupuestos, DIPRES
- División de Control de Gestión Pública, DIPRES
- Unidad Ley de Acceso a la Información Pública, DIPRES
- Oficina de Partes, DIPRES

**REFRENDACIÓN:**

- Subdirector de Presupuestos
- Jefe División de Control de Gestión Pública
- Encargado de Acceso a la Información Pública
- Abogada de Acceso a la Información Pública



2014-11-24 14:00:00

## Annex 3

[MIS DATOS](#) | [POLITICA DE PRIVACIDAD](#) | [AYUDA](#)

Bienvenido Don(a): Graciela Saavedra (Asignador) | Salir

---

Buscar Folio o Rut:

\* Solo busca folios, nombres, apellidos, correos electrónicos y contenido de la solicitud

[Alertas de solicitudes](#)

**Asignar Solicitudes**

**Gestión de Solicitudes 11**

**Ingreso de Solicitudes**

**Rectificar Solicitudes**

**Solicitudes Respondidas**

**Solicitudes Finalizadas**

**Listado de Usuarios**

**Mis Datos**

**Ayuda**

**Inicio » Gestión de Solicitudes »**

**Datos de la solicitud** Plazo: Quedan 20 días

|                                  |                                  |         |    |
|----------------------------------|----------------------------------|---------|----|
| Folio                            | AE002W-0000405 (Web)             |         |    |
| Fecha ingreso                    | 03-11-2014                       |         |    |
| Fecha estimada de término        | 01-12-2014                       |         |    |
| Responsable                      | Hector Marcelo Basaletti Salinas |         |    |
| Dirigida a                       | Ministerio de Hacienda           |         |    |
| Entidad                          | Dirección de Presupuestos        |         |    |
| Etapas actual                    | En Proceso                       |         |    |
| Estado actual                    | Asignada                         |         |    |
| Notificar por correo electrónico | No                               |         |    |
| Forma de recepción               | Correo electrónico               |         |    |
| Formato de entrega               | Formato electrónico o digital    | Firmada | no |

**Datos Personales de "Blanca Lopez "** [Mostrar datos del solicitante](#)

|                    |                   |              |
|--------------------|-------------------|--------------|
| Correo electrónico | bl84@sussex.ac.uk | Apoderado    |
| Dirección          | Número            | Departamento |
| País               | Región            |              |
| Comuna             | Ciudad            |              |

**Información Solicitada**

Buenas tardes, le agradezco mucho a la unidad su respuesta a mi solicitud anterior. En dicha respuesta se menciona que existe una evaluación comprehensiva del gasto del programa "Red de urgencia", pero no la encuentro ni en el sistema DIPRES, ni en la página del MINSAL, me pueden hacer el favor de informarme como localizarla?

Por último, la ev. comprehensiva del gasto es la misma que la ev. del gasto institucional?

Muchas gracias nuevamente, saludos cordiales.

Blanca López

**Documentos Adjuntos** [Agregar Archivo](#)

La solicitud no contiene archivos

Puede ingresar la ubicación web de documentos asociados a esta solicitud

Url documento 1

Url documento 2

Seleccione etapa En Proceso

Seleccione el estado Asignada

**Observaciones**


Annex 3

Guardar cambios

Prorrogar Solicitud

Resumen días de trámite

| Modificado por               | Perfil    | Estado     | Fecha      | Días de trámite |
|------------------------------|-----------|------------|------------|-----------------|
| Graciela Saavedra Valderrama | Asignador | En Proceso | 03-11-2014 | Menos de un día |



Modificado por:  
Graciela Saavedra Valderrama  
Asignador  
Estado : En Proceso  
Fecha: 03-11-2014  
Observación  
Asignación de responsable: Hector Marcelo Basaletti



Annex 3



ORD. N° 1606 /

**ANT.** Solicitud de acceso a la información pública N° AE002W-394, de fecha 30 de septiembre de 2014.

**MAT.** Da respuesta que indica.

**SANTIAGO, 28 OCT. 2014**

**DE : DIRECTOR DE PRESUPUESTOS**

**A : BLANCA LOPEZ**  
**bl84@sussex.ac.uk**

En el marco de la Ley N° 20.285 sobre Acceso a la Información Pública, esta Dirección recibió su solicitud de acceso identificada en el antecedente, referida a obtener:

"Le solicito de la manera más atenta, me pueda proporcionar por favor el número exacto y nombres de todos los programas del sector salud. Lo anterior, debido a que en la ley de presupuesto y la página de DIPRES no lo clarifica. Además quiero saber como consultar tanto las evaluaciones realizadas a los programas de este sector (salud) del periodo 2007 a 2014 y el seguimiento a las recomendaciones, porque tengo problemas para abrir los formatos pdf que vienen en la página de DIPRES..." [sic]

Al respecto, cúpleme informar a Ud. lo siguiente:

1. En relación a aquella parte de su solicitud que se identifica más adelante, ha sido remitida al Ministerio de Salud, organismo que a nuestro juicio, es competente para ocuparse de la misma, conforme lo establece el artículo N°13 de la Ley 20.285 Sobre Acceso a la Información Pública:

"Le solicito de la manera más atenta, me pueda proporcionar por favor el número exacto y nombres de todos los programas del sector salud. Lo anterior, debido a que en la ley de presupuesto y la página de DIPRES no lo clarifica." [sic]

No obstante lo anterior, en el sitio web de esta Dirección: <http://www.dipres.gob.cl/594/w3-propertyvalue-2129.html>, se encuentra a disposición del público y en formato digital, los programas del sector salud, mediante la Partida 16 del Ministerio de Salud de la Ley de Presupuestos del Sector Público de cada año.

2. En el sitio web de esta Dirección de Presupuestos: <http://www.dipres.gob.cl/595/w3-propertyname-562.html>, sección "Control de Gestión", "Ver todos los informes de Control de Gestión", se encuentra a disposición del público y en formato digital, los informes de Evaluación de Programas Gubernamentales y de Evaluación de Impactos de Programas, por año y ministerio.

Annex 3

Sin perjuicio de lo señalado precedentemente, se acompaña en anexo aparte al presente oficio, cuadros con los Programas Evaluados del sector salud periodo 2007-2014, y los Compromisos Asociados a los Programas Evaluados. Dichos anexos especifican, entre otra información, el sitio web que aloja el respectivo reporte para cada caso informado.

Se acompaña en anexo aparte al presente documento, copia del oficio remitido al órgano antes señalado.

Saluda atentamente a Ud.,



PWA/AAM

Incluye: lo señalado

**DISTRIBUCIÓN:**

- Destinataria
- Subdirección de Presupuestos, DIPRES
- División de Control de Gestión Pública, DIPRES
- Unidad Ley de Acceso a la Información Pública, DIPRES
- Oficina de Partes, DIPRES

**REFRENDACIÓN:**

- Subdirector de Presupuestos
- Jefa División de Control de Gestión Pública
- Encargado Ley de Acceso a la Información Pública
- Abogada Ley de Acceso a la Información Pública



## Annex 3

1º tabla: Listado de Programas Evaluados 2007-2014 (Sector Salud)

| Año  | Nombre Programa  | Tipo de Evaluación* | Servicio  | Observaciones   |
|------|--|---------------------|---|---|
| 2007 | No hay evaluaciones asignadas para el año 2007   |                     |   |   |
| 2008 | No hay evaluaciones asignadas para el año 2008   |                     |   |   |
| 2008 | Comisión de Medicina Preventiva e Invalidez (COMPIN)                                     | EGI                 | Subsecretaría de Salud Pública                          | <a href="http://www.dipres.gob.cl/595/articulos-60545_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-60545_doc_pdf.pdf</a>   |
| 2008 | Programa de Tratamiento de la Obesidad   | EPG                 | Fondo Nacional De Salud                                 | <a href="http://www.dipres.gob.cl/595/articulos-38653_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-38653_doc_pdf.pdf</a>   |
| 2009 | No hay evaluaciones asignadas para el año 2009   |                     |   |   |
| 2010 | No hay evaluaciones asignadas para el año 2010   |                     |   |   |
| 2010 | Central de Abastecimiento del Sistema Nacional de Salud (CENABAST)                       | EGI                 | Central de Abastecimiento del Sistema Nacional de Salud | <a href="http://www.dipres.gob.cl/595/articulos-89715_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-89715_doc_pdf.pdf</a>   |
| 2010 | Establecimientos Experimentales  | EGI                 | Subsecretaría de Redes Asistenciales                    | <a href="http://www.dipres.gob.cl/595/articulos-95316_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-95316_doc_pdf.pdf</a>   |
| 2010 | Programa de Prevención y Control del VIH/SIDA y Enfermedades de Transmisión Sexual (ETS) | EPG                 | Subsecretaría de Salud Pública                          | <a href="http://www.dipres.gob.cl/595/articulos-64136_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-64136_doc_pdf.pdf</a>   |
| 2011 | Bono Automático AUGE   | EPN                 | Fondo Nacional De Salud                                 | <a href="http://www.dipres.gob.cl/594/articulos-119344_doc_pdf.pdf">http://www.dipres.gob.cl/594/articulos-119344_doc_pdf.pdf</a> |
| 2011 | Fondo Nacional De Salud  | EGI                 | Fondo Nacional De Salud                                 | <a href="http://www.dipres.gob.cl/595/articulos-115534_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-115534_doc_pdf.pdf</a> |
| 2011 | Programa de Apoyo al Recién Nacido   | EPG                 | Subsecretaría de Redes Asistenciales                    | <a href="http://www.dipres.gob.cl/595/articulos-76585_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-76585_doc_pdf.pdf</a>   |
| 2012 | No hay evaluaciones asignadas para el año 2012   |                     |   |   |
| 2012 | No hay evaluaciones asignadas para el año 2012   |                     |   |   |
| 2012 | Programa Nacional  | EPG                 | Subsecretaría de Salud                                  | <a href="http://www.dipres.gob.cl/595/articulos-89691_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-89691_doc_pdf.pdf</a>   |

|      | Immunizaciones  |     | Pública                              |   |
|------|---|-----|--------------------------------------|---|
| 2013 |   | EL  |                                      | No hay evaluaciones asignadas para el año 2013  |
| 2013 | Red de Urgencia   | EGI | Subsecretaría de Redes Asistenciales | En proceso de licitación.   |
| 2013 | Compras a Privados: Programa Prestaciones Valoradas       | EPG | Fondo Nacional De Salud              | <a href="http://www.dipres.gob.cl/595/articulos-109112_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-109112_doc_pdf.pdf</a> |
| 2013 | Programa Nacional de Alimentación Complementaria**        | EPG | Subsecretaría de Salud Pública       | <a href="http://www.dipres.gob.cl/595/articulos-109111_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-109111_doc_pdf.pdf</a> |
| 2014 | Programa de Formación de Especialistas y Subespecialistas | EL  | Subsecretaría de Redes Asistenciales | No hay evaluaciones asignadas para el año 2014  |
| 2014 |   | EPG | Subsecretaría de Redes Asistenciales | <a href="http://www.dipres.gob.cl/595/articulos-122603_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-122603_doc_pdf.pdf</a> |

\*E,PG: Evaluación de Programas Gubernamentales; EL: Evaluaciones de Impacto; EGI: Evaluación de Gestión Institucional; EPN: Evaluación de Programas Nuevos.

\*\* Programa sin seguimiento de compromisos.

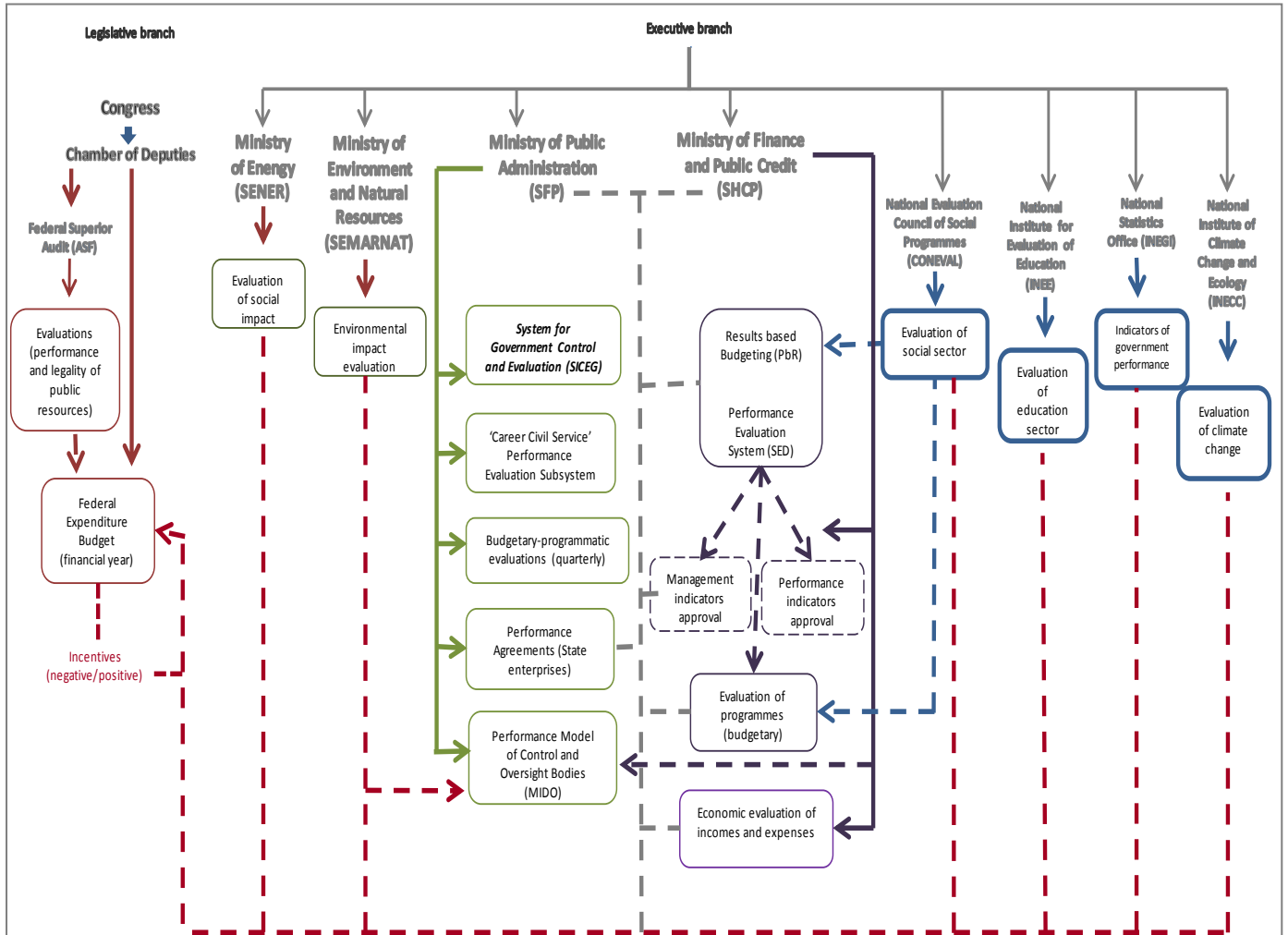
2º tabla: Compromisos Asociados a los Programas Evaluados

