ESSAYS ON IMPERPECT INFORMATION AND INDIVIDUAL CHOICE

Lucio Castro

Dissertation submitted for the degree of:

Dphil in Economics

University of Sussex

October 2013

SUMMARY

This dissertation is organized in three separate chapters. A common thread intertwined in the three of them is the role of informational flows at influencing individual´s choice. Chapter 1 provides novel empirical evidence on the role of imperfect information about price schedules on determining consumer's demand choice for the residential gas market using a natural experiment and consumer microdata from the metropolitan region of Buenos Aires, Argentina. Based on the results a large field experiment carried out in a municipality of Argentina, Chapter 2 presents new empirical results showing that taxpayers are relatively insensitive to information on the prevailing level of tax evasion and the supply and quality of local public goods, whereas raising the salience of fines and other related penalties may have a large bearing on tax compliance. Relatedly, Chapter 3 presents novel evidence, also based on a field experiment conducted in another Argentine municipality, suggesting that informational treatments influence taxpayers' beliefs about the risk of detection and the salience of penalties in the case of non-compliance, and hence, their compliance behaviour with respect to the tax code.

ACKNOWLEDGEMENTS

I want to thank my advisor Andrew Newell for his unconditional support along this very, very, long, and sometimes, disparaging journey. His cheerful, positive and joyful attitude towards economics, and more importantly, towards life, played a crucial role in every defining moment of this dissertation. I also want to thank my friends and colleagues Dr. Eduardo Levy Yeyati and Dr. Guillermo Cruces, who threatened to stop talking to me if I did not finalize my Phd. Dr. Juan Pablo Rud also played a crucial role, with his indefatigable enthusiasm about my topic of dissertation, and more generally, about my research work. I also want to thank my co-author in Chapter 2, Dr. Carlos Scartascini, for helping me to significantly shorten my learning curve up on how to write top-level academic research with concrete and relevant policy implications.

I am also grateful to Antonella Mancino, who provided excellent support with the residential gas consumption microdata utilized in Chapter 1. I would also like to thank the Majors and the staff from the Municipalities of Junín and 9 de Julio in Argentina for their courage and collaboration in the field experiments presented in Chapters 2 and 3 and Diego Jorrat, Alejo Vidal and Estefania Lottito for their assistance in the collection and processing of the municipal administrative data. I am also grateful to Paula Szenkman, who provided extraordinary support in the coordination of the field work in both field experiments. I also want acknowledge the support of CIPPEC (Center for the Implementation of Public Policies for Equity and Growth) in Argentina and the Interamerican Development Bank (IDB), and particularly, the Institutional Capacity Strengthening Fund (ICSF), funded by the Government of the Popular Republic of China, in carrying out the field experiments.

Last but not least, I want to thank my family, my wife Eugenia and my son Joaquin, for their patience and support over so many days, nights, weekends and holidays of hard work.

I also want to dedicate this dissertation to my mother, Veronica, and my Grannies, Sara and Cecilia, who always believed in me. This dissertation is also theirs.

INTRODUCTION

This dissertation is organized in three separate chapters. Chapter 1, co-authored with Dr. Carlos Scartascini, Dr. Julian Christia and Dr. Paulo Pastos, and accepted for publication at the Journal of Industrial Economics (JIE) (see Appendix 1: Supporting Documentation), applies semi-experimental methods to estimate the price elasticity of residential gas demand by applying a regression discontinuity design (RDD) to microdata of household consumption in the metropolitan area of Buenos Aires in Argentina. The chapter exploits the unique features of a recently introduced tariff schedule for natural gas in Buenos Aires to estimate the short-run impact of price shocks on residential energy utilization. The schedule induces a non-linear and non-monotonic relationship between households' accumulated consumption and unit prices, thus generating exogenous price variation, which we exploit in a regression-discontinuity design. The results indicate that a price increase causes a prompt and significant decline in gas consumption. The results also indicate that consumers respond more to recent past bills than to expected prices, which argues against an assumption of perfect awareness of complex price schedules by consumers.

Chapter 2, co-authored with Dr. Carlos Scartascini, explores the effects of informational flows on the individual taxpayer's decision on whether or not complying with the tax code. It reports the results of a field experiment that tried to affect taxpayers' compliance by affecting their beliefs regarding the levels of enforcement, equity, and fairness of the tax system. In the experiment, conducted in an Argentine municipality, taxpayers were divided in four groups; three of those groups received a specific treatment (a message on the tax bill). Results indicate that the most effective message was the one that stated the actual fines and potential legal consequences taxpayers may face in the case of non-compliance. Tax compliance among those taxpayers increased by more than 4 percentage points (and this value may be underestimating the true impact). There are no average effects for the other treatments. However, the evidence also points out that not every taxpayer updates his or her beliefs in the same direction. This evidence may help to square off some of the opposing results in the literature.

Chapter 3, using a similar approach to Chapter 2, investigates the importance of informational flows at influencing the two main determinants of the standard model of tax evasion developed by Allingham and Sandmo (1972) and Yistaki (1974). I carried out a large field experiment in another municipality of Argentina that consisted in sending messages to the entire population of taxpayers of a local property tax aimed at influencing their subjective beliefs about the probability of enforcement of the tax code and the potential economic and legal consequences of tax cheating. In the experiment, I randomly divided the population of taxpayers in four groups. Three groups received specific treatments: a penalty-based message, a monitoring-based letter, or both. The empirical results indicate that these instruments may not be substitutes (or that at least, that the underlying elasticities are different). While the penalty-based message had a positive and significant effect (tax compliance for

those taxpayers increased by around 3 percentage points with respect to the control group) I do not find significant average treatment effects for the monitoring-based letter. The average effect for those taxpayers that received both treatments is statistically similar to the group that received only the penalty-based message. I find even larger effects for non-compliant taxpayers with multiple properties.

A common thread that entwines the three chapters is the empirical analysis of the importance of informational flows at influencing individual choice. Chapter 1 provides semi-experimental evidence on the role of imperfect information about price schedules on determining consumer's demand choice in the residential gas market. Based on a large field experiment, Chapter 2, in turn, suggests that taxpayers are relatively insensitive to information on the prevailing level of tax evasion and the supply and quality of local public goods, whereas raising the salience of fines and other related penalties may have a large bearing of tax compliance. Relatedly, Chapter 3 shows, based also on a randomized field experiment, that informational treatments might influence differentially taxpayer's subjective perceptions about the risk of detection and the salience of penalties, and hence, her compliance decision with respect to the tax code.

TABLE OF CONTENTS

TABLES AND FIGURES

CHAPTER 1. Does energy consumption respond to price shocks? Evidence from a regression-discontinuity design

1.1. Introduction

Suppose that energy prices experience a shock. Does energy consumption respond? How much and how promptly? These are key questions in the study of a wide range of macroeconomic, regulatory and environmental issues, such as the transmission channels of energy price shocks, optimal taxation and pricing policies in energy markets, and interventions to address climate change. Naturally, economists have a long-standing interest in estimating consumption responses to price changes in energy markets.[1] Progress towards this aim has been complicated by an important identification challenge, however. Since consumers typically experience the same events at essentially the same time, it has been difficult to construct the equivalent of randomly assigned treatment and control groups and thereby ground the estimated price-elasticities on a well-defined counterfactual (Reiss and White, 2008).

In this chapter, we exploit unique features of a recently introduced tariff schedule for residential natural gas in Greater Buenos Aires to estimate the short-run impact of price shocks on residential gas utilization.[2] The revised schedule induces a non-linear and non-monotonic relationship between annual previously accumulated consumption and unit prices, thus giving rise to exogenous price variation. In particular, the introduction of a threshold for defining unit prices based on previously accumulated consumption approximates a randomly assigned price differential for a large number of consumers on each side the threshold, allowing us to build treatment and control groups to estimate the impact of interest. We estimate the demand effect of a price shock using a regression discontinuity (RD) design whereby the gas consumption levels of households situated barely above the tariff discontinuity are compared with those of households located barely below.

We implement the RD design on administrative records of the natural gas distribution company, gathering longitudinal data on the unit prices and consumption levels of each consumer located near the tariff discontinuity–the same information that these households received in their utility bills. We find that an increase in the average price of natural gas in the utility bill received by consumers induces a statistically significant and prompt decline in residential gas utilization: a 25% price increase reduced residential gas consumption by 3.8% in the subsequent two-month period. This

[1] Research on this topic, discussed in more detail below, dates to Parti and Parti (1980), Dubin and McFadden (1984) and Hsiao and Mountain (1985). Recent influential contributions include Reiss and White (2005, 2008).

[2] Greater Buenos Aires (*Gran Buenos Aires*) is an urban metropolitan area that comprises the city of Buenos Aires and 24 adjacent municipalities (National Institute of Statistics and Censuses, 2003). According to the 2010 population census it has 12.8 million inhabitants, nearly a third of the total Argentinean population.

result suggests that policy interventions via the price mechanism may constitute a powerful instrument to influence the patterns of residential energy utilization, even in a relatively short time span.[3]

An important feature of our research design is that it exploits the specific information set available to consumers to estimate the effect of interest. In particular, because households are reclassified every billing period on the basis of their annual accumulated consumption, fully informed consumers in the treatment and control group face the same expected prices moving forward. However, we provide survey evidence suggesting that consumers have highly imperfect knowledge about the price determination mechanism, and appear to infer prices from recent past bills. In addition, we offer strong statistical evidence that consumers do not manipulate strategically their annual accumulated consumption. For these reasons, the resulting estimates are especially relevant for residential energy markets characterized by ex-post billing where households infer changes in prices from the utility bill. Importantly, while it has long been emphasized that this feature of residential energy markets plays an important role in shaping consumption responses to price changes (Shin, 1985), there is still little direct evidence on whether and how promptly energy consumption responds to price variations inferred from utility bills.

We complement and extend several strands of literature. A number of studies on the price-elasticity of energy demand employ time series methods using data on energy prices and aggregate energy consumption (Liu and Lin, 1991; Krichene, 2002; Bushnell and Mansur, 2005; Fezzi and Bunn, 2010). A related body of work draws on cross-sectional survey data, including influential papers by Parti and Parti (1980), Dubin and McFadden (1984), Dubin (1985) and Reiss and White (2005). While these methods allow for the estimation of long-term impacts, the aggregated or cross-sectional nature of the data imposes relatively strong identifying assumptions. Furthermore, estimates yielded by cross-sectional data are, by construction, silent on the speed with which energy consumption adjusts to price shocks, an issue that is of key interest in a variety of policy contexts.

Another strand of research estimates price-elasticities in the context of tariff field experiments, including early work by Hausman, Kinnucan and McFadden (1979), Acton and Mitchell (1980), Caves and Christensen (1980) and Parks and Weitzel (1984). Whereas this approach addresses some limitations of the time-series and cross-sectional evidence, it has been criticized on the ground that the (most often voluntarily-selected) set of participants are thoroughly informed about price changes at the outset, generating an informational context that differs significantly from real-world situations in which households learn about price changes from utility bills or the press (Acton, 1982; Reiss and White, 2008).

[3] The evidence we provide may therefore contribute to the discussion on the relative importance of prices versus nudges for steering consumers' behaviors (Loewenstein and Ubel, 2010).

Two recent papers using disaggregate billing data on electricity consumption from California are perhaps the closest to our own. Reiss and White (2008) examine how price shocks and conservation appeals impact residential electricity consumption. Their estimates point to sizable short-run impacts on energy utilization. In independent work, Ito (2010) exploits a spatial discontinuity in electric utility service areas in southern California, which leads to nearly identical households experiencing different nonlinear price schedules. His contribution is highly complementary to ours. Consistent with our results for the natural gas market, he finds that residential consumers in the electricity market respond to (lagged) average price, rather than marginal or expected marginal price.[4,5]

The remainder of the chapter is structured as follows. Section 2 provides background information on the market for natural gas in Greater Buenos Aires and describes the data employed. Section 3 describes the research design and provides important complementary evidence from a survey of consumers located near the tariff discontinuity of interest. Section 4 presents the econometric results. Section 5 provides a discussion of the results in the context of the literature. Section 6 offers some concluding remarks.

1.2. Background and data

We focus on the residential market for natural gas in Greater Buenos Aires. In late 2008, the national energy regulatory agency (ENARGAS) introduced the first increase in residential gas tariff since 2002. The revised schedule was composed of eight new tariff groups, each facing different variable fees per cubic meter; see Table 1. This schedule introduced significantly higher unit prices for those consumers with higher levels of annual accumulated consumption. We exploit the resulting discontinuity in variable prices at the annual consumption level that divides categories R32 and R33 to examine consumption responses to price shocks.

Our empirical analysis draws on administrative records from MetroGAS S.A., one of the largest residential gas distributors in Argentina. The company has a client base in Greater Buenos Aires of about 2.5 million residential households, who receive their gas bills every two months. The administration of MetroGAS agreed to provide us with data on a small share of its customer base around the discontinuity R32-R33.[6] Given this constraint, we have defined the sample with the view to optimize the implementation of the RD design for this discontinuity. Specifically, consumers were

[4] Focusing on the gas market in the US, Davis and Muehlegger (2010) exploit variation in wholesale prices to estimate the price elasticity of demand in the residential sector. We believe that our RD estimates add value to this literature in that they: (1) are less likely to be affected by unobserved heterogeneity across households; and (2) emphasize the role of imperfect information in shaping consumer behavior in this market.

[5] See also Jessoe and Rapson (2011) for a recent analysis of commercial and industrial demand response under mandatory time-of-use electricity pricing in Connecticut.

[6] To preserve confidentiality of its broad customer base, the company refused to provide us with data on consumers from other categories.

selected into the sample provided by the company if: (i) they had a residential bill issued in May 2009 with an annual accumulated consumption between 1480 and 1520 cubic meters; and (ii) they had been customers for at least six bimonthly cycles by that month. The resulting estimation sample contains 7190 households.

This sample is composed of longitudinal records corresponding to the bills issued in May 2009, in the five previous bimonthly periods, and in the three subsequent ones. For each period, the data set comprises information on the amount billed, quantity of gas consumed (in cubic meters), type of reading (measured or imputed), category assigned to the consumer, and the exact dates of reading and issuance. It further contains information on the region and neighborhood of residence of each consumer.

1.3. Research design

An ideal experiment designed to estimate the impact of a price shock on residential energy consumption would randomly assign some consumers to a treatment group, facing price P_H, and other consumers to a control group, facing price P_L. Unfortunately, a large-scale experiment of this kind has yet to be implemented, making the task of estimating this behavioral response rather difficult. To approximate such an ideal experiment, we exploit unique features of the price schedule described in section 1.2, along with the specific information set available to consumers.

In May 2009, households with annual accumulated consumption above 1500 cubic meters were assigned a unit price roughly 25% higher than those with an annual accumulated consumption barely below this level. This discontinuity of the unit price schedule makes it possible to apply a RD design in which the outcome variable corresponds to the consumption level in the subsequent two-month period and the running variable to the annual accumulated consumption.

However, as we explain in detail below, the interpretation of the RD design in this application is made difficult by two important features of the price determination mechanism: (i) the category to which consumers are assigned to, and hence the unit price they are charged, is determined by the accumulated consumption of the previous 12 months; and (ii) the categorization of consumers is revised every two months, in line with the variation of the 12-month accumulated consumption over that period.

1.3.1. The price determination mechanism

Let us define the key variables underlying the determination of the amount billed in a given bimonthly period 0. The annual accumulated consumption AAC_0 corresponds to the sum of the consumption level C in period 0 and in the five previous bimonthly periods:

$$AAC_0 = \sum_{j=-5}^{0} C_j \tag{1}$$

The unit price in period 0 is a function of whether accumulated consumption is above or below a given threshold:[7]

$$P_0 = \begin{cases} P_L \text{ if } AAC_0 \leq 1500 \\ P_H \text{if} AAC_0 > 1500 \end{cases} \tag{2}$$

The total bill B in period 0 can be expressed as:

$$B_0 = FC + P_0 C_0 + \mu_0 \tag{3}$$

where FC is a fixed cost, P_0 is the unit price in that period, and μ_0 is an idiosyncratic shock capturing the fact that the bill received by consumers sometimes contains idiosyncratic adjustments and retroactive charges (e.g., taxes and other charges defined by the regulator on a rather ad hoc manner).

Finally, while consumers may target consumption levels in the following period, they are unable to perfectly control their gas consumption patterns. Hence, actual consumption will differ from targeted consumption (CT) by a random shock. That is:

$$C_1 = CT_1 + e_1 \tag{4}$$

In our setting, Period 0 corresponds to that billed in May 2009. Consumers in the treatment group are those with annual accumulated consumption barely above 1500 cubic meters in period 0, while consumers in the control group are those with annual accumulated consumption barely below this level. Whether or not we would expect this price shock to have a differential impact on gas consumption in period 1 crucially depends on the specific information set held by consumers.

Since households are reclassified every period on the basis of their annual accumulated consumption, fully informed consumers in the treatment and control group face the same expected price moving forward. Hence, if perfectly informed, both groups would have essentially the same

[7] For simplicity, in this section we focus on consumers with annual accumulated consumption between 1,251 and 1,800 cubic meters who can therefore face only two potential prices (see Table 1).

incentive to restrain consumption so as not to surpass the 1500 cubic meters threshold in period 1, despite the fact that the bill received in period 0 contained sharp differences in unit prices. However, in the light of the well-documented prominence of information imperfections in residential energy markets with ex-post billing (Shin, 1985), and considering the complexity and novelty of the price determination mechanism in the residential market for natural gas of Greater Buenos Aires, we would expect consumers to be imperfectly informed. To provide further evidence on the information set held by consumers, we have surveyed a subset of households in the estimation sample.

1.3.2. Survey evidence on consumers located near the discontinuity

We have administered a telephone survey to 353 households from the estimation sample. The sub-sample surveyed was stratified by district to ensure an adequate geographical representation. The survey was conducted in September 2010 and targeted the member of the household that was responsible for paying the gas bill.[8]

The survey questionnaire consists of two sections. The first section collected information on basic socioeconomic characteristics of the household head (age, education, occupation) and housing conditions. The second section explored knowledge about the amount billed and payment method, and assessed the extent to which consumers read their utility bills. In addition, this section collected extensive information on perceived and objective knowledge on how the amount billed is computed; specifically, on: (i) how frequently tariffs are determined; (ii) the past consumption periods used for their determination; and (iii) the consumption threshold that leads to a higher unit price.

Table 2 reports summary statistics on the member of the household surveyed. For comparison, it also provides summary statistics from the national Household Survey (*Encuesta Permanente de Hogares*) on households living in Greater Buenos Aires. Relative to the average resident, surveyed household members from our estimation sample are more likely to be female and married. They are also more likely to have tertiary education, and tend to be considerably older than the average resident. These differences are likely to reflect the fact that the survey targeted the household member that was responsible for paying the gas bill. The fact that surveyed households are more likely to have tertiary education might also suggest that they tend to have higher-than-average income. Table 2 further suggests that surveyed households are more likely to be homeowners, and have a slightly larger number of rooms and families. This evidence is consistent with the fact that our estimation sample focuses on households of relatively high gas consumption. Further data from the questionnaire suggests that natural gas is the dominant source of energy used by these households:

[8] The survey used records from 2952 households in order to obtain 353 valid responses from the member of the household that is responsible for paying the gas bills. A failure to produce a valid interview occurred when the enumerator: (i) was unable to establish a phone conversation with any member of the household; (ii) was unable to reach the person responsible for paying the gas bills; or (iii) the person responsible for paying the gas bills refused to cooperate. The response rate of the survey (12%) is well in line with that of recent phone surveys in the US – a recent study by the Pew Research Center (2012) reports that a standard survey yields a response rate of 9%.

76.1% of households use it for space heating, 95.2% use it for water heating, and 99.4% use it for cooking.

Table 3 reports the key results from the survey.[9] Nearly 92% of households reported that they were able to remember the amount charged in the last gas bill. The proportion of households who paid their bill by direct debit is relatively small (14%), which alleviates the concern that consumers might not be aware of how much they are charged every period. About 75% of consumers stated that they regularly read their gas bills.

However, knowledge about the price determination mechanism proved to be almost non-existent. Among surveyed households, 31% stated that they know how the price is determined. However, the questions aimed at assessing precise knowledge of the price determination mechanism suggest that the proportion of well-informed consumers is considerably lower. First, only 14% of households knew that consumers are re-categorized (and unit prices are determined) in each billing cycle. Second, 39% of consumers knew that their billing category is determined on the basis of the accumulated consumption over the previous year. Third, only 4% of consumers knew that the threshold that divides categories R32 and R33 is 1,500 cubic meters. Overall, less than 1% of households provided correct answers to the three objective questions posed.

In summary, the survey reveals that consumers tend to know how much they are paying for their gas consumption, but have scant information about the actual pricing scheme. Consequently, in the remainder of this chapter, we will assume that the vast majority of consumers have imperfect information about the prevailing price determination mechanism and infer prices from the utility bill.

1.4. Econometric model

Under the assumption that most if not all consumers have imperfect information about the price determination mechanism, we can estimate the short-term effects of price shocks by applying a sharp RD design. That is, we can compare gas utilization in period 1 for consumers that in period 0 had annual accumulated consumption barely above and below the 1500 cubic meters threshold, as both these sets of consumers are expected to be very similar along observed and unobserved characteristics but experienced very different unit prices. Since we can reasonably assume that households infer prices from recent past bills, differences in consumption in period 1 across both groups of consumers can be interpreted as the short-run behavioural response to the price shock.

To implement the RD design we estimate the following regression model:

$$C_{i,1} = \beta_0 + \beta_1 Treatment_{i,0} + f(\overline{AAC}_{i,0}) + \omega_{i,0} \tag{5}$$

[9] The full set of results is available upon request.

where $C_{i,1}$ corresponds to consumption in period 1 for consumer i. $\overline{AAC}_{i,0}$ corresponds to the running variable, the normalized annual accumulated consumption in period 0. That is:

$$\overline{AAC}_{i,0} = AAC_{i,0} - 1500 \tag{6}$$

The treatment variable is a binary indicator of whether individual i in period 0 was assigned the higher unit price. It is determined as:

$$Treatment_{i,0} = \begin{cases} 0 \text{if } \overline{AAC}_0 \leq 0 \\ 1 \text{if } \overline{AAC}_0 > 0 \end{cases} \tag{7}$$

Parameter β_1 in (5) captures the average effect of barely surpassing the threshold, and hence having received a substantially larger utility bill, once we flexibly control for the running variable, $\overline{AAC}_0$. The intuition behind this approach is that all observable and unobservable variables should evolve smoothly around this threshold, and hence any jump in consumption in period 1 can be attributed to the discontinuous increase in the amount billed. In other words, in a valid RD design, a key assumption is that observations are randomly distributed between the treatment and control group in a local neighbourhood of the threshold (Lee and Lemieux, 2010). In our setting, this assumption seems plausible for three reasons. First, as documented above, consumers have highly imperfect knowledge about the price determination mechanism. Second, they tend to be unaware of the location of the relevant threshold. Third, even if they had perfect information about the pricing scheme, it would be difficult and time-consuming – if at all possible – to precisely manipulate gas utilization so as to avoid passing this cutoff. Indeed, this would require obtaining access to gas readings, forecasting factors affecting future demand, and knowing exactly when the billing period ends (i.e. the exact timing of the meter reading).

We focus on behavioural reactions during the period of mid-May to mid-July 2009, as these two winter months account for the bulk of annual gas consumption (about 30% of the total). It would have been possible to examine consumption patterns of these households in later billing cycles. However, beginning in June 2009, public reactions motivated by the increase in gas prices generated a disruption in the normal billing process of the company. This generated differences in the timing of bill issuance for individuals on both sides of the cutoff, making it difficult to disentangle between timing effects and those stemming from the price shock.

Though we would like to compare observations just above and below the threshold, in practice a larger window around the cutoff has to be used to obtain precise estimates. In choosing the width of this window, researchers typically face a trade-off between bias and precision: a wider window provides greater precision but at the expense of expected higher bias. In our setting, because

the utility company had records on almost two million consumers, we were able to select a narrow window around the threshold and still have a sizable sample size of 7200 consumers.

Our baseline specification consists of a local linear regression that controls for normalized annual accumulated consumption as expressed in (6). Although in general it is recommended to control for the running variable to minimize the potential bias (Imbens and Lemieux, 2008), in this particular application the close relationship between the outcome and the running variable justifies following this approach. For robustness, we additionally report estimates: (i) reducing the window in the running variable used to select consumers; (ii) controlling for higher order polynomials of the running variable; and (iii) allowing for a differential slope between the outcome and the running variable on both sides of the cutoff.

We cluster standard errors by the running variable, as suggested by Lee and Card (2008) for cases in which this variable is discrete.[10] It has been pointed out that when the number of clusters is small, standard errors may be downward biased (Angrist and Lavy, 2009). In our baseline specification, there are 41 clusters which may be a large enough number (Cameron, Gelbach and Miller, 2008). Nonetheless, this problem becomes more serious when the number of clusters is reduced, as a narrower window in the running variable is used. However, the results are robust to two proposed solutions to this problem: (i) computing averages of the relevant variables by clusters and reproducing the analysis at this level; and (ii) selecting the highest between robust-clustered and conventional standard errors (Angrist and Pischke, 2008).

1.5. Results

1.5.1. Testing the validity of the research design

The basic identifying assumption of RD design is that the outcome variable would have been continuous at the assignment threshold in the absence of the treatment (Lee, 2008). Albeit this assumption cannot be tested directly, we provide evidence on this issue by examining whether a number of covariates are continuous at the threshold. We define a treatment group composed of consumers in the (1,20) cubic meters of normalized annual accumulated consumption by May 2009 and a corresponding control group for those in the (-20,0) cubic meters range. Given that a small bandwidth is used to select these two groups of consumers, differences in the running variable between them are minimal, as the average difference in this variable is only 20 cubic meters – which represent only 1.3% of the mean (20/1500). Hence, as a first approximation it is possible to compare average values between the treatment and control groups to inspect for evidence in favour of the identifying assumption. We additionally run local linear regressions to test for the existence of jumps

[10] Measured consumption is always rounded to the nearest integer, and hence we cluster the standard errors at the unit levels of the running variable.

in covariates at the threshold once we control for differences in normalized annual accumulated consumption.