| Año  | Nombre Programa  | Tipo de Evaluación* | Servicio  | Observaciones   |
|------|--|---------------------|---|---|
| 2008 | Comisión de Medicina Preventiva e Invalidez (COMPIN)                                     | EGI                 | Subsecretaría de Salud Pública                          | <a href="http://www.dipres.gob.cl/595/articulos-92141_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-92141_doc_pdf.pdf</a>   |
| 2008 | Programa de Tratamiento de la Obesidad   | EPG                 | Fondo Nacional De Salud                                 | <a href="http://www.dipres.gob.cl/595/articulos-117341_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-117341_doc_pdf.pdf</a> |
| 2010 | Central de Abastecimiento del Sistema Nacional de Salud (CENABAST)                       | EGI                 | Central de Abastecimiento del Sistema Nacional de Salud | <a href="http://www.dipres.gob.cl/595/articulos-117425_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-117425_doc_pdf.pdf</a> |
| 2010 | Establecimientos Experimentales  | EGI                 | Subsecretaría de Redes Asistenciales                    | <a href="http://www.dipres.gob.cl/595/articulos-117426_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-117426_doc_pdf.pdf</a> |
| 2010 | Programa de Prevención y Control del VIH/SIDA y Enfermedades de Transmisión Sexual (ETS) | EPG                 | Subsecretaría de Salud Pública                          | <a href="http://www.dipres.gob.cl/595/articulos-92129_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-92129_doc_pdf.pdf</a>   |
| 2011 | Fondo Nacional De Salud  | EGI                 | Fondo Nacional De Salud                                 | Compromisos por vencer***   |
| 2011 | Programa de Apoyo al Recién Nacido   | EPG                 | Subsecretaría de Redes Asistenciales                    | <a href="http://www.dipres.gob.cl/595/articulos-117545_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-117545_doc_pdf.pdf</a> |
| 2012 | Programa Nacional Immunizaciones   | EPG                 | Subsecretaría de Salud Pública                          | <a href="http://www.dipres.gob.cl/595/articulos-117342_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-117342_doc_pdf.pdf</a> |
| 2013 | Compras a Privados: Programa Prestaciones Valoradas                                      | EPG                 | Fondo Nacional De Salud                                 | <a href="http://www.dipres.gob.cl/595/articulos-124659_doc_pdf.pdf">http://www.dipres.gob.cl/595/articulos-124659_doc_pdf.pdf</a> |
| 2014 | Programa de Formación de Especialistas y Subespecialistas                                | EPG                 | Subsecretaría de Redes Asistenciales                    | Evaluación recién terminada. Proceso de elaboración de compromisos aún no comienza.   |

\*\*\* Este programa tiene compromisos que vencen a partir del 30-06-2014, por lo que el reporte estará disponible en la web de DIPRES dentro del mes de Octubre de 2014.

Annex 4

Evaluation systems and monitoring practices in Mexico



Source: author's own elaboration.

Annex 5

## Healthcare programmes of Maule



Sr. Juan Carlos Díaz.

A través del presente doy respuesta a su solicitud, ingresada con el Folio N° A0047T0000008 (portal transparencia) de fecha 24 de Agosto del presente año. Respecto a su primera consulta **¿cuál es la cantidad y nombre exactos de los programas del sector salud?**, puedo señalar lo siguiente: los programas que ejecuta la Secretaría Regional Ministerial de Salud del Región del Maule, mandatados por el Ministerio de Salud son 31 programas, los cuales se dividen en 2 grandes áreas, con la finalidad de cumplir las funciones esenciales materias relativas a la promoción de la salud, vigilancia, prevención y control de enfermedades que afectan a poblaciones o grupos de personas de salud pública en la región del Maule.

I.- Salud Pública: donde se ejecutan los siguientes programas.

1. Vigilancia epidemiológica transmisibles
2. Reglamento Sanitario Internacional (RSI)
3. Vigilancia Epidemiológica no transmisible (registro poblacional de cáncer)
4. Programa Nacional de Inmunización (PNI)
5. Programa de la Infancia
6. Programa de Salud de la Mujer y del Recién Nacido
7. Programa Salud Integral del Adolescente y Jóvenes
8. Programa Adulto Mayor
9. Programa de Tuberculosis
10. Programa Cardiovascular
11. Programa Salud Bucal
12. Programa Salud Mental
13. Programa Salud Sexual y Prevención de VIH/SIDA e ITS
14. Programa Prevención cáncer cervicouterino y mamas
15. Programa Nacional de Alimentación Complementaria (PNAC – PACAM)
16. Programa Prevención Consumo de Tabaco
17. Programa Promoción de Estilos de Vida Saludables (elije vivir sano)

II.- Salud Ambiental: donde se ejecutan los siguientes programas.

18. Inocuidad de los alimentos
19. Agua para consumo humano
20. Aguas recreacionales
21. Entornos saludables
22. Aguas servidas
23. Residuos sólidos domiciliarios
24. Residuos sólidos industriales no peligrosos
25. Residuos sólidos industriales peligrosos
26. Residuos de Establecimientos de Atención de Salud (REAS)
27. Residuos industriales líquidos (RILES)
28. Salud laboral
29. Zoonosis y Control de vectores
30. Programa Calidad del aire (control de fuentes de emisiones fijas)
31. Programa de profesiones Médicas y Políticas Farmacéuticas

Respecto a la segunda pregunta **sobre las evaluaciones realizadas a los programas de este sector (salud) del periodo 2007 a 2014 y el seguimiento a las recomendaciones?**

Los programas que ejecuta la Secretaría Regional Ministerial de Salud del Región del Maule, mandatados por el Ministerio de Salud, son evaluados a través del cumplimiento de METAS de desempeño colectivo, Programas de Mejoramiento de la Gestión (PMG) y compromisos de Hacienda (H), las cuales son asignadas y monitoreadas desde el nivel central. Lo cual puede ser consultado al Departamento de Control de Gestión del Ministerio de salud.

Respecto a la información adicional, respecto al sistema de consulta a través de la página del MINSAL, enviaremos desde la Oficina de Información y Reclamos su inquietud.

Agradeciendo de ante mano su requerimiento.

Atte.



VALERIA ORTIZ VEGA  
SECRETARIA REGIONAL MINISTERIAL DE SALUD  
-REGIÓN DEL MAULE

Annex 6

## Healthcare programmes of Osorno



### REPORTE DE PROGRAMAS AREA SANITARIA SERVICIO SALUD OSORNO

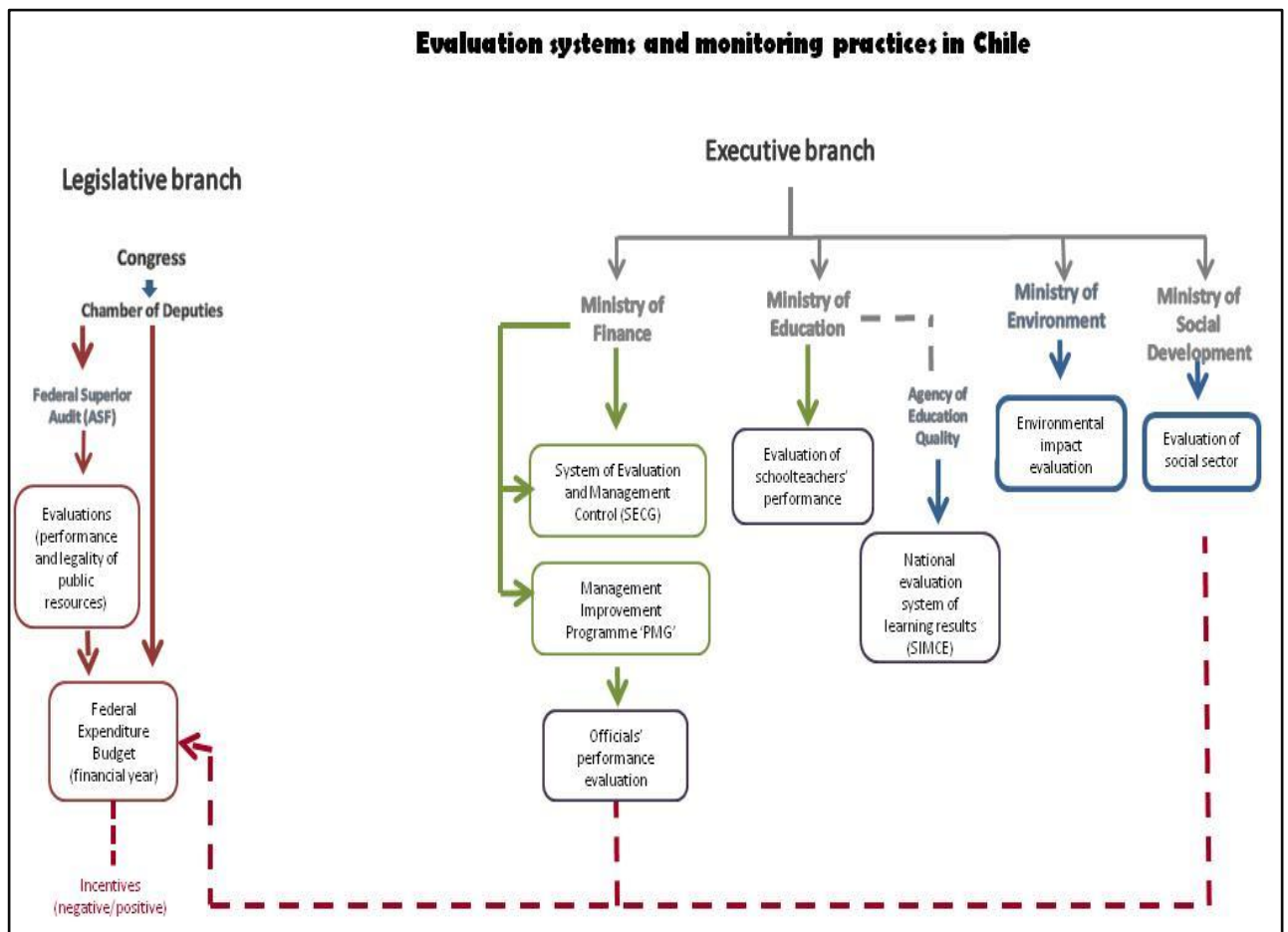
|  |   |                         |
|--|---|-------------------------|
| Nombre del programa  | Programa Nacional de Salud de la Infancia (PNSI)  |                         |
| Estado del programa  | En curso  |                         |
| Unidad responsable del programa  | Subdepartamento Atención Sanitaria y Art. de la Red / Unidad Ciclo Vital de las Personas. |                         |
| Identificación Presupuestaria  | Partida   | 16                      |
|  | Capítulo  | 38                      |
|  | Programa  | Piloto de Salud Escolar |
|  | Subtítulo   | 24                      |
| Breve descripción del programa   |   |                         |
| <p>El Programa Nacional de Salud de la infancia tiene como propósito contribuir al desarrollo integral de niños y niñas menores de 10 años, en su contexto familiar y comunitario, a través de actividades de fomento, protección, prevención, recuperación de la salud y rehabilitación, que impulsen la plena expresión de su potencial biopsicosocial y mejor calidad de vida.</p> <p>En el contexto del PNSI, el año 2013 se incorporó el Servicio de Salud Osorno al Piloto de Salud Escolar, cuyo propósito es Otorgar una atención integral y resolutive a la población de 5 a 9 años a través de controles de salud realizados en el establecimiento educacional, que contribuyan a la promoción de factores protectores de la salud y el desarrollo y a la pesquisa oportuna de problemas de salud prevalentes en este grupo de edad.</p> <p>Para el desarrollo de este último, se distribuyeron recursos por Resolución N°407 del 15 de febrero de 2013, del Minsal, que correspondieron a \$31.599.000 por subtítulo 24; los que se asignaron a los Deptos. de Salud Osorno, Purranque y Río Negro, para abarcar la atención de 2.000 niños(as) de 5 a 9 años en etapa escolar.</p> <p>Al 31 de diciembre de 2013, los recursos se encuentran transferidos a las comunas; y al 10 de enero de 2014, los Deptos. de Salud deben entregar antecedentes para evaluar cumplimiento.</p> |   |                         |
| Cobertura / Población objetivo   |   |                         |
| Niños y niñas de 0 a 9 años  |   |                         |
| Año de inicio del programa   | Año de término del programa   |                         |
|  | Continuidad   |                         |

Annex 6

|  |   |   |
|--|---|---|
| Nombre del programa:   | <b>Salud de la Mujer</b>  |   |
| Estado del programa  | En curso  | En preparación                              |
|  | X   |   |
| Unidad responsable del programa  | Subdepartamento Atención Sanitaria y Art. de la Red / Unidad Ciclo Vital de las Personas. |   |
| Identificación Presupuestaria  | Partida   | 16  |
|  | Capítulo  | 38  |
|  | Programa  | Apoyo a la gestión local en AP: Rescate PAP |
|  | Subtítulo   | 24  |
| Breve descripción del programa:  |   |   |
| <p>Se firmó convenio con el Municipio de Osorno, que se comprometió a utilizar los recursos entregados (\$584.000) para financiar los siguientes objetivos:</p> <p>1.-Rescate de PAP positivos: Cáncer Cérvico Uterino.</p> <p>2.-Difusión para incentivar toma de examen de PAP.</p> <p>De acuerdo a lo informado con estos fondos se realizaron:</p> <p>26 extensiones horarias de 3 hrs. 1 de 2 hrs. con un total de 80 hrs. y un total de 226 PAP tomados.</p> <p>Además el Programa de Cáncer cervicouterino, evaluó el grado de cumplimiento con fecha corte 31 de Diciembre, informe que se remitió a todos los Establecimientos de la red.</p> <p>La comuna de Osorno, con una población de 38.331 mujeres de 25-64 años, obtuvo una cobertura de 53,6%, excluyendo los PAP del extrasistema. Cabe destacar que el año 2012, con una población de 37.668, la comuna de Osorno obtuvo una cobertura de 54%, por lo que se hace necesario continuar con este apoyo financiero, el presente año.</p> <p>OTRAS COMUNAS:</p> <p>Comuna San Pablo: 62,2</p> <p>Comuna Entre Lagos: 67</p> <p>Comuna Puerto Octay : 51,8</p> <p>Comuna Purranque: 66,8</p> <p>Comuna Río Negro: 62,5</p> <p>Comuna San Juan de la Costa: 73,7</p> <p>Siendo el promedio del Servicio de Salud 61,1%.</p> <p>Se continuará trabajando con otras estrategias para aumentar la toma de PAP y servirá para ello la implementación del REVICAN en el presente año.</p> |   |   |
| Cobertura / Población objetivo   |   |   |
| 53,6/ 38331 (comuna Osorno solamente)  |   |   |
| Año de inicio del programa   | Año de término del programa   |   |
|  | Continuidad   |   |



Annex 7



Source: author's own elaboration.

Annex 8

Categories of recommendations across countries

| AGENCY/<br>PROGRAMME   | TYPE OF<br>EVALUATION                     | RECOMMENDATI<br>ON IMPROVED  | RECOMMENDATI<br>ON NOT<br>IMPROVED | CATEGORIES OF RECOMMENDATIONS |                                |   |  |                                   |            |  |  |
|--|---|--|------------------------------------|-------------------------------|--------------------------------|---|--|-----------------------------------|------------|--|--|
|  |   |  |                                    | DESIGN OF<br>THE<br>PROGRAMME | PROCESS OF<br>THE<br>PROGRAMME | ACTIVITIES<br>LINKED TO<br>THE STAFF<br>AND<br>PROGRAMME<br>OPERATORS | TO DEVELOP<br>A SPECIFIC<br>EVALUATION<br>OR<br>ANALYSIS | COORDINATI<br>ON AND<br>SYNERGIES | IT SYSTEMS | PROGRAMME<br>'S<br>SATISFACTIO<br>N SURVEY | TO<br>PUBLICIZE,<br>PROMOTION<br>AND MEDIA |
| MEXICAN HEALTH SECTOR  |   |  |                                    |                               |                                |   |  |                                   |            |  |  |
| Status Quo category  |   |  |                                    |                               |                                |   |  |                                   |            |  |  |
| ISSSTE/EMERGE<br>NCY<br>HEALTHCARE   | EVALUATION<br>SPECIFIC OF<br>PERFORMANCE  | To include indicators<br>(MIR) of impact,<br>services provided,<br>management and<br>quality of service                                      |                                    | X                             |                                |   |  |                                   |            |  |  |
|  | COST<br>EFFECTIVENESS                     | To modify the goals<br>according to the<br>reality of the<br>programme   |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | Implementation of a<br>pilot to standardize<br>the triage model in the<br>emergency areas  |                                    |                               | X                              |   |  |                                   |            |  |  |
|  |   | To include more<br>indicators and modify<br>the main one   |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | To develop a design<br>or impact evaluation  |                                    |                               |                                |   | X  |                                   |            |  |  |
| SSA/HEALTH<br>SERVICES<br>PROVIDED<br>ACROSS<br>DIFFERENT<br>MEDICAL SITES | SPECIFIC OF<br>PERFORMANCE<br>(EED)       | Implementing an IT<br>system for monitoring<br>the services provided<br>and results per unit<br>and at local level                           |                                    |                               |                                |   |  |                                   | X          |  |  |
|  |   | To optimize the<br>outpatient services<br>according to the<br>capacity of the<br>hospital  |                                    |                               | X                              |   |  |                                   |            |  |  |
|  |   | To finish the regional<br>hospitals under<br>construction  |                                    |                               | X                              |   |  |                                   |            |  |  |
|  |   | To implement specific<br>indicators of quality<br>and programme<br>satisfaction per unit<br>and services<br>provided                         |                                    | X                             |                                |   |  |                                   |            |  |  |
|  | DESIGN                                    | To amend the problem<br>that the programme<br>would tackle   |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | The evaluator<br>suggested to<br>consider a proposal<br>for the problem that<br>the programme would<br>tackle                                |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | To improve indicators<br>(MIR)   |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | To make coherent the<br>design of the MIR  |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | To quantify the<br>potential of the<br>recipients of<br>programmes   |                                    | X                             |                                |   |  |                                   |            |  |  |
| SSA/PROGRAM<br>ME FOR<br>VULNERABLE<br>PEOPLE AND<br>FAMILIES              | OF<br>CONSISTENCY<br>AND RESULTS<br>(ECR) | To develop national<br>and international<br>research to show the<br>correlation between<br>the programme and<br>services provided            |                                    |                               |                                |   | X  |                                   |            |  |  |
|  |   | To improve indicators<br>(MIR)   |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | To modify the problem<br>that the programme<br>would tackle  |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | To quantify the<br>programme's<br>recipients   |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | Implementation of a<br>costs system  |                                    |                               | X                              |   |  |                                   |            |  |  |
|  |   | The development of<br>strategic plans at the<br>short, medium and<br>long term   |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | An analysis of the<br>synergies between<br>the programme and<br>others   |                                    |                               |                                |   |  | X                                 |            |  |  |
|  |   | To update of the<br>diagnostic which<br>stresses the problem<br>that the programme<br>would tackle   |                                    | X                             |                                |   |  |                                   |            |  |  |
|  | SPECIFIC OF<br>PERFORMANCE<br>(EED)       | To quantify the<br>potential of<br>programme's<br>recipients   |                                    | X                             |                                |   |  |                                   |            |  |  |
|  |   | To develop an ECR  |                                    |                               |                                |   | X  |                                   |            |  |  |
|  |   | To develop an<br>evaluation of impact  |                                    |                               |                                |   | X  |                                   |            |  |  |
|  |   | Subprogrammes,<br>strategies and<br>activities needed to<br>be reorganized for<br>better coordination  |                                    |                               |                                |   |  | X                                 |            |  |  |
|  |   | To include the SSA<br>and SNDIF into the<br>decision-making<br>process to group the<br>subprogrammes for<br>better institutional<br>planning |                                    |                               | X                              |   |  |                                   |            |  |  |
|  |   | To coordinate the<br>strategic planning<br>actions at the local<br>level   |                                    |                               |                                |   |  | X                                 |            |  |  |
|  |   | To improve<br>coordination between<br>the child defence<br>programme and the<br>Ministry of Social<br>Development<br>(SEDESOL)               |                                    |                               |                                |   |  | X                                 |            |  |  |

| Amended programmes   |                               |  |  |   |   |  |   |   |  |  |   |
|--|-------------------------------|--|--|---|---|--|---|---|--|--|---|
| IMSS/PROGRAMME OF PROVISION OF HEALTHCARE SERVICES           | SPECIFIC OF PERFORMANCE (EED) | To include performance indicators (MIR) per medical unit   |  | X |   |  |   |   |  |  |   |
|  |                               | Implementation of preventive actions targeting teenagers, male and elderly people  |  |   | X |  |   |   |  |  |   |
|  |                               | Inclusion of indicators to measure such preventive actions   |  | X |   |  |   |   |  |  |   |
|  |                               | Measurement of the rate of mortality of breast cancer  |  |   | X |  |   |   |  |  |   |
|  |                               | Further increase the goal established for the detection of breast cancer in women by 25%                                     |  | X |   |  |   |   |  |  |   |
| SSA/PREVENTION OF ADDICTIONS                                 | DESIGN                        | The development of operating rules (Reglas de Operación ROP) and guidelines for the programme                                |  | X |   |  |   |   |  |  |   |
|  |                               | The improvement of indicators (MIR)  |  | X |   |  |   |   |  |  |   |
|  |                               | The implementation of a data system to exchange information with programmes with complementary actions                       |  |   |   |  |   | X |  |  |   |
|  |                               | Programme operators should consider the three main problems (tobacco, alcohol and drugs) to strengthen the services provided |  |   | X |  |   |   |  |  |   |
|  |                               | Including gender strategies and recipients living in rural areas and immigrants  |  |   | X |  |   |   |  |  |   |
| SSA-SNDIF/PROTECTION AND INTEGRATED DEVELOPMENT OF CHILDHOOD | SPECIFIC OF PERFORMANCE (EED) | A diagnostic of the programme, which serves to validate its operation  |  | X |   |  |   |   |  |  |   |
|  |                               | The development of an impact evaluation  |  |   |   |  | X |   |  |  |   |
|  |                               | The quantification of the programme recipients   |  | X |   |  |   |   |  |  |   |
|  |                               | The establishment of a standardized process to select the programme recipients   |  |   | X |  |   |   |  |  |   |
|  |                               | To publicize the programme and its results   |  |   |   |  |   |   |  |  | X |
|  |                               | The development of a strategic plan for the short, medium and long term  |  | X |   |  |   |   |  |  |   |
|  |                               | To identify the services provided to the programme recipients  |  | X |   |  |   |   |  |  |   |
|  |                               | To use the data census to quantify the programme recipients  |  | X |   |  |   |   |  |  |   |
|  |                               | To develop a diagnostic per each state of the problem that the programme would tackle  |  |   | X |  |   |   |  |  |   |
|  |                               | The quantification of the programme recipients based on the measurements of poverty  |  | X |   |  |   |   |  |  |   |
|  |                               | To strengthen synergies with the ministry of social development to coordinate efforts to tackle the problem                  |  |   |   |  |   | X |  |  |   |
|  |                               | The homogenization of the operation of the mechanism of social protection at local level                                     |  |   |   |  |   | X |  |  |   |
| SSA/PREVENTION AGAINST OBESITY                               | DESIGN                        | The improvement of indicators (MIR)  |  | X |   |  |   |   |  |  |   |
|  |                               | The inclusion of those indicators related to the gender perspective  |  | X |   |  |   |   |  |  |   |
|  |                               | The development of regulation, guidelines, operating rules and monitoring system to operate the programme                    |  | X |   |  |   |   |  |  |   |

| Cancelled programmes                                     |                                  |   |   |    |    |   |   |   |   |   |   |
|--|----------------------------------|---|---|----|----|---|---|---|---|---|---|
| SSA/NATIONAL CRUSADE FOR THE QUALITY OF HEALTH PROVISION | DESIGN                           | NO DATA   | NO DATA   |    |    |   |   |   |   |   |   |
|  | OF CONSISTENCY AND RESULTS (ECR) | NO DATA   | NO DATA   |    |    |   |   |   |   |   |   |
| SSA/EQUITABLE FUNDING OF HEALTH PROVISION                | SPECIFIC OF PERFORMANCE (EED)    | To homogenize the problem that the programme would tackle in the diverse documents issued by the agency |   | X  |    |   |   |   |   |   |   |
|  |                                  | To improve indicators (MIR)   |   | X  |    |   |   |   |   |   |   |
|  |                                  | To improve coordination and responsibilities between the different units operating the programme        |   |    |    |   |   | X |   |   |   |
|  |                                  | To implement a data system including impact indicators and results of the programme                     |   |    |    |   |   |   | X |   |   |
|  |                                  | To develop strategic plans at the short, medium and long term   |   | X  |    |   |   |   |   |   |   |
|  |                                  | To establish a procedure of accountability and transparency of the programme                            |   |    | X  |   |   |   |   |   |   |
| IMSS/ EFFICIENT CURATIVE CARE                            | COST EFFECTIVENESS               |   | To develop an evaluation of consistency and results   |    |    |   | X |   |   |   |   |
|  |                                  |   | To improve the data system of the programme to timely results   |    |    |   |   |   | X |   |   |
|  |                                  | Training of medical units in the methodology of logical framework                                       |   |    |    | X |   |   |   |   |   |
|  |                                  |   | The medical units should be faster to open the cost centers   |    | X  |   |   |   |   |   |   |
|  |                                  |   | The medical units should coordinate the catalogue of the cost centers with other departments or agencies                  |    |    |   |   | X |   |   |   |
|  |                                  |   | To develop the regulatory framework of the cost centers   |    | X  |   |   |   |   |   |   |
|  |                                  |   | To include all the elements needed to develop the cost centers study  |    |    |   | X |   |   |   |   |
|  |                                  |   | To upgrade the cost of the dietetic and nutrition center  |    | X  |   |   |   |   |   |   |
|  |                                  |   | To refine and update the accountable and budgetary categories   |    | X  |   |   |   |   |   |   |
|  |                                  |   | To consolidate the data system of results and effects of the medical interventions  |    |    |   |   |   | X |   |   |
|  |                                  |   | To develop indicators related to quality of the life of the recipients  | X  |    |   |   |   |   |   |   |
|  |                                  |   | To use the model of territorial distribution of the effects for the purpose of developing indicators                      |    | X  |   |   |   |   |   |   |
|  |                                  |   | To link the indicators of hospital mortality to matters such as demography, diseases and disorders                        | X  |    |   |   |   |   |   |   |
|  |                                  |   | The policy operators should consider the behavioural patterns of cost-effectiveness made by other departments and centers |    | X  |   |   |   |   |   |   |
|  |                                  |   | To define the universe of diseases to be monitored based on the cost-effectiveness approach                               | X  |    |   |   |   |   |   |   |
|  |                                  |   | To generalise the actions to all health centers to improve the quality in services  |    | X  |   |   |   |   |   |   |
|  | PROCESS                          | To develop indicators linked to chronic degenerative diseases   |   | X  |    |   |   |   |   |   |   |
|  |                                  |   | To strengthen the connection of the programming, budgeting and provision of services                                      |    | X  |   |   |   |   |   |   |
|  |                                  |   | To supervise the process of check-up and medical examination  |    | X  |   |   |   |   |   |   |
|  |                                  |   | To improve the IT system  |    |    |   |   |   | X |   |   |
|  |                                  |   | To develop studies of medical management, audits and provision of sanitary services                                       |    |    |   | X |   |   |   |   |
|  |                                  |   | To revise the correct allocation of economic, human and material resources to provide quality in the health services      |    | X  |   |   |   |   |   |   |
| TOTAL MEXICO   |                                  | 39  | 43  | 36 | 22 | 1 | 8 | 8 | 6 | 0 | 1 |

| CHILEAN HEALTH SECTOR  |                                     |  |  |   |   |   |   |   |   |  |
|--|-------------------------------------|--|--|---|---|---|---|---|---|--|
| Status Quo category  |                                     |  |  |   |   |   |   |   |   |  |
| UNDERSECRETARY OF THE HEALTHCARE NETWORK/EXPERIMENTAL HEALTH CENTRES | COMPREHENSIVE SPENDING REVIEW (ECG) | The consideration of a specific model and methodology by other agencies to be compared with the programme  |  |   | X |   |   |   |   |  |
|  |                                     | Implementing a costs' system to improve budgetary control, financial management and programming activities |  |   | X |   |   |   |   |  |
|  |                                     |  | Restructuring medical sites, homogenising with other sites in the health sector  |   | X |   |   |   |   |  |
|  |                                     |  | Updating the remuneration system   |   | X |   |   |   |   |  |
|  |                                     |  | Training of executives   |   |   | X |   |   |   |  |
|  |                                     |  | Transferring of the programme funds timely   |   | X |   |   |   |   |  |
|  |                                     |  | Transparency of the transferring process   |   | X |   |   |   |   |  |
|  |                                     |  | Considering a specific model (CRSCO and Hospital L. Tisné) to be duplicated in other medical sites                       |   | X |   |   |   |   |  |
|  |                                     |  | Improving coordination between agencies to enhance the implementation of health policies                                 |   |   |   |   | X |   |  |
|  |                                     |  | Focusing the strategic planning on specific issues   | X |   |   |   |   |   |  |
|  |                                     |  | To delegate to a professional of the department of Management Control the responsibility of the institutional management |   |   | X |   |   |   |  |
|  |                                     |  | The improvement of the health centres  |   | X |   |   |   |   |  |
|  |                                     |  | To design and use specific indicators (MIR)  | X |   |   |   |   |   |  |
|  |                                     |  | To define the scope of management of the health centres  |   | X |   |   |   |   |  |
|  |                                     |  | Implementing a model for recipients and the financial model of the programme   |   | X |   |   |   |   |  |
|  |                                     |  | To include the strategic objectives to all health centres  | X |   |   |   |   |   |  |
|  |                                     |  | To implement Responsibility Centres  |   | X |   |   |   |   |  |
|  |                                     |  | To implement the model GRD   |   | X |   |   |   |   |  |
|  |                                     |  | To improve IT systems  |   |   |   |   | X |   |  |
|  |                                     |  | To revise which model is more efficient to provide emergency services  |   | X |   |   |   |   |  |
|  |                                     |  | To implement performance indicators of human resources   | X |   |   |   |   |   |  |
|  |                                     |  | To strengthen the evaluation mechanism   |   |   |   | X |   |   |  |
|  |                                     |  | To evaluate the participation of the health centres with the academia  |   |   |   | X |   |   |  |
|  |                                     |  | To implement a programme satisfaction survey   |   |   |   |   |   | X |  |
|  |                                     |  | To transfer the programme satisfaction survey to other services  |   |   |   |   |   | X |  |