Table 4 presents results for key dates and period lengths. Panel A shows that the timing of events is very similar between the treatment and control groups, while Panels B and C document that the length of critical periods is almost identical across groups. Similar patterns emerge when we test for jumps in these variables, by regressing them on the treatment dummy while controlling for the normalized annual accumulated consumption in period 0 (Column 4). Results not reported (but available on request) show that there are significant differences between the analysed dates and period lengths across regions, suggesting the relevance of exploring balance in timing patterns. Inspecting balancing in period lengths for Period 1 is critical to attribute differences in total consumption during that period to changes in consumer behaviour. The results provide evidence that actual gas consumption is recorded every two months. Hence consumption reported in the administrative records corresponds to actual consumption and not to imputations by the firm. Moreover, gas bills are issued approximately one week after the final measurement for the period and should be received by consumers approximately 10 days after a period ended, according to the firm. This leaves about 50 days for consumers to react.

In Table 5 we examine differences in the geographic distribution across groups. Though in general the results point to adequate balancing along this dimension, in 4 out of 10 regions statistical differences at the 10% level are found when comparing the treatment and control groups. However, the results show that only in one case there are statistical differences at the 10% level once we control linearly for normalized annual accumulated consumption (adjusted difference column).

In Table 6 we analyse differences in historical consumption patterns. Panel A presents raw and adjusted differences between consumption in periods -5 to 0. By construction, the annual accumulated consumption of the treatment group will be higher than that of the control group. Hence the existence of some statistically significant differences when analysing raw differences is not surprising. Though expected, these findings highlight the importance of adjusting for the running variable. When doing this, we tend not to observe statistically significant differences across groups. Panel B presents a more clean test of pre-treatment differences between groups by comparing consumption in each period as a share of the total annual consumption. In this dimension, the treatment and control groups present strikingly similar patterns, suggesting once more that the RD design is able to yield unbiased estimates in this context.

Figures 1 to 3 depict the results presented in Tables 4 to 6. The same patterns highlighted in the tables clearly stand out from these figures: the covariates considered are smooth around the discontinuity. Importantly, Table 7 and Figure 4 clearly show that the average amount billed in period

0 slightly increases as normalized annual accumulated consumption raises but jumps drastically when the latter crosses the cutoff.[11]

It has been stressed in the RD design literature that this approach will not be suitable if agents can manipulate the running variable, implying that the condition that individuals on both sides of the discontinuity are similar is not fulfilled (McCrary, 2008; Lee and Lemieux, 2010). For the reasons outlined above, we would expect the scope for this manipulation to be limited in our setting. Nevertheless, to explore this issue further we follow McCrary (2008) and examine the density distribution of the running variable, in particular whether there is a jump in this density around the threshold. Figure 5 shows that the density is quite flat and does not point to the existence of any discontinuity around the threshold.

1.6. Estimating the short-run impacts of the price shock

1.6.1. Main results

Given the evidence confirming the validity of the research design, we now turn to the primary focus of this chapter: the impact of the price shock on gas consumption in the subsequent billing period. Table 8 presents the results. In specification (1), we regress gas consumption in period 1 on the treatment dummy, while controlling linearly for normalized annual accumulated consumption. The results reveal that experiencing a price shock induces a statistically significant drop in gas consumption of 15.9 cubic meters in the subsequent period (or roughly 3.8% of the average gas consumption). Figure 6 depicts these results. There is a clear positive relationship between consumption in period 1 and normalized annual accumulated consumption in period 0, as would be expected given that consumers with higher consumption in the past also tend to consume more in the future. But, most importantly, gas consumption seems to fall discontinuously at the threshold, suggesting that households react to the price shock by significantly reducing consumption in the subsequent two-month period. The estimated effect is sizable if considering that, given the short time span, it is unlikely that consumers will adjust to the new price via investments in more efficient appliances or improvements in insulation. Moreover, since consumers typically learn about the new price approximately 10 days after the beginning of the period, they have only about 50 out of approximately 60 days to adjust to the price shock.

Results from other specifications show that these estimates are quite robust. In specifications (3), (5) and (7) the control function becomes progressively more flexible as we add second, third and fourth order terms. The estimated coefficient remains remarkably stable, hovering between 15.2 and 16.5. In all cases, the results are statistically significant at least at the 10% level, although as expected

[11] To verify if the amount billed in period 0 actually followed the tariff structure prevailing at the time, we have predicted the bill by applying the prevailing unit prices, adding the fixed charge and applying the minimum billed amount. The correlation between predicted and actual bills is 0.98, suggesting that the firms bills in period 0 followed closely the prevailing price schedule.

the standard errors are larger in more flexible specifications. The remaining columns present the corresponding estimates when allowing for differential shapes of the control function at both sides of the cutoff. Once again, the estimated coefficient remains robust, although its precision falls markedly in more flexible specifications.

Table 9 presents further evidence on the robustness of the results as we use an increasingly narrow bandwidth, thus restricting our attention to observations progressively closer to the threshold. Although the coefficients become less precisely estimated when restricting to observations located closer to the cutoff, the estimated impact remains virtually unchanged. Table 1 in the Appendix 2 further shows that the estimated impacts are robust to the inclusion of neighbourhood fixed-effects.

1.6.2. Alternative hypotheses

In addition to the issues addressed in the previous section, a potential threat to the validity of our estimates is that they might be partially driven by mean reversion. If gas consumption does not follow a random walk, households who experience a positive demand shock in period 0 might be expected to reduce consumption in period 1 relative to identical households with a negative or zero consumption variation in period 0.[12] To account for this hypothesis, in Table 10 we examine differential consumption patterns in period 1 for consumers located just above and just below two placebo AAC thresholds (where consumers did not experience a price shock). If negative serial correlation were to explain our estimates, we would expect to observe significant differentials in consumption patterns around other AAC thresholds. But the estimates in Table 10 show that this is not the case. We find no differential consumption patterns across consumers located just above and just below such fake thresholds, while we do confirm the results relative to the threshold of interest.

As mentioned above, an important concern about our analysis is that the estimated treatment effects might partially reflect strategic behaviour of households in the treatment group. In particular, rather than just reacting to the price shock, these households might have an incentive to reduce consumption in order to fall below the threshold of interest in the subsequent billing period. Under this alternative scenario, those consumers in the treatment group that are located just above the threshold should be expected to reduce consumption by less than those consumers in the treatment group located that are farther away from the threshold. To examine this hypothesis, we divide the treatment group into two different sets of consumers: a "Close" group, including those consumers with AAC in period 0 between 1501 and 1510 cubic meters; and a " Far" group encompassing those consumers with AAC in the 1511 to 1520 interval. The control group continues to be composed of consumers with AAC in Period 0 between 1480 and 1500 cubic meters. To fall below the thresholds of interest, individuals in the "Close" group need to reduce their consumption in Period 1 (compared

[12] For example, if a household receives a guest in period 0 gas demand would like increase in that period, returning to normal levels in period 1 once the guest leaves.

to Period 0) by 5 cubic meters, whereas those in the "Far" group need to reduce it by 15 cubic meters, on average. If consumers in the treatment group were to behave strategically, both the "Close" and "Far" groups would be expected to reduce consumption in Period 1, but the latter group would be expected to reduce it by 10 cubic meters more on average. Alternatively, if consumption responses were mainly driven by the fact that all consumers in the treatment group experienced a price shock, both the "Close" and "Far" groups would be expected to decrease consumption by a similar amount. To distinguish between these alternative hypotheses, we estimate an equation of the form:

$$C_{i,1} = \beta_0 + \beta_1 Close_{i,0} + \beta_2 Far_{i,0} + f(\overline{AAC}_{i,0}) + \omega_{i,0} \quad (8)$$

where $Close_{i,0}$ and $Far_{i,0}$ have the meaning described above. Table 11 documents that the reduction in consumption for the "Close" and "Far" groups, as compared to the control group, is very similar (-16.3 versus -18.5) providing little support to the hypothesis that consumption responses are motivated by strategic behaviour.

1.7. Discussion

While the quasi-experimental setting we adopt provides unique features for examining short-run consumption responses to price shocks, some caution is needed in extrapolating to other situations. An important drawback of the RD design is that the results are local, in that they refer to the particular threshold studied, and may not generalize to the other points in the distribution of the running variable. This caveat applies to the present analysis. In the specific market we study, consumers with annual accumulated consumption around the 1500 cubic meters threshold have above-the-average gas consumption. The extent to which the results generalize to other consumers is an open question. Households with lower levels of annualized consumption may present higher price sensitivity, if they are lower-income consumers and, therefore, more price-conscious. On the other hand, these consumers may be less sensitive to price changes, as lower baseline gas utilization may signal a low weight of this good in their overall consumption basket. Nevertheless, the study of price sensitivity for high-consumption households may be interesting in its own right, as they account for a sizable share of overall gas consumption.

A second limitation of our analysis is that the estimated impact refers to the consumption response to a positive price shock. It is difficult to assess the extent to which the results would generalize to policy interventions inducing a fall in the price faced by consumers – e.g. binding price caps and subsidies. Indeed, existing research suggests that consumers react differently to price increases and decreases (see, e.g. Dargay, 1993).

A third caveat of our study is that, while the tariff schedule we exploit offers a unique ground for examining consumption responses to price shocks, caution is needed in extrapolating our estimates to other price schemes. Upon exceeding the 1500 cubic meter threshold, households face both a

discontinuous price shock and an increase in the marginal cost of each additional unit consumed. Furthermore, as discussed above, perfectly informed consumers could reverse this discontinuous shock on future bills with a small decrease in consumption. It is important to emphasize that this setting differs clearly from graduated price schemes, where consumers face a higher marginal cost for all consumption above a threshold but a lower rate for consumption below the threshold. Our estimates are potentially more informative on how consumers might be expected to respond to an across-the-board rate increase, in which the price of all units of gas (marginal and inframarginal) increases. However, valid extrapolation to those situations hinges on households believing that their efforts to reduce consumption have only a marginal effect, as opposed to a discontinuous effect, on subsequent bills. Although the evidence we provide suggests that consumers are generally unaware of (or do not understand well) the threshold, and also that they do not manipulate strategically their gas consumption around it, we cannot exclude the possibility that (some) households realize that a marginal effort to reduce consumption might reduce discontinuously their subsequent gas bill.

With these caveats in mind, this chapter offers solid evidence that prices do matter for energy consumption in energy markets, even over short-run durations. As Reiss and White (2008, pp.654) emphasize, this fact is important as it has not been widely recognized by policy makers, in part due to the absence of clear and unambiguous evidence of such behaviour. It is therefore interesting to compare our estimates to those obtained in recent research on consumption responses to price changes in energy markets. At about -0.15, the short-run estimated elasticity we report is in line with recent estimates for the natural gas market. Drawing on state-level panel data for the US, Bernstein and Griffin (2006) estimate a short-run elasticity estimate of -0.12. Also for the US, but using disaggregated panel data at the customer level in an instrumental variables approach, Davis and Muehlegger (2010) report an estimated price elasticity of demand of -0.28. Our estimates are also similar to those of recent studies using disaggregated data for the residential electricity market in California. Focusing on San Diego households, Reiss and White (2008) find that a price increase of 130% induced a consumption decline of 13% in about 60 days. Ito (2010) exploits a spatial discontinuity in service areas in southern California, which leads to nearly identical households experiencing different nonlinear price schedules, and estimates a short-run elasticity with respect to lagged average prices of -0.11.[13]

Another important conclusion from our analysis is that the way in which customers process information about complex tariff structures is an important driver of their behaviour. Indeed, we provide survey evidence that consumers have highly imperfect knowledge about the price determination mechanism, and appear to infer prices from past utility bills. In addition, we offer strong statistical evidence that households do not manipulate strategically their gas consumption

[13] Reiss and White (2008) survey earlier literature using cross-sectional survey data for household electricity consumption, and conclude that the cross-sectional estimates vary widely, with typical values near -0.3.

around the relevant threshold, despite strong incentives for doing so under the prevailing price determination mechanism. The evidence we provide is therefore especially relevant for residential energy markets characterized by complex tariff structures and ex-post billing. While it has long been stressed that imperfect information in residential energy markets plays an important role in shaping consumption responses to price changes (Shin, 1985), there is still little direct evidence on consumer behaviour in the presence of complex price schedules and ex-post billing. In this regard, our findings are consistent with (and complementary to) those of recent studies by Bushnell and Mansur (2005) and Ito (2010). Examining the electricity market in San Diego, Bushnell and Mansur (2005) provide evidence that customers respond more to out-dated prices from their last bill than to current market conditions. Ito (2010) finds that consumers respond to lagged average prices rather than marginal or expected marginal prices when faced with nonlinear electricity price schedules. Taken together this evidence suggest that, in order to precisely estimate the impacts of complex price schedules on energy utilization, policy makers and regulators need to move beyond the assumption of perfect awareness of such schedules by consumers.