|  |                                       |  |   |   |   |   |   |   |   |   |   |
|--|---------------------------------------|--|---|---|---|---|---|---|---|---|---|
| FONASA/PURCHASE FROM PRIVATE SECTOR                                      | EVALUATION OF PUBLIC PROGRAMMES (EPG) | The inclusion of evaluations from other medical units  |   |   |   |   | X |   |   |   |   |
|  |                                       |  | Developing a forecast of the demand for the medical services  |   |   |   | X |   |   |   |   |
|  |                                       |  | Strengthening audit activities and a study of costs   |   | X |   |   |   |   |   |   |
|  |                                       |  | The improvement of indicators (MIR)   | X |   |   |   |   |   |   |   |
|  |                                       |  | Implementing a monitoring system  |   |   |   |   |   | X |   |   |
|  |                                       |  | Implementing a monitoring system (similar to previous)  |   |   |   |   |   | X |   |   |
|  |                                       |  | Developing a comparative analysis of costs of the provision of services between the public and private sector                               |   |   |   | X |   |   |   |   |
|  |                                       |  | Restructuring the process of public bidding and procurement   |   | X |   |   |   |   |   |   |
|  |                                       |  | Outlining of the set of deficiencies of the programme   |   | X |   |   |   |   |   |   |
|  |                                       |  | Strengthening the programme   |   | X |   |   |   |   |   |   |
|  |                                       |  | Developing some mechanisms of evaluation of internal management of the programme  |   |   |   | X |   |   |   |   |
| UNDERSECRETARY OF PUBLIC HEALTH/NATIONAL PROGRAMME OF SUPPLEMENTARY FOOD | EVALUATION OF PUBLIC PROGRAMMES (EPG) |  | Reformulating the programme's aim according to the activities really performed and the inclusion of health services for people with obesity | X |   |   |   |   |   |   |   |
|  |                                       |  | Establishing interagency coordination and synergies with other programmes   |   |   |   |   | X |   |   |   |
|  |                                       |  | To develop a cost-benefit and feasibility analyses  |   |   |   | X |   |   |   |   |
|  |                                       |  | To develop IT systems and strengthen of the human resources   |   |   |   |   |   | X |   |   |
|  |                                       |  | To develop a monitoring system of indicators (MIR)  |   |   |   |   |   | X |   |   |
|  |                                       |  | To develop a market research of products for better public bidding and procurement  |   |   |   | X |   |   |   |   |
|  |                                       |  | To monitor of programme's satisfaction  |   |   |   |   |   |   | X |   |
|  |                                       |  | To develop a market research of dairy products considering international public bidding   |   |   |   | X |   |   |   |   |
|  |                                       |  | Making a benchmarking studies of the impact of poor nutrition   |   |   |   | X |   |   |   |   |
|  |                                       |  | Forecasting the demand of health services to improve the budgetary process of the programme   |   |   |   | X |   |   |   |   |
| UNDERSECRETARY OF PUBLIC HEALTH/NATIONAL PROGRAMME OF IMMUNIZATIONS      | EVALUATION OF PUBLIC PROGRAMMES (EPG) | Establishing a system for monitoring vaccine wastage and analysis of the overpurchase process                                    |   |   |   |   |   |   | X |   |   |
|  |                                       | Strengthening and promoting of the IT system   |   |   |   |   |   |   | X |   |   |
|  |                                       | Improving the budget system and financial management including a control mechanism of the public bidding process and procurement |   |   | X |   |   |   |   |   |   |
|  |                                       |  | To be aware of the demographic and epidemiological changes affecting the programme  |   | X |   |   |   |   |   |   |
|  |                                       |  | To establish an annual programme of staff training  |   |   | X |   |   |   |   |   |
|  |                                       |  | To publicize of the programme per year  |   |   |   |   |   |   |   | X |
|  |                                       |  | To improve coordination between areas inside the agency   |   |   |   |   | X |   |   |   |
|  |                                       |  | To develop a study to qualitatively measure the operation of the recipients in terms of vaccination   |   |   |   | X |   |   |   |   |
|  |                                       |  | To quantify the programme recipients  | X |   |   |   |   |   |   |   |

| Amended programmes  |  |   |  |   |   |   |   |   |   |  |  |
|---|--|---|--|---|---|---|---|---|---|--|--|
| FONASA,<br>UNDERSECRETARY OF PUBLIC<br>HEALTH AND<br>OF<br>HEALTHCARE<br>NETWORKS/OF<br>TREATMENT OF<br>OBESITY | EVALUATION OF<br>PUBLIC<br>PROGRAMMES<br>(EPG) | The improvement of<br>interagency<br>coordination with<br>other sector<br>programmes with<br>similar objectives   |  |   |   |   |   | X |   |  |  |
|   |  | The implementation of<br>a study of cost-<br>effectiveness for the<br>medical and surgical<br>treatment   |  |   |   | X |   |   |   |  |  |
|   |  | The improvement of<br>indicators (MIR)  |  | X |   |   |   |   |   |  |  |
|   |  | The improvement of<br>techniques to<br>measure the<br>programme   |  |   | X |   |   |   |   |  |  |
|   |  | Implementing a<br>monitoring system   |  |   |   |   |   |   | X |  |  |
|   |  | Expanding the<br>number of recipients   |  |   | X |   |   |   |   |  |  |
| UNDERSECRETARY OF PUBLIC<br>HEALTH/COMMISSION OF<br>PREVENTIVE<br>MEDICINE AND<br>DISABILITY                    | COMPREHENSIVE<br>SPENDING<br>REVIEW (ECG)      | The development of a<br>feasibility and<br>economic impact<br>analyses  |  |   |   |   | X |   |   |  |  |
|   |  | The implementation of<br>an IT system for<br>processing health<br>services online,<br>which includes the<br>registration of time<br>employed in activities<br>and processes |  |   |   |   |   |   | X |  |  |
|   |  | Increasing the hours<br>of specialists doing<br>expert's report of<br>medical leaves  |  |   | X |   |   |   |   |  |  |
|   |  | Establishing a<br>guideline to define<br>days to rest up per<br>pathology   |  |   | X |   |   |   |   |  |  |
|   |  | Identifying tendency<br>of doctors who<br>issued a high number<br>of medical leaves   |  |   | X |   |   |   |   |  |  |
|   |  | Increasing audits for<br>checking compliance<br>of days to rest up<br>through home visits   |  |   | X |   |   |   |   |  |  |
|   |  | Increasing the<br>number of months for<br>calculating the income<br>to pay medical leaves   |  |   | X |   |   |   |   |  |  |
|   |  | Developing a risk<br>indicator of medical<br>leaves   |  | X |   |   |   |   |   |  |  |
|   |  | Developing a study<br>measuring time to<br>process medical<br>leaves  |  |   |   |   | X |   |   |  |  |
|   |  | Implementing a<br>manual of procedures<br>of all health services<br>of the commission   |  |   | X |   |   |   |   |  |  |
|   |  | Establishing an audit<br>programme of the<br>provision of the sub-<br>products  |  |   | X |   |   |   |   |  |  |
|   |  | The implementation of<br>the human resources<br>policy established in<br>the modernization<br>plan  |  |   |   | X |   |   |   |  |  |
|   |  | The completion of the<br>transferring process<br>of personnel to other<br>agencies  |  |   |   | X |   |   |   |  |  |
|   |  | The development of a<br>study of the staff<br>size  |  |   |   | X |   |   |   |  |  |
|   |  | The process to<br>transfer funds should<br>be timely  |  |   | X |   |   |   |   |  |  |
|   |  | The distribution of<br>economic resources<br>should be per region<br>and disaggregated<br>per item  |  |   | X |   |   |   |   |  |  |
|   |  | To move offices to<br>new facilities  |  |   | X |   |   |   |   |  |  |
|   |  | Implementing a<br>management model in<br>two regions  |  |   | X |   |   |   |   |  |  |
|   |  | The development of a<br>technical feasibility<br>study of transferring<br>some health services<br>to other agencies   |  |   |   |   | X |   |   |  |  |
|   |  | The implementation of<br>the electronic medical<br>license  |  |   | X |   |   |   |   |  |  |
|   |  | The department<br>(COAF) should not<br>receive incomplete<br>forms of medical<br>leaves   |  |   | X |   |   |   |   |  |  |

|  |                                       |   |  |   |   |   |   |   |   |  |  |
|--|---------------------------------------|---|--|---|---|---|---|---|---|--|--|
| UNDERSECRETARY OF PUBLIC HEALTH/OF PREVENTION AND CONTROL OF HIV | EVALUATION OF PUBLIC PROGRAMMES (EPG) | The establishment of a unit to coordinate all sector agencies                                 |  |   |   |   |   | X |   |  |  |
|  |                                       | To strengthen preventive actions to vulnerable recipients                                     |  |   | X |   |   |   |   |  |  |
|  |                                       | To implement a multisectoral collaboration in the preventive strategy                         |  |   |   |   |   | X |   |  |  |
|  |                                       | To develop a mechanism to strengthen citizenry participation                                  |  |   | X |   |   |   |   |  |  |
|  |                                       | To assure that the preventive actions have enough economic resources to meet their goals      |  |   | X |   |   |   |   |  |  |
|  |                                       | To find a funding mechanism to have a stock of condoms to vulnerable recipients               |  |   | X |   |   |   |   |  |  |
|  |                                       | To proceed with the preventive test of HIV  |  |   | X |   |   |   |   |  |  |
|  |                                       | To strengthen an integrated healthcare model of HIV   |  |   | X |   |   |   |   |  |  |
|  |                                       | To continue developing specific studies used for evaluations and the redesign of strategies   |  |   |   |   | X |   |   |  |  |
|  |                                       | To decentralize the decisions related to therapies and control of stock                       |  |   | X |   |   |   |   |  |  |
|  |                                       | To improve the mechanism of purchase and pay of medicines and other medical consumables       |  |   | X |   |   |   |   |  |  |
|  |                                       | To link the IT system to register statistics and health data                                  |  |   |   |   |   |   | X |  |  |
| NATIONAL HEALTH FUND (FONASA)/NATIONAL HEALTH FUND               | COMPREHENSIVE SPENDING REVIEW (ECG)   | The improvement of indicators (MIR) of six different types (one recommendation per indicator) |  | X |   |   |   |   |   |  |  |
|  |                                       | The improvement of indicators (MIR) of six different types (one recommendation per indicator) |  | X |   |   |   |   |   |  |  |
|  |                                       | The improvement of indicators (MIR) of six different types (one recommendation per indicator) |  | X |   |   |   |   |   |  |  |
|  |                                       | The improvement of indicators (MIR) of six different types (one recommendation per indicator) |  | X |   |   |   |   |   |  |  |
|  |                                       | The improvement of indicators (MIR) of six different types (one recommendation per indicator) |  | X |   |   |   |   |   |  |  |
|  |                                       | The improvement of indicators (MIR) of six different types (one recommendation per indicator) |  | X |   |   |   |   |   |  |  |
|  |                                       | The development of an analysis of the strategic agency goals                                  |  |   |   |   | X |   |   |  |  |
|  |                                       | The development of a profile of personnel required  |  |   |   | X |   |   |   |  |  |
|  |                                       | The development of mechanisms to transfer financial resources                                 |  |   | X |   |   |   |   |  |  |
|  |                                       | The availability of data to the center desk   |  |   | X |   |   |   |   |  |  |
|  |                                       | The list of recipients  |  | X |   |   |   |   |   |  |  |
|  |                                       | The development of a fiscal system  |  |   | X |   |   |   |   |  |  |
|  |                                       | The improvement of workload planification   |  |   | X |   |   |   |   |  |  |
|  |                                       | The development of a diagnostic of the agency tasks   |  |   |   |   | X |   |   |  |  |
|  |                                       | The development of a comprehensive IT healthcare system                                       |  |   |   |   |   |   | X |  |  |
|  |                                       | The development of an audit system  |  |   | X |   |   |   |   |  |  |
|  |                                       | The monitoring of waiting time and patient queues in medical services                         |  |   | X |   |   |   |   |  |  |
|  |                                       | The development of a mechanism to control the programme costs                                 |  |   | X |   |   |   |   |  |  |



|  |                                       |   |   |    |    |    |    |    |    |   |   |
|--|---------------------------------------|---|---|----|----|----|----|----|----|---|---|
| CENABAST/NATIONAL SUPPLY CENTER                              | COMPREHENSIVE SPENDING REVIEW (ECG)   | To keep running two main activities, which represent the core business of the programme   |   |    | X  |    |    |    |    |   |   |
|  |                                       | To include the job position of contract management at the same level of importance than the other positions   |   |    |    | X  |    |    |    |   |   |
|  |                                       | To review the contracts with customers and suppliers  |   |    | X  |    |    |    |    |   |   |
|  |                                       | To strengthen the programme of fiscal discipline  |   |    | X  |    |    |    |    |   |   |
|  |                                       | To calculate the fee charged by the center  |   |    | X  |    |    |    |    |   |   |
|  |                                       |   | To redefine the vision of the programme   | X  |    |    |    |    |    |   |   |
|  |                                       |   | To develop the activity of logistic support   |    | X  |    |    |    |    |   |   |
|  |                                       |   | To consider the proposal of logistic support made by the evaluator                                      |    | X  |    |    |    |    |   |   |
|  |                                       |   | To improve the activities to storage medicines  |    | X  |    |    |    |    |   |   |
|  |                                       |   | To pay off debts with suppliers   |    | X  |    |    |    |    |   |   |
| <b>Cancelled programmes</b>                                  |                                       |   |   |    |    |    |    |    |    |   |   |
| UNDERSECRETARY OF HEALTHCARE NETWORKS/NEWBORN HEALTH SUPPORT | EVALUATION OF PUBLIC PROGRAMMES (EPG) | Transferring of all health services of this programme to the programme of biopsychosocial support   |   |    | X  |    |    |    |    |   |   |
|  |                                       | Developing an assessment by the Ministry of Health and Ministry of Cooperation and Planning regarding the conditions of the Chilean childhood to strengthen its development               |   |    |    |    | X  |    |    |   |   |
|  |                                       | Transferring roles and functions to the programme of biopsychosocial support  |   |    |    | X  |    |    |    |   |   |
|  |                                       | Including in the current hospital services some logistic improvements, the control of storage of medicines (quantity and quality) and the improvement of the activity of logistic support |   |    | X  |    |    |    |    |   |   |
|  |                                       | Developing a satisfaction survey once the new database of programme recipients has been defined   |   |    |    |    |    |    |    | X |   |
|  |                                       |   | Redefining the quantification of the programme recipients based on new criteria to set those recipients | X  |    |    |    |    |    |   |   |
| TOTAL CHILE  |                                       | 57  | 72  | 18 | 59 | 9  | 20 | 6  | 12 | 4 | 1 |
| TOTAL  |                                       | 96  | 115   | 54 | 81 | 10 | 28 | 14 | 18 | 4 | 2 |

Source: author's own elaboration.

Table 1

### Planning, control, monitoring and evaluation practices in Mexico

| Government Institution/System/Regulation   | Year of publication                 | Objectives/Goals   | Year of conclusion | Arguments/Reasons for disappearance  |
|--|-------------------------------------|--|--------------------|--|
| Efficiency Commissions [Comisiones de Eficiencia]  | 1926                                | To organize public services to streamline government agencies.   | No data available  | SPP, 1988  |
| Planning of the Republic Law [Ley de Planeación General de la República]   | 1930                                | To establish the National Planning Office as an advisor. To coordinate government institutions' activities regarding geographic, economic and social characteristics.  | No data available  | SPP, 1988. Oropeza, 2005: 85   |
| Internal Commission of the 'Partido Revolucionario'  | 1933                                | Integrated by government officials and party members, which were responsible for the 1st. Sexennial Plan.  | 1940               | This was not an economic development plan. The Plan and Programme Commission were dissolved in the next administration.  |
| Plan 1934-1940   | 1933                                | To establish a first planning plan by the National Revolutionary Party.  | 1940               |  |
| Advisor Committee [Comité Asesor]  | 1935                                | To coordinate national planning policies, it was created by the President Lázaro Cardenas  | 1940               |  |
| Plan 1940-1946   | 1940                                | This is the last military government (post-revolutionary). In those years emerged the first efforts to assess public resources.  | 1946               | SPP, 1988  |
| Development Fund [Fondo para el Industrial]  | 1941                                | To support industrial sector and the 'Bank of Mexico' was responsible for it.  | 1944               | This lacked of funds to promote enterprises and disappeared shortly afterwards due to the Industrial Commission.   |
| Commission of Economic Planning [Comisión Federal de Planificación Económica]  | 1942                                | Agency of the Ministry of National Economy.  | No data available  | SPP, 1988  |
| Industrial Development Commission [Comisión Federal de Fomento Industrial]   | 1944                                | For planning industrial development in the country and promoting of new industries. It was created to strength activities coordinated by the Industrial Fund.  | 1947               | Transformations driven under this Commission was hampered by two factors: technical staff and public funding. It was replaced by the Investment Commission.  |
| Office of National Property and Administrative Inspection [Secretaría de Nacionales e Inspección Administrativa]                         | 1946                                | To coordinate management performance of government institutions.   | No data available  | SPP, 1988  |
| Investment National Commission [Comisión de Inversiones]   | 1947                                | To coordinate public investment, it was attempt by President Miguel Aleman to restore Industrial Commission.   | 1952               | In 1954, the Commission change its name to 'Comisión de Inversiones' under responsibility of the President. In the next administration, the law diminished their activities until it vanished. SPP, 1988<br>This law was replaced by the Federal Law of State enterprises in 1986. |
| Law of the Decentralized Agencies and Enterprises [Ley para el Control de las Empresas Descentralizadas y Empresas de Propiedad Estatal] | 1947                                | To monitor results, which should be consistent with purposes they were created.  | 1952               |  |
| Investment Committee [Comité de Inversiones]   | 1953                                | To organize investment in public and private sector, under responsibility of the Ministry of Finance and the Ministry of Economy. It was created by President Adolfo Ruiz.   | 1954               | Public sector disregarded for the Committee, and disagreements between institutions responsible. It was replaced by the Investment Commission.   |
| Programme of Public Investment [Programa Nacional de Inversión Pública]  | 1953-1958                           | To establish goals for economic growth utilised in the 'national accounts'.  | 1958               | Oropeza, 2005: 85  |
| Investment Commission [Comisión de Inversiones]  | 1954                                | To plan public expenditure, to assign priority for investment projects, and to establish a programme of public investment in the medium-term. It was created to substitute previous Investment Committee by President Adolfo Ruiz. | 1958               | In spite of being one of the most successful initiatives, it diminished control when its structure was moved to the Ministry of Presidency. INAP, 1972: 19.  |
| Ministry of Presidency [Secretaría de la Presidencia]  | 1958                                | To establish general plan of public expenditure and Executive branch investments, to plan and control of public investment of decentralized agencies and state enterprises. It was created by the President Adolfo Lopez.          | 1976               | This ministry disappeared due to the establishment in 1976 of the Ministry of Programming and Budget SPP. INAP, 1972: 47. Lanz, 1987.  |
| Office of National Heritage [Secretaría de Patrimonio Nacional SEPANAL]  | 1960 (in the latter of this decade) | To coordinate investment planning and current spending undertaken by the Sub-commission for Investment-Finance, integrated by general directors of the Ministry of Treasury, Presidency and SEPANAL.                               | 1970               | At the beginning of the 1970's Treasury controlled fiscal and monetary policy in Mexico in a largely conventional, market-oriented manner, but it confronted centers of contention in other agencies. Bailey: 1984: 79   |

|  |           |  |                   |   |
|--|-----------|--|-------------------|---|
| ncy Commission to formulate Development Plans [Comisión Interministerial para la Formulación de planes nacionales de Desarrollo]                       | 1962      | To establish national plan for economic and social development in the short and medium-terms, to evaluate investment projects in terms of quality and importance, under responsibility of the Ministry of Finance and the Ministry of Presidency. It was created by President Adolfo Lopez and had similarities with the Investment Commission (1954). | No data available | It was diminished its activities until vanished in the next administration due to disagreements between institutions, as well as lacking of experienced technical staff to operate it. However, according to the INAP there is no evidence of its disappearance.<br>INAP, 1972: 22. |
| the Action Plan [Plan de Acción]   | 1962-1964 | This plan allocate investment programmes within the mid-term of economic and social policy.  | 1964              | Oropeza, 2005: 85   |
| Planning Law (draft project) [Ley de Planeación]   | 1963      | The Comisión de Planeación [Planning Commission] was responsibility of Congress (Senate) to draft the Federal Planning Law. The law was formulated to coordinate and unify different planning criteria of government institutions. Also consider mid and long term planning for national economic development.   | 1963              | Did not thrive. This law initiative did not achieve consensus and lack of internal organization of activities.<br>Oropeza, 2005: 86. SPP, 1988.   |
| Administration Commission [Comisión de Administración Pública]   | 1965      | The Ministry of Presidency was responsible of this Commission. Its aim is for reorganization of federal government according to main national sector activities. To determine whether should be the 'adequate' structure of public administration, including regulation and administrative reforms.  | 1974              | In 1971 emerged the General Office of Administrative Studies in the same Ministry of Presidency to replace this Commission.<br>Lanz, 1987. Ministry of Presidency, 1967: 21.  |
| Plan of Economic and Social Development [Plan Nacional de Desarrollo Económico y Social]   | 1966-1970 | To establish guidelines for public investment, long term development and public and private planning.  | 1970              | Oropeza, 2005: 85-86. SPP, 1988   |
| Agreement of Federal Investment [Acuerdo de Programación de Inversión Federales por Entidad Federativa]  | 1973      | Investment programme for every State to assess social impact.  | No data available | SPP, 1988.  |
| Law of Economic and Social Development (project) [Ley General de Desarrollo Económico y Social]  | 1975      | To establish planning basis of federal public administration (APF) and coordination within the state/local governments and private/social sectors.   | 1975              | Did not thrive.<br>SPP, 1988.   |
| Administrative Reform Programme [Programa de Reforma Administrativa]   | 1976-1982 | To link state enterprises activities with their sector, which have the same objectives. The administrative reform focused on the reorganization of state enterprises and ministries, as well as on programme budgeting.  | 1982              | Bailey, 1980: 11. SPP, 1988. Lanz, 1987.  |
| Planning System [Sistema de Planeación Nacional]   | 1976      | This is the first effort to build a mechanism for planning, programming-budgeting, evaluation and control government actions as a whole based on four modalities: obligatory, coordination, concerted and inductive.   | 1983              | SPP, 1988.<br>This was replaced by the National Democratic Planning System in 1983.   |
| Office of Programming and Budget [Oficina de Programación y Presupuesto]   | 1976      | Responsible for planning, programming, budgeting, monitoring and evaluation of public expenditure.   | 1992              | Its disappearance was due to conflicts with the SHCP. On one hand, the expenditure operation, on the other hand, the income. Most of the main tasks were transferred to the SHCP.<br>SPP, 1988.   |
| Budget, Accounting and Public Administration [Ley de Presupuesto, Contabilidad y Gasto Público]  | 1976      | To improve public spending control oriented based in planning goals and priorities.  |                   | SPP, 1988.<br><b>This law was replaced by</b> The Federal Law of Budget and Fiscal Responsibility.  |
| Government Organic Law [Ley de la Administración Pública]  | 1976      | To organize federal public administration and state enterprises.   |                   | Currently exists.   |
| Office of Presidency [Assessment and Evaluation Office]  | 1977      | For conducting analysis progress of government programs and national assessment sessions with responsible.   | 1983              | This office faced problems with the Undersecretary of Evaluation (SPP), therefore, its efforts were slowly fading. It vanished when the SECOGEF emerged in 1983.  |
| Monitoring and Control Financial System of Public Expenditure (COSEFF) [Sistema de Control y Seguimiento Físico y Económico del Gasto Público Federal] | 1976      | This system belong to the SPP to ensure further control, monitoring, analysis and evaluation of public spending, to identify deviations between scheduled and executed programmes and projects.  | 1992              | This system disappears along with the SPP in 1992.<br>Carrillo and García, 1981.  |
| Internal Expenditure Budget [Presupuesto de Egresos de la Federación]  | 1976      | To perform corrective, relevant and timely actions.  |                   | Currently exists  |
| Intergovernmental Agreement [Convenios de Coordinación]  | 1977      | For strengthen federalism and decentralization, the relationship between federal and local level (planning, programming and budgeting).  |                   | PROPMAP, Official journal, 28/05/1996.<br>Galvez, 1985.   |

|  |           |  |      |   |
|--|-----------|--|------|---|
| ment Agreement [Convenios de Coordinación]   | 1977      | For strengthen federalism and decentralization, the relationship between federal and local level (planning, programming and budgeting). Afterwards, this 'Agreements' were often utilized by the SECOGEF to support the System for Government Control and Evaluation (SICEG) at local level (Galvez, 1985).  |      | PROPMAP, Official journal, 28/05/1996.<br>Galvez, 1985.<br>Currently are the Social Development Agreements.   |
| retary of Evaluation (General of Analysis for Evaluation) [Secretaría de Evaluación] | 1977      | These units were created in the SPP, to design, execute and coordinate national integral evaluation system, to design evaluation regulations, to analyse public expenditure and the achievement of budgets and programmes.   | 1983 | At the end of 1977, evaluation functions were translated from the SPP to the Ministry of Presidency, however, structure remained in the former.<br>According to Internal organization of the SPP (Official Journal, 25/01/1983), this Undersecretary disappeared in 1983, and only remained the General Office of Documentation and Evaluation in charge of the Undersecretary of Development Planning.<br>Bailey, 1980: 27. Moctezuma, 1989: 88. |
| velopment Plan [Plan Global de]  | 1980-1982 | To link sector programmes within global vision of problems and continuing the establishment of the national evaluation system.   | 1982 | SPP, 1988.  |
| State Development Committees [Comités de Planeación para el Estado]                  | 1981      | To promote and assist in the formulation, implementation and evaluation of State development plans, aimed by state enterprises. They operate in substitution of the COPRODE which operated between 1980-1981   |      | Currently exists.<br>SPP, 1988.   |
| of the General Controller of the [Secretaría de la Contraloría de la Federación]     | 1982      | This ministry regulates monitoring and evaluation activities, as well as responsible for auditing, monitoring and evaluation.  | 1994 | Official Journal, 29/12/1982.<br>The evaluation results process was so far as it was planned, and more correlated with the auditing process. In 1994, became into the Ministry of the Controller and Administrative Development (SECODAM). Afterwards the Ministry of Public Administration (SFP) until its disappearance in 2013.  |
| em for Government Control and [Sistema de Control y Evaluación Nacional SICEG].      | 1982      | It was established under the supervision of the Ministry of General Controller of the Federation 'SECOGEF' (LOAPF, article 32 BIS), operated as a preventative system on the basis of Commissioners and Delegates:<br>'...Its aim is to provide elements of judgement for control and monitoring of objectives and policies achieved, as well as to promote permanent the improvement of public management'.<br>Since the reforms in 2009 (Official journal, 17/06/2009), it was under the control of the Evaluation Management and Government Performance Unit in the SFP (before SECOGEF). | 2013 | Official journal, 29/12/1982. Galvez, 1985.<br>The system disappeared along with the Ministry of Public Administration (Official journal, 02/01/2013) and its functions will be distributed amongst the SHCP, and the Presidency Office.  |
| Law [Ley de Planeación]  | 1983      | Art.3 ...will coordinate actions and evaluate results.<br>Art.9 The Federal Executive shall establish an Evaluation and Compensation System for Performance to measure the units of the central federal administration...<br>Art.13 makes it mandatory activities leading to plans and programs formulation, implementation, monitoring and evaluation.  |      | Currently exists.<br>Official journal, 05/01/1983.  |
| Constitution [Constitución Política de los Estados Unidos Mexicanos]                 | 1983      | To give State authority for conducting and promoting economic, political and social development, also planning activities, and to establish laws of national planning of economic and social development (articles 25, 26, 73)   |      | Currently exists.<br>Official journal, 03/02/1983.  |
| Democratic Planning System [Sistema Nacional de Planeación Política SNPD]            | 1983      | Constitutional reform (art. 26) of 03/02/1983. The system establishes a referendum mechanism to collect societal participation in the formulation, monitoring and evaluation stages. Regarded as mandatory planning in the public sphere, the intergovernmental consensus and getting involve with the private and social sectors.   |      | Currently exists.<br>Official journal, 03/02/1983.  |
| Development Plan [Plan Nacional de Desarrollo PND]                                   | 1983-1988 | This plan is mandatory by Planning Law, to determine national objectives, strategies and priorities, estimation of resources allocated, instruments and responsible of   |      | Currently exists and changes its objectives every six-year presidential term.   |