1.8. Concluding remarks

We have exploited unique features of the tariff schedule for natural gas in Greater Buenos Aires, along with survey evidence on the specific information set possessed by consumers, to estimate the short-run effect of price shocks on residential gas consumption. The revised tariff schedule induced a non-linear and non-monotonic relationship between annual accumulated consumption and unit prices, thus generating exogenous price variation. Drawing on administrative records on the utility bills of residential consumers, we have estimated the short-run consumption response to a price shock using an RD design whereby two-month consumption levels of households situated barely above an important tariff discontinuity are compared with those of consumers located barely below–hence focusing on a large group of relatively homogeneous consumers facing sizable differences in prices.

We provide evidence that a price increase in the utility bill received by consumers causes a prompt and significant decline in gas consumption. We also show that customers appear to respond more to recent past bills than to expected prices moving forward. Our estimates therefore suggest that policy interventions via the price mechanism are powerful instruments to influence residential energy utilization patterns, but argue against an assumption of perfect awareness of complex price schedules by consumers.

CHAPTER 2. Tax Compliance and Enforcement in the Pampas. Evidence from a Field Experiment

2.1. Introduction

Tax evasion is a pervasive problem in many countries. In particular, some developing countries do not even collect half of what they are supposed to according to the written letter of the law.[14] The academic literature has not been oblivious to the need for explaining why people pay (or don't pay) taxes, and the striking differences in compliance across countries. Most of the literature has relied on the standard model of tax compliance developed by Allingham and Sandmo (1972), which suggests that each individual taxpayer faces a tradeoff between the monetary benefit of evading and pocketing the money, and the potential costs of being detected and having to pay the evaded tax plus a penalty. According to these models, how much each person evades is the result of the optimal decision of risk adverse individuals that maximize their expected utility. Tax evasion is supposed to decrease whenever the expected penalty (determined by the size of the penalty and the probability of getting caught) increases.

Even though this model, along with Yitzhaki (1974), has been the workhorse behind most academic research and its results have guided public policies, some of its implications do not square well with real world evidence. In particular, with the fact that people tend to comply with the law more regularly than what the model would suggest given that in most countries fines and audit probabilities are rather low (Alm et al 1992, Andreoni et al 1998, Dhami and al-Nowaihi 2007, Frey and Torgler 2007). To explain this puzzle, two main strands of literature have developed by grounding the research in behavioral economics. One strand modifies the axioms of expected utility theory by adding risk and uncertainty to the models (Yaniv 1999, Snow and Warren 2005, Dhami and al-Nowaihi 2007 and 2010). The other incorporates additional costs to the expected utility function by making the decision process of the individual taxpayer to be affected also by social interactions (Cowell and Gordon 1988, Gordon 1989, Myles and Naylor 1996, Kim 2003, Bayer 2006, Fortin et al 2007, Eisenhauer 2008, Dell'Anno 2009, Traxler 2010). In these latter models, taxpayers do not only weigh the monetary payoffs but also the moral implications of their decisions. What other taxpayers are doing and how the government uses public revenues have a high burden on tax evasion decisions (Torgler 2002).

Combining these two strands provides predictions more in line with the empirical evidence: most people declare more than what the standard model would suggest and not everybody engages in tax evasion (Hashimzade et al 2012). In terms of their policy implications, levels of compliance seem

[14] For example, that is the case for personal income taxation and corporate income taxation in Argentina, Dominican Republic, Ecuador and Guatemala, according to recent studies summarized in IDB (2013).

to depend not only on actual rates but also on people's subjective beliefs about the levels of enforcement and penalties, the behavior of other taxpayers, the use of public monies by the government, and on some additional intrinsic motivations. As such, it may be possible to increase the levels of tax compliance by raising the salience of enforcement and penalties, the level of compliance of their neighbors, and the efficiency and/or effectiveness of government expenditures, where trust on government authorities may be a mediating factor.[15]

While the theoretical literature has advanced steadily, the empirical literature has not reached the same level of consensus yet. For example, while most studies to date find a positive result for reducing tax evasion by increasing the rates or salience of penalties and audits (e.g., Blumenthal et al 2001a), there is still controversy about the effect of messages that appeal to moral considerations, and those reflecting the use of public monies by the government. For example, while Alm et al (1992), Torgler (2003) and several survey-based studies find confirmatory evidence, Blumenthal et al (2001b), Torgler (2004), and Fellner et al (2011) do not find common effects across all type of taxpayers.

One way to test the hypotheses and shed new light to this literature is by affecting taxpayers' beliefs with information (IDB 2008; Coleman, 1996, 2007, Thaler and Sunstein, 2008) in a large-scale randomized field experiment. For that purpose, we conducted an experiment that affected the collection of the most relevant municipal tax in a municipality of Argentina. This tax, locally known as *tasas*, is levied upon individuals according to the size of the property and the services they receive from the local government such as street-lighting, and trash collection and street cleaning. Approximately 23,000 individual taxpayers of this *tasa* in the Municipality of Junín, a midsize and largely urbanized district located in the upper north of Argentina´s main province Buenos Aires, who are billed bimonthly, were randomly divided into 4 groups. One of the groups received no treatment (the control group); the other 3 were treated by including messages in their tax bill. The treatments were designed to test the main determinants of tax compliance according to the literature: deterrence (or beliefs about enforcement and fines), equity (or beliefs about other taxpayers' behavior), and fairness (or the beliefs about the use of resources by the government) (Congdon, Kling, and Mullainathan, 2011, Hashimzade, Myles, and Tran-Nam, 2012).

The results from the experiment indicate that introducing messages in the tax bill may be a good instrument for affecting taxpayers' behavior. Still, not all the messages seem to have the same effect. The most effective message was the deterrence one, which stated the actual fines and potential administrative and judicial steps that the municipality might follow in the case of non-compliance. More precisely, tax compliance among the taxpayers that received this deterrence message increased

[15] Some studies looking at this issue include Aguirre and Rocha (2010), Alm and Martinez-Vazquez (2007), Torgler et al. (2008), Alm and Torgler (2004), Cummings et al. (2005), Dell'Anno (2009); Frey and Torgler (2005), Scholz and Lubell (1998a), Scholz and Pinney (1995), and Murphy (2004).

by almost 5 percentage points with respect to the control group (which is equivalent to reducing tax evasion by more than 10%).[16] We don't find any average treatment effects for the other two messages. This result, which is in line with Torgler (2004) and Fellner et al (2011), contradicts the evidence in the survey-based literature.[17] While people tend to answer that they would be willing to pay more taxes, this behavior is not corroborated in practice. Interestingly, we do find some heterogeneous effects across the population, which indicates that not everybody reacts to the messages, and among those who responses can go on either direction (higher or lower compliance).

In addition to its policy relevance (a low-cost high-return policy innovation), this chapter contributes to the existing literature on the importance of information flows for influencing the individual taxpayer´s compliance decision (e.g., Blumenthal et al, 2001a, 2001b; Pomeranz, 2010; Kleven et al, 2011; and Fellner, Saugruber and Traxler, 2011) in several ways. First, having grounded the research design on both, the behavioral economics and tax evasion theoretical literature, this chapter gets closer to performing a more comprehensive test of the different hypotheses out there by evaluating the three main hypotheses (equity, fairness, and deterrence) at the same time. This way, it is possible to compare their relative significance better than if done across studies.

Second, it provides further evidence regarding which (if any) of the mechanisms usually identified by the theoretical literature is relevant for explaining taxpayer behaviors. As such, this chapter is influential not only because of its findings but also because of what it does not find. As mentioned above, it reinforces the notion that enforcement matters while reducing the confidence in across-the-board mechanisms that stress equity and fairness considerations. If the evidence keeps mounting, these results should encourage researchers to evaluate how much to stress the use of tax morale mechanisms in the theoretical models as a way to square off the theory with the stylized facts with, and to weigh appropriately the findings of the survey-based literature. For example, some theoretical papers require a high value of the parameter of tax morale to revert the Allingham and Sandmo's result on the effect of changes on tax rates on evasion. The existence of heterogeneous effects may also encourage researchers to develop more complex models of social interaction based on heterogeneous agents over a wider range of parameters.

Third, the experiment was carried out on a property-based tax. Instead, the previous literature has focused on income taxes (e.g. Blumenthal et al, 2001a, 2001b; Kleven et al, 2011), VAT (Pomeranz, 2010), or very specific fees (Fellner, Saugruber and Traxler, 2011). Because the revenue of this tax is directly related to the provision of highly-visible locally-provided public services, it

[16] Our results may be underestimating the true effect. Even though every taxpayer received the message, it's not certain that all of them have read it.

[17] Finding different results between opinion-based surveys and behavior-based experiments is not unique to this case. Lacetera et al (2013) report similar results in the area of blood donations. There, while people report in surveys that they would be less likely to donate if offered a monetary reward, the field experiment evidence does not support those findings.

should work better for measuring the impact of messages that target the fairness of the tax system on actual payments (Torgler 2004).

Fourth, by concentrating the research design on studying the decision of paying or not paying an amount that is "exogenously" determined, as we do here, helps to estimate more precisely the impact of the policy (because it reduces the possibility of changes in behavior in other markets –e.g. labor market-, and increases in tax elusion that may affect some of the estimations based on declared income) and provides a direct measure of reductions in tax evasion, which changes in reported income may not capture. Moreover, this chapter uses as outcome variable actual payment behavior instead of willingness to pay. While survey responses could be very valuable for understanding changes in perceptions (a relevant policy experiment by itself), it is only through measuring changes in actual payment behavior that researchers can feel confident that they are actually testing the hypotheses in the theoretical models.

Fifth, while previous studies typically rely on mailing letters to potential evaders, we include the messages and images directly into the tax bill, ameliorating some (but not all) of the concerns regarding intent-to-treat (ITT) bias (there will still be some taxpayers, particularly among the group of tax evaders who don't even read the bill).[18]

Sixth, instead of sending the messages to a selected sample of known evaders, which may lead to problems of sample selection and external validity, we have randomized the treatment over the entire population of compliant and non-compliant taxpayers. This way, we included every type of taxpayer -which may have the added bonus in terms of policy implications of not only helping to generate positive incentives for those who had failed to pay in the past but also to maintain compliance of the formerly "good" taxpayers. This strategy shows also relevant as we uncover heterogeneous effects across this splitting. As such, it helps to guide better policies that may be otherwise aimed at the universe of taxpayers (with negative consequences).

Finally, focusing on a developing country helps to validate the results in the literature in a region, Latin America, where tax evasion is rampant. Therefore, understanding the relative effects of different policies is more relevant. Moreover, this is a region where trust in government is lower than in the developed world, which again, could further our understanding of the effect of messages in such an environment. As a side product, we are also providing with an exact measure of tax evasion on property-like tax for an average size Municipality in the region.

The rest of the chapter is organized as follows: Section 2.2 reviews the literature on the determinants of tax compliance and discusses current approaches on how to increase compliance

[18] In a survey conducted by the authors on a sample of representative taxpayers from another Municipality, 82 percent answered that they had read at least some parts of the bill. This percentage drops to 56% when asked about whether they read any other accompanying material sent by the Municipality.

based on the standard tax evasion model and recent developments on behavioral economics related to the effects of tax morale. Section 2.3 describes the field experiment and our main results as well as some additional empirical analyses. Section 2.4 presents our main findings. Section 2.5, concludes and discusses some potential avenues for future research.

2.2.Why do people pay taxes?

As it was already mentioned, the standard model of tax compliance (Allingham and Sandmo, 1972 (AS); Yitzhaki, 1974 (Y)) suggests that taxpayers face a tradeoff between the monetary benefits of evading and pocketing the money, and the potential costs of being detected and having to pay the evaded tax plus a penalty. Tax evasion levels are the optimal decision of risk adverse individuals that maximize their expected utility. Given the assumptions of this model, tax evasion decreases whenever the expected penalty (determined by the amount of the penalty and the probability of getting caught) increases.[19]

One way the literature has found to explain why people comply more than what the actual rates of enforcement and penalties would suggest has been to add a term to the standard formulation that captures the additional individual (psychic, moral, social) costs that the taxpayer faces when evading (Gordon 1989, Myles and Naylor 1996, Kim 2003, Fortin et al. 2007, Traxler 2010). The parameters included in this additional term usually depend on the beliefs of the taxpayer regarding the fairness and equity of the system.[20] Then, following Traxler (2010), the optimization problem can be formulated as follows

$$U(e_i) = E[u(e_i)] - \theta_i e_i c(n)$$

where *e*, the share of tax being evaded, is the variable of choice of the individual, the first term is the traditional expected utility term coming from the AS (1972) and Y (1974) models, and the second is a term that captures the moral cost of tax evasion. In this formulation, the "moral" cost depends on how much each individual internalizes the social cost [θ_i], which may depend on how much each individual evaluates the degree of fairness of the tax system, and a function *c(n)* that captures the strength of the norm for a given share of evaders *n* (or the equity of the system).[21]

The policy implications of this setting are obvious. Tax evasion drops as penalties and enforcement increases (as in the traditional model), and it also decreases as θ_i (fairness) increases and *n* decreases (or equity increases) –for formal proof see Traxler (2010, pages 92-93). Overall, the empirical evidence (in particular survey and cross-section studies) tends to support these results.

[19] For space limitations, and because these are very well known models in the economics profession, we don't reiterate their formulation here. For comprehensive overviews see Traxler (2010) and Hashimzade et al (2012) among others.

[20] Dell'Anno (2009) presents a very detailed and schematic table summarizing the literature on tax morale and compliance.

[21] As mentioned, this is one of the many ways to model the moral costs. See Hashimzade et al (2012) for a survey of different alternative models.

People tend to comply more if they believe that others comply too, and if they believe that the government makes a good use of the money it collects. According to survey evidence, there is a significant correlation between tax morale and the size of the shadow economy (Alm and Torgler, 2004) and individuals who have heard about tax avoidance have a significantly lower tax morale than others (Torgler, 2005). It has also been found that a higher sense of obedience leads to higher compliance, so that if an individual believes that others comply with the tax code then their tax evasion will diminish (Scholz and Pinney, 1995, Scholz and Lubell, 1998; Torgler et al 2008).[22] It has also been observed from laboratory experiments and surveys that there is a positive correlation between taxpayers' priors about the tax code fairness and tax compliance (Rawling, 2004, Cumming et al, 2006). In connection to the behavior of public officials, there is a positive correlation between tax morale and the belief that the government will spend the money wisely. For example, Barone and Mocetti (2009) find that that the attitude towards paying taxes is more positive when resources are spent more efficiently.[23]

Factors affecting tax morale can explain differences in tax compliance across countries or regions with similar tax systems. For example, Bergman and Narvaez (2006) rely on differences in tax morale to explain the success of Chile and the failure of Argentina in collecting taxes over the nineties. Recent empirical evidence also shows that there is significantly lower tax morale in South America and Mexico than in Central America and the Caribbean (Torgler, 2005).

Evidence compiled at the Inter-American Development Bank (IDB) using data from the Americas Barometer for a selected group of countries for which data is available show similar results (Scartascini, 2011).[24] Basically, people who are more satisfied with the provision of public services at the local level, who trust the municipality more, who have not been exposed to corruption, and who consider the level of interpersonal trust to be higher tend to answer that they would be more willing to pay higher taxes to the municipality. Similar evidence has been found by Ortega, Ronconi and Sanguinetti (2012) based on household survey data for 17 Latin American cities. In their study, the net share of people responding that they would be willing to pay more taxes if the government improved their performance is positive and significant in most cities and for most measures of better performance. Moreover, they find that reading a brief piece of factual information was able to generate significant changes in perceptions in some but not all of the cities in the study. In one of them they find evidence of reciprocity and of a fairly sizable magnitude.

[22] Relatedly, Traxler and Winter (2012) find that the more commonly a norm violation is believed to occur, the lower is the individual's inclination to punish it.

[23] It is important to notice that this does not capture trust in government in general, but the individual's opinion about government's spending behavior (Aguirre and Rocha, 2010; Alm and Martinez-Vazquez, 2007; Murphy, 2004).

[24] Survey waves 2008 and 2010 for Colombia, Dominican Republic, Guatemala, Honduras, Peru, and Uruguay.

2.2.1. Existing experimental evidence

As Ortega et al (2012) show, in order to change people´s perceptions it may not be necessary to change actual policies. The way in which people form their beliefs depends on the information they receive and the way it is presented, and their underlying system of beliefs (IDB 2008, Thaler and Sunstein 2008). Coleman (1996, 2007) conducted a randomized controlled trial in Minnesota using messages to affect tax compliance. The State Revenue Service sent a letter to 20,000 taxpayers with a message designed to correct the erroneous perception of many taxpayers that cheating on taxes was common. The impact of the message was significant: the mean increase in federal declared taxable income for the group receiving the letter was $2,390. Wenzel (2001, 2005) found similar results in Australia.

Blumenthal et al (2001a) studied the impact of messages sent to 1724 randomly selected taxpayers in Minnesota as well. The message consisted of informing that the returns they were about to file would be closely examined. The result was that those in the treatment group (individuals who received the letter) increased their average tax payments compared to previous years. In particular, the effect was stronger for the self-employed, who may have a higher opportunity to evade. Blumenthal et al (2001b) does not find significant average effects when the messages present normative appeals instead.

Torgler (2003) conducted a lab experiment in a small village of Costa Rica where only 37 volunteers –actual taxpayers- participated. The experiment consisted on three types of interventions in order to study tax compliance. The first was related to public goods provision financed from the taxes paid, the second was a moral suasion, and the third was a positive reward for tax compliance. Results show that the three interventions had a positive effect on tax compliance. The moral suasion message, even though large and positive for explaining timelines of payments, was not statistically significant when utilized in a randomized controlled trial with 580 taxpayers in a small town in Switzerland (Torgler 2004).

In the same vein, Kleven et al (2011) conducted a randomized field experiment over a representative sample of 40,000 taxpayers in Denmark. In the first stage, half of the population was randomly selected to be audited, while the rest was deliberately not audited. In the second stage, they sent threat-of-audit letters to a random sample in the two groups. Both the audits and the letters had a significant positive effect on self-reported income. Similar results were found by Pomeranz (2010), who finds that random audit announcements are transmitted up the VAT chain and increase compliance by firms' suppliers.

Finally, Fellner, Sausgruber and Traxler (2011) carried out a large field experiment with potential evaders of TV license fees in Austria sending different messages to taxpayers. They find a strong effect on compliance, particularly for a treatment aimed at raising the salience of detection

risks. Neither appealing to morals nor providing information about other taxpayers' compliance behavior had any significant effect on compliance. A potential limitation of this study is that it does not provide a precise measure of evasion, as TV license avoidance is hard to detect.

Still, in spite of the growing number of articles on the topic, this is still a nascent literature and more work is warranted. First, while some results seem to have become robust findings, such as the positive effect of the threat to audit on compliance, there is still no definitive answer about some of the other potential determinants. For example, while some papers find positive effects for moral suasion treatments (Torgler, 2003) others do not (Fellner et al, 2011). Similarly occurs with the fairness treatments. Torgler (2003) finds positive effects; Ortega et al (2012) does too, but only under some restrictive conditions (which makes the results significant for the group in only 1 out of 17 cities in which they performed the survey experiment).

Second, most of the studies have been written for developed countries, where tax evasion is low, enforcement is a regular activity of the tax administration, and trust in government is high. Pursuing these studies in developing countries would allow to test the value of the hypotheses in contexts in where tax evasion is rampant, governments have fewer resources for enforcing laws, and trusts in the institutions is lower. Third, while most studies have focused on income taxes (e.g. Blumenthal et al, 2001a, 2001b; Kleven et al, 2011), VAT (Pomeranz, 2010), or very specific fees (Fellner, Saugruber and Traxler, 2011), few have dealt with taxes that are directly linked to the provision of a public good. Hence, the possibility of testing for the relevance of public goods provision on compliance has been indirect at best.

Fourth, given that the judge is still out there regarding which –if any- of the policy options matter for affecting compliance, it would be important to evaluate them within the same experimental framework in order to ensure that differences in their effectiveness come from the policy itself and not the context in which they are applied. Finally, in order to increase the external (and internal) validity, studies would benefit from working on the overall population affected by a specific tax rather than on a selected sample of taxpayers. The experimental design of the experiment described in this chapter has attempted to deal with these issues.

2.3.The experiment

With the objective of raising compliance with the tax and evaluating the effectiveness of sending different messages to the taxpayers, the Municipality of Junín agreed to include three different messages in the tax bill through a randomized experimental design. These messages were expected to change taxpayers' beliefs regarding the equity, fairness, and enforcement of the tax. In terms of academic relevance, the experiment would provide further evidence regarding hypotheses that have shown elusive to corroborate, it may provide the basis for further theoretical elaboration, and test the external validity of previous research.

2.3.1. Background on the CVP in Junín

The "Public Space Conservation" tax (*Tasa de Conservación de la Via Pública*, or CVP henceforth) is a tax that is levied upon real estate. Taxed property includes homes, farms, business premises, and most other real estate. The tax is computed by taking into account the linear frontage, in meters, of the taxed real estate, the number of streetlights around the property, and the type of trash collection and street cleaning services it receives.[25] Property owners are billed every two months. Taxpayers have approximately 10 days to pay from the moment they receive the bill to the first due date. In case of payment delays, a cumulative compound monthly interest rate of 2% is applied to the outstanding liabilities.

By August 2011, there were around 26,000 individual taxpayers registered to pay the CVP, equivalent to a third of the population of Junín, according to the last 2010 Argentine census.[26] While the tax is levied every two months, the Municipality also allows taxpayers to pay it on a yearly or monthly basis. However, only around 12% of the taxpayers select either of these modalities. The large majority of the taxpayers (around 80%) pay their monies personally at the Municipality or other premises of the local government instead of using direct deposit or other automatic payment modalities.[27] Most taxpayers own one relatively small taxed property and only 20% of the taxpayers own more than one registered taxed property. Also, there are 2.7% of the registered taxpayers with postal address outside the municipality of Junín.[28]

Over the last four years, tax compliance, defined as total tax payments over tax liabilities, hovered around 40% percent on average in each bi-monthly billing cycle. Compliance exhibited a declining trend from 2010 onwards, only to show a gradual recovery in 2011. Moreover, payments display a marked seasonal behavior with larger payments made in January and February (bimester 1), mainly as a result of taxpayers paying in advance some their ensuing liabilities for the remaining of the year (see Figure 7).

Junín is divided in 15 districts and 10 sections (see maps in Appendix 3). Districts are larger and generally encompass more than 1 section. Tax compliance, public services' provision as well as the number and size of taxed real estate vary widely across districts as a reflection of heterogeneous characteristics of the underlying taxpayer population. As the capital city of the municipality, Junín is the largest district in terms of the number of taxpayers and taxed properties as well as the amount of taxes collected. Within the city of Junín itself, there is also a large heterogeneity in taxpayers´ observable characteristics across sections. While the remaining districts are largely rural, they also

[25] While there is available information on the fiscal valuation of the majority of the taxed real state for 2009, the data is not utilized for the computation of CVP´s liabilities as it is deemed unreliable and outdated by the local authorities. Hence, we don't use it in our baseline estimations.
[26] http://www.censo2010.indec.gov.ar/definitivos_bajarArchivo.asp?idc=93&arch=x
[27] Unfortunately, the municipality does not collect this data at the individual level.
[28] It includes, mostly, people with investment properties or second homes in the city.

include some small cities. As it can be observed in Table 12, Junín is quite similar (although slightly richer) to the average municipality in the Province of Buenos Aires (the most populated province in Argentina).

The units of analysis are individual taxpayers. We excluded private companies and social organizations from the sample because for those taxpayers somebody different than the person receiving the treatment may make the payment decision, which could bias the results.[29] Among the individual taxpayers of the CVP in Junín, the entire universe was included in the randomized field experiment with the exception of 3,000 individuals that had previously paid in advance their tax liabilities, and therefore were not billed in the period the experiment took place. For each taxpayer, we have collected the administrative information on tax liabilities, tax arrears, the number and size of taxed real estate (e.g. average linear front meters) and some public goods provided by the Municipality (e.g. streetlights, garbage collection, and street cleaning services). The definition of each one of these variables is provided in Table 13.

For selecting the taxpayers into each treatment, we followed a stratified or block randomization strategy due to the presence of highly correlated taxpayers´ observable characteristics at the section level.[30] Therefore, we use the geographical location of each taxpayer to define the strata or blocks. In the case of taxpayers listed with a mailing address from outside the Municipality, we created an additional artificial block as they may show highly idiosyncratic unobservable characteristics. Further, taxpayers registering multiple properties but lacking a unique mailing address were assigned to the section in which the property with the highest tax liability prior to the experiment was located. We also found that 4% of the taxpayers had multiple fiscal identifications (ID) numbers but the same exact name, surname and mailing address. Because despite consulting with the Municipality it was not possible to establish the cause of this situation, we decided to treat these "duplicated" taxpayers as single taxable units, to reduce potential assignment errors. Hence, a fictitious unique identification number was created and utilized in the randomization procedures for those taxpayers

Within each block, for each taxpayer randomly assigned to a treatment a taxpayer was assigned to the control group. Thus, because we have three treatments, sixty percent of the taxpayers were randomly assigned to the control group (about 15,000), and the remainder was equally distributed to each of the treatment groups within each strata (about 2,500 each). Even though this decision may have lowered power for each treatment, and it may work against finding statistically significant results, it would also reduce the possibility of cross-treatment contamination (which could

[29] For example, the bill may be received by a clerk who had been instructed in advance by the firm´s owner about whether to pay the bill or not. The decision then would have been unaffected by the treatment.

[30] For instance, the calculation of a simple measure of intra-section correlation yields a coefficient of 0.5 for trash collection and other locally provided public services, 0.25 for the number of streetlights and 0.2 for the average tax liabilities and the number of unrecoverable debtors.

have affected the ability of finding different results across treatments). Contamination of the control group instead would render the estimates as a lower bound. This decision of "undertreating" was also in line with the municipal government's preferences.

In order to ensure the right balance across groups, we ran the randomization 1000 times and selected the random draw that showed the best balance for all the pre-experimental covariates (including the pre-experimental outcome variable) controlling for the strata dummies. The final draw was selected according to two criteria: (a) the minimum statistical difference between treatment and control for draws statistically significant at a five percent or lower; and, (b) the minimum maximum t-stat (Bruhn and McKenzie, 2008).