|   |           |   |      |  |
|---|-----------|---|------|--|
| Law of State enterprises [Ley de Entidades Paraestatales]   | 1986      | This law replaced the Control Law of the Decentralized Agencies and State Enterprises. To give management autonomy to state enterprises in exchange for performance assessment.   |      | Currently exists.<br>Official journal, 14/05/1986.   |
| General Programme for the Modernization of the Federal Public Administration [Programa General de Modernización de la Administración Pública] | 1989      | Its main objectives were: i) Consolidate an administrative culture which encourages efficiency and productivity in the management of government institutions; ii) Promotion the quality and transparency of the “evaluation” in the public services provided; iii) Prevention and fighting corruption.  | 1995 | Official journal, 9/02/1989.<br>In 1995, the federal government started a reform to change the planning-programming-budgeting comprehensive system. Indeed three new programmes were put into operation: the New Programmatic Structure, the Performance Evaluation System and the Public Administration Modernization Programme. This reasons forced to concluding the programme. |
| Programme for Modernization of State Enterprises [Programa Nacional para la Modernización de la Empresa Pública]                              | 1990-1994 | To restructure of state enterprises, to strengthen management autonomy, as well as their efficiency, effectiveness and productivity; to improve incentives and the performance agreements; and to develop an evaluation system to boost efficiency.   | 1994 | Official journal, 16/04/1990.<br>The results were poor due to the overlapped regulations, such as budgetary and performance agreements.  |
| Control Programme [Programa de Control Social]  | 1991      | This is a tool for control and evaluation to tackle corruption. At regional and local level, activities to strengthen the SICEG were focused on the mechanism of Social Controller in the National Programme of Solidarity and on the effectiveness of resources.   |      | Currently exists.<br>SFP, 1991.  |
| Ministry of the Controllorship and Administrative Development [Secretaría de Administración y Desarrollo Administrativo]                      | 1994      | This ministry has the same functions of the SECOTEF but reinforce the administrative capacities of public administration.   | 2003 | Official Journal 28/12/1994.<br>In 2003 year, this ministry became into the Ministry of Public Administration (SFP). Official Journal, 10/04/2003.   |
| Programmatic Structure [Nueva Estructura Programática NEP]  | 1998      | This is one of the two elements of the National Programme of Development Financing ‘PRONAFIDE’ 1997-2000. It is a system of strategic indicators (performance) for monitoring and evaluation of programmes and main government activities, but it also plans to organize them – at federal, state and local level - on the basis of budgetary programmes to classify public resources into categories ‘estructura programática’   | 2000 | Decree of the PRONAFIDE. Official Journal 7/11/1997.<br>According to the CEFP (2008: 33), the SIMEP interrupted the SED’s consolidation, including the NEP.  |
| Performance Evaluation System [Sistema de Evaluación del Desempeño SED]   | 1998      | The SHCP and the SECODAM sought the implementation of an “integral data system”, which later become the SED, an essential mechanism inside the strategic planning phase for modernization. The SED was a system to transform the NEP into a performance measurement model using a results-based approach.<br>Its objectives were designed to strengthen governance, achieve results and citizens’ satisfaction, link agencies and state enterprises’ performance to institutional objectives, and to increase the effectiveness and efficiency of public resources. The mission of the system was to ‘contribute to raising the productivity of fiscal resources... for better decision-making in terms of the planning and execution of expenditure, to strengthen the programme evaluation process and responsible performance of bureaucracy’ (Mejía, 2005: 18). |      | Currently exists with some adjustments.<br>Decree of the PRONAFIDE. Official Journal 7/11/1997.<br>Mejía, 2005: 18.<br>According to the CEFP (2008: 33), the SIMEP interrupted the SED’s consolidation, including the NEP.   |
| Programme for Modernization of the Federal Public Administration [Programa de Modernización de la Administración Pública]                     | 1995-2000 | The SECODAM (before SECOTEF) established this mid-term programme, which comprises four programmes and comprehensible institutional objectives: i) Transform the Federal Public Administration into an organization that acts efficiently and effectively through a new culture of service: ii) Tackle corruption and impunity through control mechanisms and effective sanctions.   | 2000 | Decree of the PROMAP, Official Journal 28/05/1996.<br>The UDA (Unit responsible of the PROMAP), and its plan for administrative modernization of the Mexican bureaucracy, also had to face same difficulties which led to the programme producing few results. Cejudo (2008: 116-7)  |
| Superior Audit of the Federation [ASF]  | 2000      | The Supreme Audit Law of the Federation established the creation of the ASF to replace the Treasury Accounting created since 1824, to confer authority to the Legislative   |      | Currently exists.<br>Guerrero, 2002: 176.<br>Official Journal, 29/12/2000.   |

|   |      |  |      |  |
|---|------|--|------|--|
| Superior de la Federación<br>Audit of the Federation ASF]   | 2000 | <p>The Supreme Audit Law of the Federation established the creation of the ASF to replace the Treasury Accounting created since 1824, to confer authority to the Legislative branch.</p> <p>Its mission statement is 'To assess goals, objectives set in federal programmes on the basis of strategic indicators approved in the budget to verify their performance and legality of public resources'.</p> <p>It was created for external auditing of federal public resources in the three Government branches, autonomous constitutional organs, states and local governments, as well as individuals when they receive such resources.</p> <p>Even more, has no authority for intervention in projects, programmes and on-going expenditure, only for the Public Account (annually) elaborated by the Ministry of Finance and Public Credit (SHCP).</p> <p>Its regulation is the Law Enforcement and Accountability of the Federation [Ley de Fiscalización y Rendición de Cuentas de la Federación] (Official Journal 18/06/2010), and one of their objectives is related to evaluation:</p> <p>'To assess goals, objectives and performance achieved by federal programmes and utilisation of public resources on the basis of strategic indicators approved in the budget, National Development Plan, sector and regional plans, annual programmes, among others'.</p> |      | Currently exists.<br>Guerrero, 2002: 176.<br>Official Journal, 29/12/2000.   |
| Government Agenda [Agenda de Gobierno ABG]  | 2001 | The agenda operates through: Government that costs less, Quality Government, Professional Government, Digital Government, Deregulated Government, and Honest and Transparent Government. It was part of the Strategic Model for Government Innovation and Quality [Modelo Estratégico para la Innovación y Calidad Gubernamental], which was looking for radical transformation from multiple perspectives of the federal public administration.   | 2006 | This agenda disappeared in 2006. The main reason for their disappearance was the overlap between the SFP (the lead department) and the Presidency Office. Another reason was goals being modified to reflect the extent to which they could be achieved.<br>Pardo, 2007: 899.  |
| Strategic Goals System [Sistema de Metas Presidenciales SIMEP]  | 2002 | It is a top management evaluation system which identifies high impact processes and projects linked to the PND and its sector programmes. Its main characteristic was the agreement signed between the head of the ministry or state enterprise and the president, which established goals and deadlines, under supervision of the Unit of Control and Evaluation of Public Management in the SFP.   | 2006 | Accountability Report 2000-2006: 157. Diener, 2007: 85. Medina, 2007: 50-1. Mejía, 2005: 21-2. Pardo, 2007: 899.<br><br>This system disappeared in 2006. The main reason for their disappearance was the overlap between the SFP (the lead department) and the Presidency Office. Another reason was goals being modified to reflect the extent to which they could be achieved. |
| National Institute for Evaluation of Education [Instituto Nacional para la Evaluación de la Educación INEE] | 2002 | It is a public, decentralized agency with financial, technical, and operational autonomy for the assessment of the education sector. Its roles are linked to the creation and development of educational indicators, statistics, tools, research and knowledge, as well as the measurement and evaluation of education issues.   |      | Currently exists.<br>Official journal, 8/08/2002.  |
| Planning and Compensation System for Performance [Sistema de Evaluación y Compensación por el Desempeño]    | 2003 | The Planning Law in its article 9 pointed out:<br>'The Executive shall establish an Evaluation and Compensation System for Performance to measure progress of ministries of the Federal Public Administration on achieving goals and objectives of the Plan - PND - and its sector programmes...'  |      | Currently exists.<br>Official journal, 10/04/2003.   |
| Ministry of Public Administration [Ministerio de la Función Pública SFP]                                    | 2003 | This ministry used to be the SECODAF, afterwards the SECODAM until this year. The reform to the organic law (LOAPF) transformed the SECODAM into the SFP.<br><br>The aim was to include more functions such as the operation of the career civil service of the Federal Public Administration, for planning, approval human resources structure and establishing regulations as well.  | 2013 | Official journal, 10/04/2003.<br><br>The SFP disappeared (Official journal, 02/01/2013) and its functions will be distributed amongst the SHCP, the new National Anti-Corruption Commission and the Presidency Office.   |
| System of performance evaluation of public officials.   | 2003 | This is a subsystem of the Professional Career Civil Service Law [Ley del Servicio Profesional de Carrera en la Administración Pública Federal] was issued by the SFP (Unit of Human Resources Policy in the Federal Public Administration).   |      | Currently exists.<br>Official journal, 10/04/2003.   |

|  |      |   |  |   |
|--|------|---|--|---|
| Integral Performance Model of Control Oversight Bodies [Modelo Integral de Vigilancia y Control de Órganos de Vigilancia y Control MICO] | 2004 | It was implemented by the SFP, on the basis of five indexes. Its aim is to ensure the OICs, Delegates and Commissioners help to reduce risks of corruption and opacity with a preventive approach to control, proper public management oversight, and effective corrective measures.  |  | SFP, 2004. Franco, 2006: 69. Official journal, 7/10/2004.<br>In the end, the MICO could not handle these issues and in 2009 (Official journal, 8/05/2009) it disappeared. However, a few months later, the SFP attempted to re-establish the system under a new supervision team, but it had a minor impact compared to previous years. |
| National Evaluation Council of the Development Policy [Consejo de Evaluación de la Política de Desarrollo Social CONEVAL]                | 2005 | To coordinate policy evaluation of this sector and to measure poverty, setting up criteria and methodologies for both purposes.   |  | Currently exists.<br>Official journal, 24/08/2005.  |
| Performance-based Budgeting 'PbR'  | 2006 | The Federal Law of Budget and Fiscal Responsibility [Ley Federal de Presupuesto y Responsabilidad Hacendaria LFPRH] establishes the strategy known as PbR, seeks to link governmental planning, design, and implementation of public programmes with the budgetary process through evaluation and the Performance Evaluation System 'SED'. It has two elements, the first is the quality of public expenditure through the Medium-Term Programme (PMP), which eventually (2008) became the Management Improvement Programme [Programa de Mejora de la Gestión 'PMG'] under the leadership of the SFP. A second element is the Annual Evaluation Programme (PAE), jointly issued by the SHCP, SFP and CONEVAL. |  | Currently exists.<br>Official journal, 30/03/2006.  |
| National Statistics Office [Instituto de Estadística y Geografía INEGI]  | 2006 | One of the subsystems of The National System of Statistics and Geographical Information is the government, public safety and justice system operated until 2009 according to the agency, which produces data regarding public administration.   |  | Currently exists.<br>Official journal, 7/04/2006.   |

Source: author's own elaboration.

Table 3

## The effectiveness of policy evaluation in Mexico

| Public organism                          | Key budgetary programmes | Budgetary programmes   | Year in the federal budget               | Indicator of budgetary process | Type of evaluation performed   | Amendments = A; Cancellation = C; Learning process = L; Status quo = S; Allocation of financial resources (Increase = I; Decrease = D; Unchanged = U) |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 | Recommendations of evaluations |  |   |
|--|--------------------------|--|--|--------------------------------|--|---|-----------------|------|----------------|-----------------|------|-----------------|-----------------|-------|-----------------|-----------------|------|----------------|-----------------|------|----------------|-----------------|-----------------|-----------------|-----------------|--------------------------------|--|---|
|  |                          |  |  |                                |  | Project budget  | Budget approved | 2007 | Project budget | Budget approved | 2008 | Project budget  | Budget approved | 2009  | Project budget  | Budget approved | 2010 | Project budget | Budget approved | 2011 | Project budget | Budget approved | 2012            | Project budget  | Budget approved |                                | 2013   | 2014  |
| IMSS, ISSFAM, ISSSTE, SSA, SEDENA, PEMEX | A001                     | Otras actividades  | 2007                                     | X (ISSFAM)                     |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
|  | A009                     | Servicios de salud al personal militar y sus derechohabientes  |  |                                |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
|  | A012                     | Atención médica  |  |                                |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
|  | A013                     | Otorgamiento de prestaciones socioeconómicas   |  |                                |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
| SSA                                      | B001                     | Construcción de centros de especialidad para pacientes ambulatorios y de unidades de atención hospitalaria |  |                                |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
| IMSS                                     | E001                     | Atención a la salud pública y salud reproductiva   |  |                                |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
| IMSS                                     | E001                     | Atención a la salud pública*   | 2009, 2010, 2011, 2012, 2013, 2014       | X                              | Ev. of design 2009, Ev. of gender 2009, Specific ev of performance (EED) 2009, EED 2010, M&E card 2014         |   |                 |      |                |                 |      | 2,467,957,962   | 2,467,957,962   | A/S/U | 2,543,807,036   | 2,543,807,036   | A/U  |                |                 |      |                |                 |                 |                 |                 |                                | The INFOMEX report says data is in the SHCP. There is no data of Ev. of design 2009. Diagnostic of gender 2008 and Ev. of gender 2009 was the same and performed by NMUIERES, 5 recommendations were pointed out by evaluators; however those were not implemented by the health sector. The M&E card 2014 has not finished yet. In the EED 2009 all recommendations (4) were implemented. The recommendations (3) of the EED 2010 were attended. The M&E card 2014 will not be included in the analysis.  |   |
| ISSSTE                                   | E001                     | Control de Enfermedades Prevenibles por Vacunación   | 2008, 2009, 2010, 2011, 2012, 2013, 2014 | X                              | M&E card 2014  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                | The M&E card 2014 will not be included in the analysis.  |   |
| IMSS                                     | E002                     | Atención curativa eficiente*   | 2008, 2009, 2010, 2011, 2012, 2013, 2014 | X                              | Ev of gender 2009, Specific ev of performance (EED) 2009, EED 2010, Ev. cost-effectiveness 2013, M&E card 2014 |   |                 |      |                |                 |      | 104,312,435,882 | 104,312,435,882 | S/U   | 106,358,417,373 | 106,358,417,373 | A/U  |                |                 |      |                |                 | 146,721,031,540 | 146,721,031,540 | U               |                                | The INFOMEX report says data is in the SHCP. Diagnostic of gender 2008 and Ev. of gender 2009 was the same and performed by NMUIERES, 5 recommendations were pointed out by evaluators; however those were not implemented by the health sector. The Ev. Cost-effectiveness 2013 is in process. The M&E card 2014 has not finished yet. In the EED 2009 recommendation was not implemented. The EED 2010 shows that only 1 of 3 recommendations is still improving, the rest are not mentioned in the mechanism to follow up evaluation results. The M&E card 2014 will not be included in the analysis. |   |
| ISSSTE                                   | E002                     | Control de Enfermedades Transmisibles*   | 2008, 2009, 2010, 2011, 2012, 2013, 2014 | X                              | Ev. Cost-effectiveness 2013, M&E card 2014   |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 | 197,037,330     | 197,037,330     | U                              |  | The M&E card 2014 will not be included in the analysis.   |
| IMSS                                     | E003                     | Atención a la salud en el trabajo*   | 2008, 2009, 2010, 2011, 2012, 2013, 2014 | X                              | M&E card 2014  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                | The M&E card 2014 will not be included in the analysis.  |   |
| ISSSTE                                   | E003                     | Detección oportuna de Enfermedades*  | 2008, 2009, 2010, 2011, 2012, 2013, 2014 | X                              | M&E card 2014  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                | The M&E card 2014 will not be included in the analysis.  |   |
| IMSS                                     | E004                     | Investigación en salud en el IMSS*   | 2008, 2009, 2010, 2011, 2012, 2013, 2014 | X                              | M&E card 2014  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                | The M&E card 2014 will not be included in the analysis.  |   |
| ISSSTE                                   | E004                     | Orientación para la Salud*   | 2008, 2009, 2010, 2011, 2012, 2013, 2014 | X                              | M&E card 2014  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                | The M&E card 2014 will not be included in the analysis.  |   |
| IMSS                                     | E005                     | Cobertura de cotizantes  | 2009                                     |                                |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
| ISSSTE                                   | E005                     | Control del estado de salud de la embarazada*  | 2008, 2009, 2010, 2011, 2012, 2013, 2014 |                                | Ev. Cost-effectiveness 2013, M&E card 2014   |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 | 169,106,501     | 169,106,501     | U                              |  | There is NO DATA of the ev. cost-effectiveness 2013. The M&E card 2014 will not be included in the analysis.                              |
| IMSS                                     | E005                     | Ingresos obrero-patronales   | 2008                                     |                                |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
| IMSS                                     | E006                     | Recaudación eficiente de ingresos obrero patronales  | 2009, 2010, 2011, 2012, 2013             | X                              | Ev of process 2012   |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 | 5,132,788,106   | 5,132,788,106   | U                              |  | The ev. of process 2012 was performed until 2014. Thus, all 18 recommendations are still improving.                                       |
| ISSSTE                                   | E006                     | Atención Materno Infantil*   | 2008, 2009, 2010, 2011, 2012, 2013, 2014 |                                | M&E card 2014  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                | The M&E card 2014 will not be included in the analysis.  |   |
| IMSS                                     | E007                     | Servicios de guardería*  | 2008, 2009, 2010, 2011, 2012, 2013, 2014 |                                | Ev. of design 2009, Ev of gender 2009, Specific ev of performance (EED) 2009, EED 2010, M&E card 2014          |   |                 |      |                |                 |      | 8,116,057,509   | 8,116,057,509   | S/U   | 8,403,627,554   | 8,403,627,554   | A/U  |                |                 |      |                |                 |                 |                 |                 |                                | The INFOMEX report says data is in the SHCP. There is no data of Ev. of design 2009. Diagnostic of gender 2008 and Ev. of gender 2009 was the same and performed by NMUIERES, 5 recommendations were pointed out by evaluators; however those were not implemented by the health sector. The M&E card 2014 has not finished yet. In the EED 2009 all recommendations (3) were not implemented. The EED 2010 shows that only 1 of 3 recommendations is still improving, the rest were not implemented. The M&E card 2014 will not be included in the analysis.  |   |
| ISSSTE                                   | E007                     | Consulta bucal*  | 2008, 2009, 2010, 2011, 2012, 2013, 2014 | X                              | M&E card 2014  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                | The M&E card 2014 will not be included in the analysis.  |   |
| ISSSTE                                   | E008                     | Atención y Mejoramiento Nutricional  | 2008, 2009                               |                                |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
| SSA                                      | E008                     | Vigilancia y control epidemiológico  | 2008                                     |                                |  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                |  |   |
| IMSS                                     | E008                     | Atención a la salud reproductiva*  | 2009, 2010, 2011, 2012, 2013, 2014       | X                              | Ev. Cost-effectiveness 2013, M&E card 2014   |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 | 5,188,010,968   | 5,188,010,968   | U                              |  | The ev. cost-effectiveness 2013 has 16 recommendations which are still improving. The M&E card 2014 will not be included in the analysis. |
| IMSS                                     | E009                     | Prestaciones sociales eficientes*  | 2009, 2010, 2011, 2012, 2013, 2014       | X                              | M&E card 2014  |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 |                 |                 |                                | The M&E card 2014 will not be included in the analysis.  |   |
| ISSSTE                                   | E009                     | Consulta Externa General*  | 2011, 2012, 2013, 2014                   | X                              | Ev. Cost-effectiveness 2013, M&E card 2014   |   |                 |      |                |                 |      |                 |                 |       |                 |                 |      |                |                 |      |                |                 |                 | 5,021,978,604   | 5,021,978,604   | U                              |  | There is NO DATA of the ev. cost-effectiveness 2013. The M&E card 2014 will not be included in the analysis.                              |