After the randomization and random assignment of taxpayers, we also carried out a comprehensive set of actions aimed at minimizing potential administrative errors during implementation that could affect the randomized design of the experiment. First, we sent to the local authorities, three weeks before they started to print the tax bills, a file with a code assigning each taxpayer to either the treatment groups or the control. Based on that codification, the Municipality generated an administrative record that we compared against our files in order to detect and correct any mistakes with the randomization. Second, we also conducted, one week prior to distribution of the tax bills, a physical check with a random sample of 1,000 printed bills at the Municipality in Junín. No assignment errors were found in this final control.

The tax bills containing the treatment messages were sent in August 2011, corresponding to payments due in the September-October (Bim 5) billing period of the same year. That period might be considered "normal" as there are no pronounced seasonal variations in tax payments like in the January-February (Bim 1) billing period in which a large number of tax payments is made in advance for the rest of the year.

2.3.2. The treatments

We describe below the treatment messages that were included in the tax bill. The wording of the messages and the selection of the images was prepared by a communications team in coordination with the municipal authorities and tested in a small focus group. In every case, and following standard recommendations in the literature, the objective was to provide the combination of words and pictures that could have the maximum impact while conveying the message that was intended.

Deterrence

The first treatment included a message in the tax bill that provided taxpayers with a simplified example of the cost of non-compliance by computing how much the cost of a hypothetical debt of AR$ 1,000 would be after a year (given a cumulative monthly interest rate of 2%). Specifically, the message stated that, for a liability of such amount, the taxpayer would have to pay AR$ 269 in arrears

after that period of time. The message also warned the taxpayers that the Municipality would take administrative and legal steps in case of non-compliance. An image of a judicial hammer was also included in the tax bill with the intention of reinforcing the deterrence message. Table 14 provides the translated text of the messages and an example of a tax bill is included (in Spanish) in the Appendix 4.

The aim of the message was two-fold: (a) raising the taxpayers' subjective perception of the probability of receiving a fine and other possible administrative or legal penalties in the case of non-compliance, and; (b) reducing the likely computational costs derived from the calculation of arrears on unpaid tax liabilities using a compounded interest rate. The literature on tax compliance that incorporates the insights of behavioral economics points to the importance of limited computational capabilities, particularly in the case of relatively complex intertemporal tax calculations for explaining taxpayers´ responses to fines and other penalties (Congdon et al, 2011). This literature also suggests that, in the presence of taxpayers with limited attention, raising the salience of fines and legal actions in the case of non-compliance might have an effect on individual behavior with respect to tax payments (Bernheim and Rangel, 2007 and 2009).

Fairness

The second treatment introduced a message in the bill with information about the actual use of revenues by the Municipality. It highlights the number of streetlights, and water sewerage connections installed by the local government of Junín in the previous six months. The message was also accompanied by an image of an easily recognizable "men at work" traffic signal (Table 14 and sample tax bill in the Annex 3). This message was intended to modify taxpayer´s perceptions about the fairness of the tax by influencing their priors about the supply and quality of public services provided by the Municipality (Cowell and Gordon, 1996).

Equity

The third treatment introduced a message in the tax bill about the levels of CVP compliance. The message asked whether the taxpayer was aware that only three out of ten taxpayers did not pay their tax liabilities.[31] It also added a sentence questioning whether the taxpayer was currently paying her liabilities ("What about you?"), which attempted to capture the essence of the moral suasion arguments. In order to reinforce the message, the bill also contained an image with seven larger figures, personifying the tax payers who pay and do it on time, and three smaller ones, representing the people who don't do it on time (Table 14 and sample tax bill in the Appendix 4). The goal of the message was to influence taxpayers´ perceptions about the extent of evasion in the local community as suggested by Dell´Anno (2009).

[31] The message corresponds with the fact that 30 percent of the taxpayers are considered "unrecoverable debtors" (have basically never paid their tax bill). The other 70 percent have paid with some recurrence even if they had not done it in the last bimester before the experiment.

We are aware that even though extreme care has been taken in selecting the factual information and wording so the messages would tend to encourage people to update their beliefs upward (and reduce evasion), it is still possible that some people may have updated them in the opposite direction. For this reason, in the empirical work, we have taken special care in looking into potential heterogeneous effects. Of course, it may still be the case that those who usually do not pay are also skeptical of government messages, and may not update their beliefs at all regardless of the message they receive.

2.3.3. The dependent variables

We have defined the outcome variable in several alternative ways (but always dichotomously) in order to ensure we capture changes in behavior as precisely as possible. The main outcome variable, *paid*, takes the value 1 only if the taxpayer has paid in full the total tax liabilities for the period of the experiment. Two variables record the timeliness of the payment: *paid_by1D* takes the value of 1 if the payment took place before the 1st due date; *paid_by2D* takes the value of 1 if the payment took place before the 2nd due date. Because the taxpayer may decide to pay in addition some arrears and reduce its debt or to pay in advance future liabilities we created the variable *overpaid*, which takes the value 1 if the taxpayer has paid not only her liabilities for the period in full but she has also made some additional payments. For taxpayers with multiple properties, all these variables take value 1 if the taxpayer paid all of the bills in full.[32] Finally, the variable *addpayments* takes a value 1 when the taxpayer has made advanced payments or paid down part of her debt, regardless of whether they have paid their current liabilities in full.

Table 15 presents the summary statistics for the three treatment groups, compared to the control group in the pre-experimental Bim3 billing period. As expected, given the random assignment, average observable characteristics for each one of the groups is very similar.[33] Most importantly, there are no statistically significant differences among the groups in terms of the dependent variables. Table 15 indicates that previous to the experiment, around 21 percent of taxpayers paid before the 1st due date. This figure increased to 33 percent before the 2nd due date. By the end of the bimester, 40 percent of taxpayers had paid the liabilities for the period in full. Additionally, 3 percent of those taxpayers had made additional payments during the period, such as paying down their debt. In total, approximately 4.5 percent had made some kind of additional payments.

How do these percentages translate into money? Basically, while the average liability for the CVP amounts to AR$ 122, the mean tax payment is equivalent to AR$ 56. On average, the typical

[32] We have also used different cutoffs such as 60% of the taxed properties without finding any significant changes in the results.
[33] Only the variable monthly payments, for one of the treatment groups, is statistically significant at the 5% level.

individual taxpayer owns more than one property of 16 linear front meters, her trash is collected regularly, and has almost 3 street lights surrounding her property. Finally, we also include in the Table 15 the balance of randomization for the public services variables in the period in which we conducted the experiment to show that no public works affected the groups differentially.

2.3.4. Empirical strategy

To estimate the causal effects of the treatment messages on tax compliance, we employ a probit model. Formally, we estimate the following specification:

$$Prob(Y_i = 1|X) = \Phi\left(\alpha + \beta_1 T_{1i} + \beta_2 T_{2i} + \beta_3 T_{3i} + \gamma Z_i + \delta_{is}\right) \quad (1)$$

Where *Y* is the binary outcome variable equal to one if the individual taxpayer *i* meets her tax obligations in the bi-monthly billing period 5 (September-October) according to any of the definitions provided before; *T* are binary variables representing the three treatment messages (T_1=Deterrence; T_2=Fairness; and T_3=Equity), *Z* is a vector of control variables comprising taxpayer´s observable characteristics (e.g. income proxies, payment frequency and compliance status) and the amount and type of public services provided by the municipality (e.g. number of streetlights and trash collection and street cleaning services), and δ is a set of strata fixed effects. Following Duflo et al (2008) and Angrist and Pischke (2008), we include the lagged outcome variable as an additional control to avoid potential serial correlation concerns, mainly because there is high persistence on payment behavior.

2.4.Results

Table 16 presents the average treatment effects (ATE) of the probit estimation described in equation (1) –table shows marginal the effects computed at the mean.[34] Baseline estimations include the three treatment messages, the lagged outcome variable (because compliance shows great persistence), and strata fixed effects. The estimations with controls also include the variables for public service provision (trash collection and street lightning services during the period), the number of properties that each taxpayer has, the average linear front size of the properties, and a dummy that controls for those taxpayers who selected to pay monthly.[35]

The result that comes clearly out of the estimations is that the deterrence message had a positive, and statistically and economically significant effect on compliance while the other two messages had no significant average effect. The size of the effect differs according to the dependent variable used. As expected, the size of the effect is larger for our main variable of interest, *paid*, that considers compliance with the tax regardless of the date in which payments took place. In this case,

[34] Full regression tables are provided in the Appendix 4.

[35] Public services are included contemporaneously instead of lagged to control for any effect on payments that could have been caused by an increase or decrease of public services during the period of the experiment. Results are identical if we include these variables lagged instead. Also, results are basically identical if we run the regressions without including fixed effects as shown in the Tables in the Appendix 4.

the message increased compliance by almost 5 percentage points, which represents an increase in compliance rates of approximately 12 percent.

Results are smaller in magnitude for the other dependent variables but no less interesting in terms of the policy implications. The message increased compliance before the first due date by almost 2 percentage points, and compliance before the second due date by 3 percentage points. Moreover, it does not only had an effect on the compliance for current liabilities: taxpayers who received this message were also more likely to pay arrears and future liabilities in advance. These results confirm the hypotheses coming from the theoretical literature: increasing the salience of penalties affects compliance behavior.

Results remain statistically identical when we include the additional controls, which include proxies for taxpayers' observable characteristics such as the number of associated taxed properties and monthly payment frequency, as well as indicators of public services provided by the Municipality (the number of streetlights and trash collection/street cleaning services). These variables display the expected signs (as shown in Tables 3 to 6 in the Appendix 2). In addition to the persistence in payment behavior, those who paid in the past are more likely to keep paying in the present, people who receive more public services from the municipality do too. As suggested by recent theoretical work, the perception of an adequate provision of public goods with respect to tax payments or "fairness" might provide an incentive to comply with the tax code (Hashimzade et al, 2013). Still, this is evidence of a correlation and not necessarily of causation.

Taxpayers who pay monthly are also more likely to comply more. This may be explained by selection bias because people have opted for this type of billing. Hence, they may be more aware of their responsibilities with the municipality. On the other hand, taxpayers with multiple properties and/or larger properties tend to display lower compliance. One explanation may be that wealthier individuals are more likely to engage in riskier activities (Sandmo, 2005). Of course, it may also be that people with multiple properties may be more likely to misplace some of the bills, or have rented properties (and rely on the tenants for the payment).

So far, we have described the results for the marginal effects at the mean level. We have also checked for whether the size of the effect differs at different points in the distribution of the covariates. While the marginal effects tend to change, these differences tend be very small. As an example, and always using *paid* as the dependent variable, the effect of the deterrence message tends to decrease as the number of properties increase (from 0.049 for people with only one property – which includes the majority of the population- to 0.034 in the upper tail of the distribution) and as the size of the property increases (from 0.05 to 0.041). It increases, however, as the provision of trash collection improves (from 0.042 for those with no collection to 0.049 for those with the best service). Figure 1 in Appendix 2 illustrates these.

While the results for the deterrence message are large, we find no average behavioral changes for people who have received either of the other two messages. There are several potential explanations for this non-result. First, invoking moral reasons, despite evidence in the literature that it may switch perceptions at the margin (as Ortega et al, 2012 have found), it may not affect payment behavior because the perceptions elasticity for changing tax compliance is low (Blumenthal et al 2001b). This is not uncommon. For example, Fryer (2013) finds the same in an experiment in the education sector.[36]

Second, people who do not comply may have lower trust in the government, which may translate into lower effectiveness of the messages (those who don't trust the government may don't trust the messages they receive about the levels of tax evasion and public works). Finally, the average effects may be masking differences across individuals. For example, while we expected that people would tend to evaluate the information about actual use of the money by the local government in positive terms, people who had priors about government works that were "too high" may have revised them downwards instead of upwards. We also expect that people would react similarly regarding the rate of compliance of their neighbors.

We have checked for heterogeneous effects across the observable characteristics but we haven't been able to find any significant results for most of the variables of interest.[37] For example, we find no differential effect for the treatments neither across different levels of provision of public services by the municipality at the individual level nor by including a variable that measures the difference between the public goods they receive compared to the median public services provided in their strata. Therefore, we can't show evidence that people develop different priors according to the services they receive and have adjusted their behavior differently after receiving the message. This (non) result weakens some of the claims in the studies that emphasize public service provision at the individual level as a source of people's beliefs about public sector efficiency (and the fairness of the tax system). We do find some differences worth mentioning however.

First, we find that past payment behavior seems to affect the relevance and effect of the deterrence and equity messages. As it can be observed in Figure 1a in the Appendix 2, which summarizes the marginal effects for the interaction between the deterrence message and the lagged outcome variable, the message has a positive impact on those taxpayers who had not paid in the previous period but it has no statistically significant effect for those who had. The opposite occurs regarding the equity message as shown in Figure 1b in the Appendix 2. While the effect of the message is zero for those who had no complied before, it is negative for those who had. That is, the

[36] A moral suasion message changed reported student's beliefs about the relationship between education and outcomes. They also reported being more focused and working harder. However, there were no measurable changes in attendance, behavioral incidents, or test scores.