[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

|           |      |  |  |   |  |                |                |        |                |                |       |                |                |           |                |                |       |                |                |            |                |                |               |   |   |  |
|-----------|------|--|--|---|--|----------------|----------------|--------|----------------|----------------|-------|----------------|----------------|-----------|----------------|----------------|-------|----------------|----------------|------------|----------------|----------------|---------------|---|---|--|
| SSA-SNDIF | S150 | Programa de Atención a Familias y Población Vulnerable*                            | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 | X | Ev. Consistency and results 2007, Diagnostic of gender equality 2008, Specific ev of performance (IED) 2008, EED 2009, EED 2010, Ev. Consistency and results 2011, EED 2012, M&E card 2014   | 530,776,330    | 525,865,738    | A/D    | 511,543,394    | 489,815,146    | A/S/D | 501,841,902    | 641,069,788    | A/I       | 381,960,068    | 381,960,068    | A/U   | 397,690,601    | 397,690,601    | A/U/I**    | 259,494,531    | 279,494,531    | S/I           | Diagnostic of gender 2008 and Ev. of gender 2009 was the same and performed by INMUJERES. 5 recommendations were pointed out by evaluators; however those were not implemented at the health sector. In the EcyE 2007 recommendations (3) were implemented. Recommendation of the EED 2008 was implemented. There was only one recommendation of the EED 2009 and it was implemented. The only recommendation of the EED 2010 is still improving. In the PPEF 2011 public resources of this programme were increased in 15%. The only recommendation of the EcyE 2011 is still improving, however there are more pointed out by evaluators, which are the same identified by other evaluations. There are no recommendations mentioned in the mechanism to follow up evaluation results of the EED 2012, however they are the same pointed by evaluators in evaluations previously cited. The M&E card 2014 will not be included in the analysis.   |   |  |
| SSA       | S174 | Programa de estancias infantiles para apoyar a madres trabajadoras                 | 2009, 2010, 2011, 2012, 2013, 2014             |   | Ev. of design 2007, Impact ev. 2007, Diagnostic of gender equality 2008, Ev. Of perception 2009, Specific ev of performance (IED) 2009, EED 2010, Ev. Consistency and results 2011, Impact Ev. 2011, EED 2012, Ev. of process 2012, Meta ev. 2007-2012 | 1,000,000,000  | 1,000,000,000  | A/I    | 1,441,034,409  | 1,499,172,686  | I     | 128,739,936    | 122,863,566    | S/S/A/D   | 124,273,664    | 124,273,664    | S     | 202,400,000    | 202,400,000    | S*/S/U/I** | 209,484,000    | 209,484,000    | A/S/U         | This programme is classified in the budgetary process until 2009. However, it exists time before and has developed evaluations since 2007. There is no data of impact Ev. 2007. Diagnostic of gender 2008 and Ev. of gender 2009 was the same and performed by INMUJERES. 5 recommendations were pointed out by evaluators; however those were not implemented at the health sector. In 2007 and 2008 it was classified in the social sector as S174, later on transferred to health sector. Recommendations (10) of the Ev. of design 2007 were implemented, however these are different from those identified by evaluator (6). In the Ev. of perception 2009 there are no recommendations (4) in the mechanism to follow up evaluation results. Only one recommendation of 2 of the EED 2009 was implemented. There are no recommendations mentioned in the mechanism to follow up evaluation results of the EED 2010, however there are the same pointed by evaluators in evaluations previously cited. In the PPEF 2011 the resources of this programme were increased by 61%. Recommendations (4) of the EcyE 2011 were implemented, however there are more pointed out by evaluators (15). There are no recommendations mentioned in the mechanism to follow up evaluation results of the impact Ev. 2011, however there are some pointed by evaluators (5). Recommendations (1) of the EED 2012 are improving, however these are different from those identified by evaluator (6). There are no recommendations as result of Process Ev. 2012 (7) and meta evaluation 2007-2012 in the mechanism to follow up evaluation results. |   |  |
| SSA       | S200 | Caravanas de la Salud*   | 2008, 2009, 2010, 2011, 2012, 2013, 2014       | X | Ev. of design 2007, Complementary ev. 2007, Diagnostic of gender equality 2008, Specific ev of performance (IED) 2008, EED 2009, Complementary ev. 2009, EED 2010, Ev. Consistency and results 2011, EED 2012, M&E card 2014                           | 500,000,008    | 500,000,008    | S/S*/U | 520,000,000    | 514,542,236    | A/D   | 615,840,529    | 614,129,674    | S/S*/S*/D | 760,745,734    | 734,312,234    | A/D   | 726,939,144    | 726,939,144    | S*/U/I/*   | 793,792,186    | 704,096,038    | A/D           | This programme is classified in the budgetary process until 2008. In 2007 was classified as R205. However, it exists time before and developed an evaluation in 2007. Diagnostic of gender 2008 and Ev. of gender 2009 was the same and performed by INMUJERES. 5 recommendations were pointed out by evaluators; however those were not implemented by the health sector. Recommendations (7) of the Design Ev. 2007 are not mentioned in the mechanism to follow up evaluation results. In the Ev. complementaria 2007 only 2 recommendations of 29 are improving. The recommendations (2) of the EED 2009 were implemented. Only one recommendation of the EED 2009 was implemented, however there are more pointed out by evaluators (6). In the Ev. complementaria 2009 only 3 recommendations of 33 are improving. In the EED 2010 only one recommendation of three was implemented. In the PPEF 2011 public resources of this programme were reduced -4.8%. In the EcyE 2011 only 3 recommendations of 20 are still improving. In the EED 2012, 3 recommendations were pointed out by evaluators and still improving, however their recommendations are the same since evaluations of 2007. The M&E card 2014 will not be included in the analysis.  |   |  |
| SSA       | S201 | Seguro Médico para una Nueva Generación  | 2008, 2009, 2010, 2011, 2012                   | X | Ev. of design 2007, Specific ev of performance (IED) 2008, EED 2009, EED 2010, Complementary ev. 2010, Ev. Of Consistency and Results 2011, EED 2012   | No data        | No data        | A      | 2,641,185,204  | 2,641,185,204  | A/S/U | 2,641,636,393  | 2,641,532,522  | S/D       | 2,450,000,000  | 2,450,000,000  | A/A/U | 2,563,310,066  | 2,563,310,066  | S*/U       | 2,653,025,918  | 2,653,025,918  | A/U           | This programme is classified in the budgetary process until 2008. In 2007 the programme classification was not found. However, it exists time before and developed an evaluation in 2007. There is no data of Complementary Ev. 2010. In the Ev. of design 2007, only 2 recommendations of 4 were implemented, this evaluation was developed until 2008. In 2007 the programme classification was not found. In the EED 2008, only one recommendation of 3 was implemented. Recommendations are the same in both evaluations. There are no recommendations mentioned in the mechanism to follow up evaluation results of the Ev. complementaria 2009 and EED 2009. In the EED 2010 were identified 5 recommendations, however these are different from mechanism to follow up evaluation results. Recommendations (3) of the Ev. complementaria 2010 were implemented. Only 3 recommendations were implemented of the EcyE 2011, there are 12 more recommendations not mentioned in the mechanism to follow up evaluation results. Two recommendations were implemented of the EED 2012, however there are 4 more pointed out by evaluation. Recommendations are the same of EED 2010, EcyE 2011 and EED 2012.  |   |  |
| SSA       | S201 | Seguro Médico Siglo XXI*   | 2013, 2014                                     | X | Ev. of design 2013, Ev. Cost-effectiveness 2013, M&E card 2014   |                |                |        |                |                |       |                |                |           |                |                |       |                |                |            |                |                | 2,519,425,918 | 2,519,425,918   | U   | There is NO DATA of the ev. of design and cost-effectiveness 2013. The M&E card 2014 will not be included in the analysis. The latter programme "Seguro médico para una nueva generación" emerged into this "Seguro médico siglo XXI". |
| SSA       | S202 | Sistema Integral de Calidad en Salud*  | 2008, 2009, 2010, 2011, 2012, 2013, 2014       | X | Ev. of design 2007, Ev. of consistency and results 2007, Specific ev of performance (IED) 2008, Complementary ev. 2008, EED 2009, EED 2010, Complementary ev. 2010, EED 2010, Complementary ev. 2010, Ev. Consistency and results 2011, EED 2012       | 116,414,247    | 115,622,808    | A/D    | 121,500,000    | 115,718,875    | S/D   | 188,804,081    | 164,304,081    | A/D       | 162,587,627    | 166,587,627    | I/D** | 163,860,246    | 163,860,246    | S/D        | 163,860,246    | 163,860,246    | S/D           | This programme is classified in the budgetary process until 2008. However, it exists time before (Programa Cruzada Nacional por la Calidad de los Servicios de Salud) and developed one evaluation of two programmed in 2007. See programme 3041. There is no data of EcyE 2011. It used to be "Programa Cruzada Nacional por la Calidad de los Servicios de Salud", see programme 3041. In the EED 2008 only one recommendation of 2 was implemented. In the Ev. complementaria 2008, only 2 of 17 recommendations were implemented. The recommendations (3) of the EED 2009 are not mentioned in the mechanism to follow up evaluation results. In the EED 2010 only one recommendation of 3 was implemented. In the PPEF 2011 public resources of this programme were reduced 4.9%. There is no data of EcyE 2011. There are no recommendations (4) mentioned in the mechanism to follow up evaluation results of the EED 2012. The M&E card 2014 will not be included in the analysis.  |   |  |
| SSA-SNDIF | S250 | Programa de fortalecimiento a la Procuraduría de la Defensa del Menor y la Familia | 2014   |   | Diagnostic 2014  |                |                |        |                |                |       |                |                |           |                |                |       |                |                |            |                |                |               |   | The diagnostic 2014 will not be included in the analysis.   |  |
| SSA-SNDIF | S251 | Programa de desarrollo comunitario "Comunidad Diferente"                           | 2014   |   | Diagnostic 2014  |                |                |        |                |                |       |                |                |           |                |                |       |                |                |            |                |                |               |   | The diagnostic 2014 will not be included in the analysis.   |  |
| SSA       | U004 | Financiamiento equitativo de la atención médica                                    | 2008   |   | Specific ev. Of performance (IED) 2008   | 25,459,280,465 | 25,204,642,096 | A/D    |                |                |       |                |                |           |                |                |       |                |                |            |                |                |               |   | This evaluation is part of the U005. All recommendations (6) of the EED 2008 were implemented. This evaluation is part of the U005. |  |
| SSA       | U005 | Seguro Popular*  | 2009, 2010, 2011, 2012, 2013, 2014             | X | Specific ev of performance (IED) 2008, EED 2009, EED 2010, Ev. of Consistency and Results 2011, Ev. of process 2012, EED 2012, Ev. Cost-effectiveness 2012 and 2013, M&E card 2014   | 41,368,161,540 | 41,368,161,540 | S/U    | 52,378,927,045 | 48,842,127,045 | S/D   | 57,267,820,560 | 56,946,687,061 | A/I**/D   | 65,652,458,192 | 64,402,458,192 | S/D   | 66,791,938,639 | 66,791,938,639 | U          | 66,791,938,639 | 66,791,938,639 | U             | This programme is classified in the budgetary process until 2008. However, it exists time before and developed an evaluation in 2007. It is the same as the "Sistema de Protección Social en Salud". The EED 2008 is the same of the U004. There is no data of Process Ev. 2012. There are no recommendations in the Ev. cost-effectiveness 2012 pointed out by evaluators. In the EcyE 2007, 14 recommendations were implemented, only one is still improving. There are no recommendations (3) as result of EED 2009 in the mechanism to follow up evaluation results. In the PPEF 2011 public resources of this programme were increased in 12.7%. 7 recommendations of the EcyE 2011 are improving, however there are more (4) not included in the mechanism to follow up results. There are no recommendations (4) as result of the EED 2012 in the mechanism to follow up evaluation results. There is NO DATA of the ev. cost-effectiveness 2013. The M&E card will not be included in the analysis.   |   |  |
| SSA       | U006 | Fortalecimiento de las Redes de Servicios de Salud*                                | 2009, 2010, 2011, 2012, 2013, 2014             |   | Specific ev of performance (IED) 2010, Ev. of Consistency and Results 2011, EED 2012, M&E card 2014  |                |                |        |                |                |       |                |                |           | 142,170,689    | 2,611,170,689  | I     | 332,000,000    | 4,675,200,002  | I          | 50,000,000     | 2,251,892,641  | I             |   | There is no data of any evaluation (EED 2010, EcyE 2011, EED 2012). The M&E card 2014 will not be included in the analysis.         |  |
| SSA       | U007 | Reducción de la mortalidad materna*  | 2012, 2013, 2014                               |   | M&E card 2014  |                |                |        |                |                |       |                |                |           |                |                |       |                |                |            |                |                |               |   | This is the same programme as P018. The M&E card 2014 will not be included in the analysis.   |  |
| SSA       | U008 | Prevención contra la obesidad*   | 2012, 2013, 2014                               |   | M&E card 2014  |                |                |        |                |                |       |                |                |           |                |                |       |                |                |            |                |                |               |   | This is the same programme as P019. The M&E card 2014 will not be included in the analysis.   |  |
| SSA       | U009 | Vigilancia epidemiológica*   | 2012, 2013, 2014                               |   | M&E card 2014  |                |                |        |                |                |       |                |                |           |                |                |       |                |                |            |                |                |               |   | The M&E card 2014 will not be included in the analysis.   |  |