[37] Regression results for the models with interactions are presented in Table 7 in the Appendix 2. The figures summarize the marginal effects of the interactions.

message seems to have had a disincentive effect on those who had complied in the past and had been overstating other people's rates of compliance. There is of course, an alternative explanation. The fact that the government was advertising the degree of evasion may lead some people to believe that enforcement is lax and nothing would happen if they evade too.

The same results can be observed if we use, instead of the lagged outcome variable, a variable that classifies taxpayers according to whether they have debts with the municipality or not. Again, only those with debts seem to be affected by the deterrence message (Figure 9a) and those with no debt tend to present a disincentive effect when handed the equity message (Figure 9b). Interestingly, not everyone with arrears behaves in the same way when presented with the equity message. When we incorporate a variable that takes into account the stock of debt of each taxpayer, we find that while people with zero or very small levels of accumulated debt tend to present the mentioned disincentive effect, the message has a positive effect on those with higher levels of debt (Figure 10). This result may imply that those who have rarely complied may have adapted their beliefs upward regarding compliance levels when confronted with the message.

Results seem to differ too according to property sizes (Figure 11). The deterrence message seems to have a positive effect particularly on people with smaller property sizes (Figure 11a). Potential explanations include the existence of differences in risk aversion across income levels, higher sophistication in interpreting current penalty levels and enforcement abilities of the municipality, or the fact that penalties and interests have a lower relative burden as income increases. Additionally, while past payment behavior had no effect on the fairness message, there seems to be a differential effect however, when the treatment is interacted with the average size of the properties. The message seems to have a positive effect for people with smaller properties but a negative effect for people with larger ones (Figure 11b). One potential interpretation would be that expectations about the use of public monies by the municipality differ according to property size, and those with larger properties have a higher expectation about what the government should do. Hence, presented with actual information, they adapted their priors and their behavior. This result does not change when controlling for having one or more properties.

A final check on heterogeneous effects and formation of beliefs has been to run the regressions separately for the group of taxpayers who live outside the city given that they may hold different priors than the rest of the population. Interestingly, as a first approximation to the potential set of priors of this group of taxpayers, compliance is 25 percent lower than the average compliance in the city. The results, presented in Table 17, show that both, the deterrence and the fairness message tend to be positive (and larger than the results for the overall population). This would tend to indicate that people outside the city might have reacted more to the information about public services than people in the city, who are direct witnesses of the effort by the Municipality.

A potential concern with the results is that some of the taxpayers may have not read the messages, and therefore were not treated, introducing a downward bias in our estimates. In particular, there is a group, the "unrecoverable debtors", defined as taxpayers that did not registered any payments on their CVP liabilities between 2007 and 2011, who are more likely to ignore the messages included in the bill.[38] To investigate this potential concern, we estimate in Table 18 the same empirical model as before but excluding the taxpayers that were unrecoverable debtors by the Bim 3, 2011 from the sample. The estimations show a slightly larger effect of the deterrence message treatment on tax compliance for some of the dependent variables, suggesting that we could be in fact underestimating its impact due to the presence of a potential intended-to-treat (ITT) downward bias (Dufflo et al 2008; Angrist and Prischke, 2008). Table 19 shows that this change in the magnitude of the deterrence message coefficient is not driven by a potential bias in the underlining distribution of taxpayers amongst treatments and control groups. The balancing between the treatment messages and the control group remains virtually unchanged when unrecoverable debtors are excluded from the sample.

To provide some intuition for the findings in terms of revenues collected instead of individual level decisions, Figure 12 plots the tax compliance ratio –computed as tax payments over tax liabilities-, for the three treatment groups and the control group in Bim 5. It also displays the tax compliance ratio for the period when randomization was conducted (Bim 3 2011) as a benchmark as well as for the same billing period but one year before (Bim 5, 2010). This figure suggests that tax compliance for taxpayers that received the deterrence message was 3 percentage points higher than for taxpayers in the control group (no message). This difference is statistically significant according to a nonparametric test of proportions. Notice that tax compliance increases for both the treatment and the control groups with respect to the previous billing period and the same period a year before. A potential explanation is that the introduction of a simplified tax bill with an improved design by the Municipal authorities in the same bimonthly billing period than the experiment was conducted provided an incentive for this generalized increase in tax compliance across the treatment groups and the control.[39] It is noteworthy that this effect does not affect the validity of our results as the same re-designed bill was distributed to every taxpayer.

[38] We have tried with other proxies too, such as identifying potential "untreated" subjects according to different number of unpaid bills. The results show a non-linear relationship. As the definition of "unrecoverable debtor" is broaden then fewer non-compliant taxpayers who may be affected by the treatment are included in the sample, which reduces the size of the coefficients.

[39] It may also be the case (even though we have no evidence) that some of the people in the control group may have learnt about the existence of the deterrence (or the other) messages from some of their neighbors; which may have ended up affecting our estimates downward. Therefore, we can't rule out that the increase in compliance in the control group was affected by enforcement spillovers. Rincke and Traxler (2011) find sizeable spillover effects of actual enforcement –one additional household compliance for every three additional units of enforcement.

Summarizing, the evidence indicates that increasing the salience of enforcement has a positive effect on compliance. Messages that appeal to equity and fairness considerations do not necessarily show the same average effects. Still, information about public goods provision and the compliance behavior of other taxpayers seems to matter in some margins. While it is not necessarily clear that it would encourage compliance, under certain conditions it may discourage it.

Unfortunately, we could not evaluate the persistence of the effects over time because the Municipality launched a moratorium during the following bimester. These types of payment plans facilitate the payment of arrears, and reduce outstanding liabilities (in particular they pardon some of the interests and penalties). Additionally, as mentioned previously, many people who received the messages decided to pay some of their future liabilities in advance, which could skew the results for the following bimester.[40]

2.5.Conclusions and further research directions

This chapter explores empirically whether providing information to taxpayers influences or not the individual taxpayer´s compliance decision. Specifically, we conducted a large randomized field experiment to test whether including messages in the tax bill about the levels of fairness, equity and enforcement of the tax would modify taxpayers´ behavior by affecting their beliefs about the use of public monies by the municipality, the prevailing level of compliance, and the severity of enforcement. This way, we attempted to provide evidence about the relative importance of these complementary policy instruments.

For conducting the experiment, we randomized (within 25 geographical blocks) the entire universe of individual taxpayers of a property-based tax of a municipality in Argentina into 4 groups. Extending the experiment to the universe reduces problems of sample selection and increases the external validity of the experiment. Having conducted it in an average city of a developing country complements existing studies by testing the hypotheses in a context in which tax enforcement and trust in governments is lower, and tax evasion is rampant. Using a local property-based tax allows to test hypotheses about fairness more directly than using direct taxes collected by the federal government.

Results indicate that the marginal probability or impact effect on tax payments for the group of taxpayers that received a deterrence message was on average around 4 percentage points higher than for the control group of taxpayers that did not received any message. To put these findings in context, extending the use of this message to the entire population of taxpayers may have resulted in a sizeable increase in revenues equivalent to more than 10%. Moreover, the empirical results may be understating the actual impact of such a policy, as some people who have not being paying the tax

[40] More than 700 people made advance payments in the period. This is more than 30 percent higher than in the pre-experimental bimester.

may have not read the message. Additionally, once people have accumulated debt, the incentive to comply in the current period does also fall.

We find no average treatment results for the messages that were designed to affect the beliefs about the equity and fairness of the taxes. The reasons behind this are multiple. First, even if messages affect perceptions these may not translate into changes in behavior. Second, lower trust in government may translate into lower effectiveness of the messages (those who don't pay taxes because they don't trust the government may also do not trust the messages about levels of tax evasion and public works). Third, the average effects may be masking differences across individuals. Some of these differences can be uncovered with observable data. In fact, we find that past payment behavior (measure as either having paid in the previous period o having debts) tend to have an effect on how people react to the messages. In particular, those who had complied in the past tend to react negatively to the information about other people's compliance levels.

While our results provide novel empirical evidence about the effect of influencing taxpayers´ perceptions for affecting tax compliance, more in depth research is needed to understand the causal channels through which these effects operate. In particular, it would be relevant to disentangle the relative effect of the information about how to calculate fines on unpaid liabilities that was included in the deterrence message from the cautioning messages about the consequences of non-compliance and the images that were also included in the treatment. This would allow assessing what is more important: addressing the potential limited computational capabilities of the taxpayers or raising the salience of the monetary and legal consequences of non-compliance. Additionally, it may be important to disentangle whether the fairness message is having an effect (when it does) because of social norms or because it provides indirect information about stringency of enforcement by the government.

Another research avenue would be to concentrate further on using this technique for evaluating the determinants of the individual taxpayer's compliance decision in the standard model of tax evasion by concentrating on the penalty versus detection relationship (Are they substitutes? Complementary?). Chapter 3 attempts to provide some evidence in that direction based on the results of a field experiment in another municipality of Argentina, 9 de Julio.

Finally, it may also be worth running experiments like this in contexts where additional information is available at the individual level, such as political preferences, to study if the moral suasion messages affect different people differently, particularly for those messages that appeal to moral arguments or could be read through a particular political lens. Moreover, combining this type of experiments with large-scale surveys could provide a measure of the perceptions-elasticity of compliance.

Beyond the academic interest, this chapter presents additional evidence about the important role of managing adequately the many opportunities that policymakers have to influence citizens' beliefs and how this cheap policy alternative could yield substantive benefits to the government. In the case of taxation, and the fight against tax evasion, this policy alternative minimizes administrative costs, a feature usually ignored by the literature (Sandmo, 2005).

CHAPTER 3: To Monitor or to Punish? Analysis of the Effects of Informational Treatments on Tax Compliance in a Large Field Experiment

3.1 Introduction

What are the determinants of the individual´s decision to comply with her tax obligations? Economists have been attempting to answer this question since Adam Smith (1776) stated his *"four maxims of taxation"*. Yet, only in the 1970s, with the adaptation of the classic Becker (1969) ´s paper on crime to the economics of tax evasion by the seminal work of Allinghan and Sandmo (1972) and Yitzhaki (1974), economists were able to model tax compliance within the framework of the traditional theory of individual´s decision-making under uncertainty. Allinghman and Sandmo (1972) and Yitzhaki (1974) (AS-Y henceforth), portray the individual taxpayer´s decision about whether paying or not taxes as a gamble between the monetary benefits of evading and the monetary costs faced in case of detection. In their framework, the level of tax evasion depends on the probability of being caught and the resulting fines and legal penalties associated to non-compliance.

As it was mentioned in Chapter 2, analyses of actual compliance and monitoring efforts by the government have shown that compliance is usually higher than what the standard AS-Y model would predict. Part of the explanation lies on the taxpayers' intrinsic motivations to pay taxes or tax morale (Torgler, 2003). Part of the explanation has to do with the individual taxpayer's perceptions about the salience or prominence of taxes and fines and the probability of detection and enforcement that influence heavily her behavioral responses to the tax code (Chetty et at, 2009; Congdon et al, 2011). Those perceptions are not static.

For instance, recent theoretical work based on non-expected utility models points to the importance of subjective perceptions about the possibility of being caught and the level of fines in determining the tax compliance decision of the individual taxpayer (Hashimzade, Myles and Tran-Nam, 2013). The results of laboratory tax experiments reviewed in Alm (2010) also show that individuals generally make poor predictions of the probability of detection and the magnitude of fines and other penalities. In turn, a growing literature provides evidence stemming from field experiments that taxpayers' beliefs about the monitoring efforts of the tax authority, and consequently, their subjective perceptions about the probability of detection and enforcement in the case of non-compliance, are significantly influenced by informational flows (Schawarz and Orlean, 1967; Alm et al, 1993; Blumenthal et al, 2001; Kleven et al, 2011; Pomeranz, 2010; Fellner et al, 2011; Sanjit and al-Nowaihi, 2007and 2010).

Following this strand of the literature, in this chapter we attempt to influence taxpayers for increasing compliance by sending alternative messages about the level of fines and other penalties and monitoring and enforcement efforts by the government. By carrying out a large randomized field

experiment (more than 10 thousand individual taxpayers) in a municipality of Argentina, 9 de Julio, we can test the absolute and relative importance of information about the two traditional determinants of compliance in the standard AS-Y model of tax evasion (the level of fines and monitoring and enforcement efforts of the tax authority) on taxpayers' behavior.

As a motivation for the field experiment, we sketch a simple empirical model, adapted from the literature on signaling games and electoral outcomes (Kendall et al, 2013), to estimate the effects of the informational treatments on the probability of complying with the tax code of the individual taxpayer. Our results indicate that, relative to a control group that received a placebo message, the treatment that contained deterrence-based information was the most effective at influencing a higher level of tax compliance. More precisely, taxpayers that received that message increased tax compliance by around 3 percentage points with respect to the control group.

Similar results were obtained when the taxpayer received, in addition to the deterrence message, a letter indicating the intensified monitoring efforts of compliance with the tax code, and hence, suggesting an increased probability of detection of tax cheating. Further, we find even stronger effects for relatively wealthier taxpayers that were not complaint with tax code in the pre-experimental period. Tax compliance for those taxpayers that received the deterrence-based message was 6 percentage points higher than for the control group. The monitoring letter, in turn, had a slightly larger and more robustly estimated effect.