[illegible]

S\* = recommendations are classified in status quo when they represent less than 30% of total.

I\*\* = in 2011 public resources were increased compared to previous year based on evaluation results.

Note: In the PAE 2007 the “Programa Primer Empleo” of the IMSS has to develop an impact evaluation but there is no evidence that this programme actually exists in the IMSS neither in the STPS.

\*These programmes are considered PART of the social sector and coordinated by the CONEVAL.

In the “Exposición de motivos del PPEF 2011”, the public resources of the programme U001 (Programa de atención a situaciones de contingencia laboral) were reduced from 30.1 to 29.5 (-5.8) million of mxn pesos. However, this programme does not appear in the PEF 2010.

Source: author's own elaboration.



Table 6

## The effectiveness of policy evaluation in Chile

| Public organization   | Key budgetary programmes | Budgetary programmes   | Year in the federal budget                     | Indicator of budgetary process   | Type of evaluation performed   | Amendments = A) Cancellation = C) Learning process = L) Status quo = S) Allocation of financial resources (Increase = I) Decrease = D) Unchanged = U) |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 | Recommendations of evaluations |               |  |   |            |           |  |   |  |
|---|--------------------------|--|--|--|--|---|-----------------|------|----------------|-----------------|------|----------------|-----------------|------|----------------|-----------------|------|----------------|-----------------|------|----------------|-----------------|------|----------------|-----------------|--------------------------------|---------------|--|---|------------|-----------|--|---|--|
|   |                          |  |  |  |  | Project budget  | Budget approved | 2007 | Project budget | Budget approved | 2008 | Project budget | Budget approved | 2009 | Project budget | Budget approved | 2010 | Project budget | Budget approved | 2011 | Project budget | Budget approved | 2012 | Project budget | Budget approved |                                | 2013          | 2014   |   |            |           |  |   |  |
| Undersecretary of Healthcare Networks   |                          | Red de urgencia  |  | This programme is not listed in the budget 2007-2014 and DPEPS' website. The DPEPS' office letter says this evaluation is in the process of public action. | Comprehensive Spending Review (SCQ) 2012 Review a no es la ex. Del gasto institucional |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               | The government action process mentions this evaluation was contracted during 2013. Cost of the evaluation is \$118,000,000.00 Chilean pesos (\$192,186.00 USD).  |   |            |           |  |   |  |
| National Health Fund (FONASA), Undersecretary of Public Health and of Healthcare Networks |                          | Programa tratamiento de la obesidad  |  | This programme is not listed in the budget 2007-2014. The programme is run only by FONASA.   | Evaluation of Public Programme (EPG) 2008  |   |                 |      |                |                 | A    |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               | Cost of the evaluation is \$11,800,000.00 Chilean pesos (\$19,202 USD). From the total of 14 recommendations suggested by evaluators, 3 were implemented whereas 5 of them (no. 1, 3, 5, 6) did not achieve and cancelled in the report to following up results.   |   |            |           |  |   |  |
| Undersecretary of Healthcare Networks   |                          | Establecimientos experimentales (Hospital Padre Hurtado, Centro Mapu y Cordillera Oriente) |  | This programme is not listed in the budget 2007-2014. This is a set of programmes  | Comprehensive Spending Review (SCQ) 2010   |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      | S              |                 |      |                |                 |                                |               | Cost of the evaluation is \$70,000,000.00 Chilean pesos (\$113,877.00 USD). The finances each year are integrated by the three hospitals. From the total of 25 recommendations suggested by evaluators, two were not implemented (no. 14, 16); the latter neither related between recommendation and action taken. There are 23 recommendations not mentioned in the report to following up results (no. 1 to 13, 15, 17 to 25). |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Fondo Nacional de Salud  | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  | Comprehensive Spending Review (SCQ) 2011   |   | 1,880,042,234   |      |                | 2,138,276,030   |      |                | 2,525,362,695   |      |                |                 |      |                |                 |      |                | S               |      |                | 3,587,539,216   |                                | 3,970,506,948 | 4,423,129,516  | The government action process mentions this evaluation was contracted during 2012. Cost of the evaluation is \$80,000,000.00 Chilean pesos (\$146,413.00 USD). It is also named as "Ex. de gastos institucionales". Organizational spending in the report issued by the organization is quite general and does not mention specific actions as result of 17 recommendations suggested by evaluators. Only 2 of these recommendations were implemented. Report available at: <a href="http://www.dipres.gob.cl/595/articles-124665_fonasa.pdf">http://www.dipres.gob.cl/595/articles-124665_fonasa.pdf</a> |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Caja de prevision de la Defensa Nacional   | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Programa de Desarrollo Indígena MIDEPLAN-80  | 2007, 2008, 2009                               |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Sistema Chile Solidario. In 2013 it's "Ley no. 20.595 and Sistema Chile Solidario"         | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Sistema de Protección Integral a la Infancia   | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Curativo cargo usuario   | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Otros  | 2007, 2008                                     |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Beneficiarios de salud   | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Subsidio cajas de compensación de asignación familiar                                      | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Programa de atención primaria (1)  | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Programa de prestaciones vulnerables (2)   | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   | 596,747,650     |      |                | 651,070,947     |      |                | 704,919,362     |      |                |                 |      |                |                 |      |                |                 |      |                | 759,005,607     |                                | 938,136,303   | 1,139,270,829  | 1,335,510,368   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Programa de prestaciones institucionales (3)   | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Compra privada: Programa prestaciones vulnerables  |  | THIS PROGRAMME INCLUDES 4 DIFFERENT PROGRAMMES OF THE SAME DIFFERENT ORGANIZATIONS. This programme is not listed in the budget 2007-2014.                  | Evaluation of Public Programme (EPG) 2013  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                | S             |  | This is a programme under responsibility by the FONASA, MINGSA, providers which do not belong to the Sistema Nacional de Servicios de Salud (SNS), its public network hospitals and the Purchasing or procurement office and public contracts (Dirección de compra y contratación pública). Integrated by four programmes: compra de prestaciones diarias, de camas integradas, de prestaciones bono AUGE, de otras prestaciones. There are 11 recommendations suggested by evaluators, however there is no report to following up results.   |            |           |  |   |  |
| National Health Fund (FONASA)   |                          | Bono AUGE (Programa de Acceso Universal de Garantías Explícitas)                           | 2011, 2012, 2013, 2014                         | It belongs to the programme "PRESTACIONES VALORADAS"   | Evaluation of New Programme (EPN) 2013   |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                | 6,269,428       |                                | 4,112,000     |  | 1,232,926   | A          | 3,051,814 | The final report says to impact evaluation 2013 but the government action process mentions this evaluation is of impact and performed during 2011. Cost of the evaluation is \$80,000,000.00 Chilean pesos (\$130,145.00 USD). There are 3 recommendations suggested by evaluators but there are still improving. There is only a report by the organization, which mentions some actions taken. |   |  |
| Institute of Public Health  |                          | Instituto de Salud Pública (4)   | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| Institute of Public Health  |                          | Secretaría y administración general hacienda   | 2011   |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| Institute of Public Health  |                          | Portal integrado de comercio exterior (DICE)   | 2011   |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| National Supply Center  |                          | Central de Abastecimiento del Sistema Nacional de Salud                                    | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  | Comprehensive Spending Review (SCQ) 2010   |   | 4,713,138       |      |                | 5,108,635       |      |                | 5,492,046       |      |                |                 |      |                |                 |      | S              |                 |      |                | 6,210,093       |                                | 6,330,720     |  | 6,068,430   |            | 6,106,199 | 7,300,501  | Cost of the evaluation is \$70,000,000.00 Chilean pesos (\$113,877.00 USD). From the total of 19 recommendations suggested by evaluators, one was implemented (no. 12) but there is no relation between recommendation and action taken, 2 were implemented (no. 2, 7), as well as the recommendation no. 8, which has 3 different suggestions but only one was implemented. There is one recommendation not implemented (no. 10) and five are not mentioned in the report to following up results (no. 1, 3, 4, 5, 9). |  |
| Undersecretary of Public Health   |                          | Programa de fortalecimiento de la estrategia digital de Chile                              | 2007, 2008                                     |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| Undersecretary of Public Health   |                          | Ley de accidentes del trabajo  | 2007, 2008, 2009                               |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| Undersecretary of Public Health   |                          | Subsidios por accidentes del trabajo   | 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014 |  |  |   |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |      |                |                 |                                |               |  |   |            |           |  |   |  |
| Undersecretary of Public Health   |                          | Programa nacional de alimentación complementaria   | 2009, 2010, 2011, 2012, 2013, 2014             |  | Evaluation of Public Programme (EPG) 2013  |   | 30,095,146      |      |                | 33,171,792      |      |                | 35,228,443      |      |                |                 |      |                |                 |      |                |                 |      |                |                 | 35,756,870                     |               | 44,381,835   |   | 45,624,326 |           | 46,947,637   | S   | There are 12 recommendations (3 of design, 3 of management, 3 of efficacy/quality, 2 of efficiency and 1 of economic) suggested by evaluators, however there is no report to following up results. |

[illegible]



Source: author's own elaboration.

S\* = recommendations are classified in status quo when they represent less than 30% of total.

(1) This programme provides services to the 28 healthcare centers (Arica, Iquique, Antofagasta, Atacama, Coquimbo, Valparaíso-San Antonio, Viña del Mar-Quillota, Aconcagua, Libertador General Bernardo O'Higgins, Maule, Ñuble, Concepción, Talcahuano, Bío-Bío, Arauco, Araucanía Norte, Araucanía Sur, Valdivia, Osorno, Chiloé, Aysén del General Carlos Ibáñez del Campo, Magallanes, Metropolitano Oriente, Metropolitano Central, Metropolitano Sur, Metropolitano Norte, Metropolitano Occidente, Metropolitano Sur-Oriente) and the programme of operational contingencies (Programa de contingencias operacionales). In 2009 was included the hospital Padre Alberto Hurtado and the healthcare center of Reloncaví.

(2) This programme provides services to the 28 healthcare centers (Arica, Iquique, Antofagasta, Atacama, Coquimbo, Valparaíso-San Antonio, Viña del Mar-Quillota, Aconcagua, Libertador General Bernardo O'Higgins, Maule, Ñuble, Concepción, Talcahuano, Bío-Bío, Arauco, Araucanía Norte, Araucanía Sur, Valdivia, Osorno, Chiloé, Aysén del General Carlos Ibáñez del Campo, Magallanes, Metropolitano Oriente, Metropolitano Central, Metropolitano Sur, Metropolitano Norte, Metropolitano Occidente, Metropolitano Sur-Oriente), the programme of operational contingencies (Programa de contingencias operacionales), Hospital Padre Alberto Hurtado, Centro de referencia de salud de Maipú, Centro de referencia de salud de Peñalolén Cordillera Oriente and the agreements of provision of health benefits (convenios de provisión de prestaciones médicas). In 2009 was included the healthcare center of Reloncaví. In 2011 was included Bono AUGE.

(3) This programme provides services to the 28 healthcare centers (Arica, Iquique, Antofagasta, Atacama, Coquimbo, Valparaíso-San Antonio, Viña del Mar-Quillota, Aconcagua, Libertador General Bernardo O'Higgins, Maule, Ñuble, Concepción, Talcahuano, Bío-Bío, Arauco, Araucanía Norte, Araucanía Sur, Valdivia, Osorno, Chiloé, Aysén del General Carlos Ibáñez del Campo, Magallanes, Metropolitano Oriente, Metropolitano Central, Metropolitano Sur, Metropolitano Norte, Metropolitano Occidente, Metropolitano Sur-Oriente), the programme of operational contingencies (Programa de contingencias operacionales), Hospital Padre Alberto Hurtado, Centro de referencia de salud de Maipú, and Centro de referencia de salud de Peñalolén Cordillera Oriente. In 2009 was included the healthcare center of Reloncaví.

(4) This programme includes the programme Undersecretary of Public Health and the programme to international organisations. In 2014 the programme to international organisations does not appear.

(5) This programme provides services to the 29 healthcare centers (Arica, Iquique, Antofagasta, Atacama, Coquimbo, Valparaíso-San Antonio, Viña del Mar-Quillota, Aconcagua, Libertador General Bernardo O'Higgins, Maule, Ñuble, Concepción, Talcahuano, Bío-Bío, Arauco, Araucanía Norte, Araucanía Sur, Valdivia, Osorno, Chiloé, Reloncaví, Aysén del General Carlos Ibáñez del Campo, Magallanes, Metropolitano Oriente, Metropolitano Central, Metropolitano Sur, Metropolitano Norte, Metropolitano Occidente, Metropolitano Sur-Oriente). It also includes the healthcare center of Maipú in 2007, 2008, 2011 but does not appear in 2009 and 2010. In 2011 included Centro de referencia de salud de Peñalolén Cordillera Oriente.