This chapter contributes and relates to several threads of the literature. First, it provides novel empirical evidence to the growing body of studies about the role of information at shaping individual choice with respect to tax policies (Bartolome, 1995; Chetty et at, 2009; Chetty and Saez, 2010; Jones, 2010). This emerging set of studies suggests that policies that provide information about tax incentives should help individuals to make better choices (Chetty et al, 2010).

This chapter also contributes to a specific literature that provides experimental evidence on the importance of information flows for influencing the individual taxpayer´s compliance decision (e.g., Schawarz and Orlean, 1967; Alm et al, 1993; Blumenthal et al, 2001; Kleven et al, 2011; Traxler, 2010; Pomeranz, 2010; Fellner et al, 2011; Sanjit and al-Nowaihi, 2007and 2010) by implementing an empirical test of the effects of informational treatments on the traditional determinants of tax compliance in the standard tax evasion model. First, we provide evidence on the potential presence of substitution effects between deterrence-based and monitoring-based information in influencing tax compliance. While since the original AS-Y model of tax compliance the theoretical literature has leaned towards treating these instruments as perfect substitutes, there is growing body of evidence that suggest that the magnitude, and possibly the sign, of the compliance effects of these informational treatments could be different (See Section 3.2 for a review of the existing evidence). Second, we improve the existing methods for estimating the effects of deterrence and monitoring-based

informational treatments on tax compliance, first, by mitigating (at least partially) concerns about potential intended-to-threat (ITT) effects, and second, by sending the information to the entire universe of taxpayers instead to a selected sample of taxpayers as in previous studies by applying the empirical methods developed in Chapter 2.

The remainder of the Chapter is structured as follows. Section 3.2 briefly reviews the existing evidence on the effects of deterrence and detection information on the compliance decision of the taxpayer. Section 3.3 describes the field experiment and our main results as well as some additional empirical analyses. Section 3.4 presents our main findings. Section 3.5, concludes and discusses some potential avenues for future research.

3.2 Existing evidence

This section presents a brief review of the existing body of evidence about the influence of information on fines and other penalties and the probability of detection on compliance with the tax code. The review is not exhaustive but it is restricted to those empirical analyses that consider only these two traditional determinants of tax compliance put forward by the standard AS-Y model. Thus, other hypotheses tested in the literature, particularly related to the use of informational treatments related to the concept of tax morale, are not covered here[41].

The early evidence

An initial set of studies employed semi-experimental methods using instrumental variables to investigate the potential effects on tax compliance of raising the level of fines and the probability of monitoring and enforcement. Using income tax data for the United States, Witte and Woobry (1985) and Dubin and Wilde (1988) found a positive correlation between the risk of audit and compliance with the tax code. However, Pommerehn and Frey (1992), in a study of the Swiss cantons, found that, while intensified audit threats increases compliance, there are no significative effects of a higher level of fines. A limitation of these earlier studies is that they do not satisfactorily answer the methodological concerns stemming from the endogenous relationship between tax payments and income levels[42].

Alm (2010) reviews, in turn, the evidence stemming from laboratory experiments on tax compliance. He finds that nearly all experiments have found that a greater probability of detection leads to more compliance, although the deterrence effect weakens with higher audit rates. There is also experimental evidence that many individuals tend to overweight the probability of detection and enforcement, and hence there is more compliance than is predicted by the standard AS-Y model. This experimental evidence also suggests that audit rates have a greater deterrent effect than fines despite

[41] Section 2.2. in Chapter 2 provides a succinct review of the literature on tax morale.
[42] Andreoni et al (1998) presents a through critique of the semi-experimental literature on tax compliance.

their theoretical equivalence in the AS-Y model. A common problem with these experimental tests is that they are buttressed by the typical shortcomings of laboratory experiments related the difficulties to generalize their findings and the artificiality of the control group[43].

Evidence from field experiments

Partly as a response to the limitations of the experimental and the semi-experimental literatures, another set of studies employs field experimental methods to test the effects of deterrence-based information and monitoring-related messages. In a pioneering study, Schawarz and Orleans (1967) conducted a field experiment to test the impact of messages that made salient sanctions in case on non-compliance relative to the effects of messages with moral appeals. They found no average effects of the deterrence-based message on compliance. A problem with this study is the difficulty to identify meaningful average treatment effects due the enormous variability of the dependent variable utilized (e.g. monetary tax returns).

In the same vein, Blumenthal et al (2001) informed by letter a group of 1724 randomly selected taxpayers in Minnesota that the income tax returns they were about to file would be closely monitored. Compared to a control group that did not receive this letter, low and middle-income taxpayers in the treatment groups on average increased tax payments compared to the previous year. However, the audit threat exerted the opposite effect on high income taxpayers, who reduced tax payments.

Wenzel and Taylor (2004) carried out a field experiment in Australia to evaluate the influence on compliance of letters with warnings about the potential application of sanctions in case of inaccurate reporting of tax deductible expenses related to rental properties. They find no significative effects on the magnitude of the claims. A shortcoming of this study is that it does not separately test the potential effects of deterrence-based information (e.g. the warnings about the likely sanctions) and the possible impacts of influencing taxpayer´s subjective beliefs about the probability of detection and enforcement of the tax code.

In turn, Kleven et al (2011) conducted a two-stage randomized field experiment over a representative sample of 40,000 taxpayers in Denmark. In the first stage, half of the population was randomly selected to be audited, while the rest was deliberately not audited. In the second stage, they sent threat-of-audit letters to a random sample in the two groups. Both the audits and the letters had a significant positive effect on self-reported income.

Pomeranz (2010) analyzed how information influences tax enforcement for the Value Added Tax (VAT) through two randomized controlled field experiments with over 445,000 Chilean firms.

[43] See Andreoni et al (1998) for a complete review of the limitations of the tax compliance laboratory experiments.

She found in the first experiment that the impact of a random audit announcement transmitted up the VAT chain increases compliance by firms' suppliers. The second experiment finds that papers trails acts as a substitute to firms own audit risk. Messages announcing intensified tax enforcement efforts had smaller effects.

Finally, Fellner et al (2011) carried out a large field experiment with potential evaders of TV license fees in Austria sending different messages to taxpayers. They find a strong effect on compliance, particularly for a treatment aimed at raising the salience of detection risks. As mentioned in Chapter 2, a potential limitation of this study is that included information about fines and other penalties in every treatment impeding an adequate identification of the effects of the deterrence information.

In the sum, the evidence stemming from empirical work in not conclusive about the effects of deterrence and monitoring-related informational flows on tax compliance. First, the existing studies do not provide definitive evidence on whether deterrence-based messages related to the salience or prominence of potential penalties or information about intensified monitoring and enforcement efforts are more effective at inducing compliant behavior. Second, the judge is still out there with respect to the magnitude and the sign of the effect of messages containing deterrence-based information. For instance, Wenzel and Taylor (2004), Schwartz and Orleans (1967) and Slemrod et al (2001) find mixed effects for different groups of taxpayers. Third, the existing body of evidence does not empirically explore the potential presence of substitution or complementary effects between these two policy instruments. Since the original AS-Y model the majority of the studies have threated deterrence-based and monitoring-based treatments as perfect substitutes, without providing a rigorous test for the potential presence of these substitution effects. The experimental design of the experiment described in this chapter has attempted to address some of these shortcomings of the existing literature.

3.3 The Experiment

The experiment was conducted in 9 de Julio, a largely rural and middle-income municipality located in the North-center of Argentina´s main province, Buenos Aires (see Map 4 in Appendix 3). Compared with average municipality in Buenos Aires, 9 de Julio is slightly less populated and urbanized, more educated and wealthier (Table 20).

3.3.1 Background on the TSRU in 9 de Julio

The main source of own revenues for the municipality collected from the individual taxpayers is the Urban Service Retributive Tax (*Tasa Retributiva de Servicios Urbanos* or henceforth TRSU by its Spanish acronym). The TRSU is a local tax levied upon the real state, and computed according to the surface and the services provided by the municipality of 9 de Julio. Taxed property includes homes,

farms, business premises, and most other real estate. The tax is computed taking into account the linear frontage, in meters, and the surface, in square meters, of the taxed real estate, the number of streetlights around the property, the number of paved roads around the property, and the type of street cleaning services it receives. Property owners are billed monthly. Taxpayers have approximately 15 days to pay from the moment they receive the bill to the first due date. In case of payment delays, a cumulative compound monthly interest rate of 2% is applied to the outstanding liabilities.

Over the last two years, tax compliance, defined as total tax payments over tax liabilities, was around 65% percent on average in each monthly billing cycle. Compliance has exhibited an increasing trend from 2010 onwards, and there is no marked seasonal behavior in the collection of the TRSU (See Figure 13).

As in the Chapter 2, the unit of analysis is individual taxpayers. For each individual taxpayer we have collected administrative data on tax payments, dues, arrears, number of properties taxed, and public services provided by the Municipality for the period 2010-2012. The definition of each one of these variables is provided in Table 21. By August 2011, there were around 15,842 taxed units. The number of individual taxpayers is slightly higher than 10 thousand because a number of taxpayers have multiple properties.[44] This number is equivalent to about a fifth of the population of 9 de Julio, according to the last 2010 Argentine census.[45]

We excluded from the sample private firms, non-governmental organizations, and other legal entities paying the tax.[46] We also excluded individual taxpayers that pay their dues on a bi-monthly basis.[47] Taxpayers with a mailing address located in sections other than the capital of the district, that it is also called 9 de Julio, were also excluded. In these sections, the municipal government distributes the tax bill to local offices and not personally to every taxpayer. Consequently, individual taxpayers are required to physically attend these official premises to pick up their bill. Unfortunately, the municipality does not keep records about whether people collect their bills or not. Therefore, we could have not monitored which taxpayers had actually received the treatments.

[44] The municipality does not utilize an individual identification (ID) system for the TSRU (or for any of the local taxes). Instead, owners of each property are identified according to name and address. To circumvent this limitation, we applied a text recognition algorithm to the administrative data to identify individual taxpayers with the same name, surname and address. Then, we generated a fictitious tax ID for each TSRM's individual taxpayer identified by the algorithm. Using this artificial tax ID, the population of 15,842 taxpayers originally encountered in the municipality's administrative records was reduced to the around 10,300 unique individual taxpayers that were finally included in the experiment.

[45] http://www.censo2010.indec.gov.ar/definitivos_bajarArchivo.asp?idc=119&arch=x

[46] For example, the bill may be received by a clerk who had been instructed in advance by the firm´s owner about whether to pay the bill or not.

[47] While the tax is levied over a monthly billing period, the Municipality also allows taxpayers to pay it on a yearly or bi-monthly basis. However, only around 14% of the taxpayers select either modality. They were excluded because the experiment was carried out in a month (August) in which bi-monthly and annual bills were not distributed.

3.3.2 The treatments

With the objective of raising compliance with the tax and testing the relative importance of the determinants of tax evasion, the Municipality of 9 de Julio agreed to send a message and a notification letter attached to the tax bill of the TRSU through a randomized control trial (RCT). The causal mechanism was to change taxpayers' beliefs regarding the level of the penalties for evasion and the monitoring efforts of compliance with the tax code. As in the field experiment presented in Chapter 2, the wording of the message and the letter and the selection of the image was prepared by a communications team in conversations with the municipal authorities. It was also influenced by the experience collected in a previous experiment through the use of a focus group as well as in the field experiment presented in Chapter 2. We describe below the treatment message and the letter as well as the placebo message.

Deterrence message

The first treatment, a message attached on top of the tax bill similar to the one used in the experiment described in Chapter 2, provided taxpayers with a simplified example about the costs of non-compliance, which indicated how much the cost a hypothetical debt of AR$ 1,000 would be after a year (given a cumulative monthly interest rate of 2%). Specifically, the message stated that, for a liability of such amount, the taxpayer would have to pay AR$ 269 in arrears. The message also warned the taxpayers that the municipality would take administrative actions in case of non-compliance.

An image of a judicial hammer was also included in the message attached to the tax bill with the intention of reinforcing the deterrence message. Image 4 in the Annex 2 provides a sample of a message (in Spanish) and the text translated to English. The message was also signed by the Undersecretary of Revenue and Control of the municipality with the purpose of personalizing the message.

The aim of the message was raising the salience of the consequences of non-compliance by simplifying the calculation of arrears and unpaid tax liabilities using the compounded interest rate, and highlighting the probability of administrative and legal sanctions for non-compliers. As mentioned above in Chapter 2, the literature on tax compliance that incorporates the insights of behavioral economics points to the importance of limited computational capabilities, particularly in face of relatively complex intertemporal tax calculations for explaining taxpayers´ responses to fines and other penalties. This literature also suggests that, in the presence of taxpayers with limited attention, raising the salience of fines and legal actions in the case of non-compliance might have an effect on individual behavior with respect to tax payments (Congdon et al, 2011; Bernheim and Rangel, 2007 and 2009).

Monitoring letter

The second treatment consisted of a notification letter addressed to every individual taxpayer. The letter was designed following some of the recommendations commonly found in the behavioral economics literature: (a) the text was written in a clear and concise manner with the key ideas included in the first paragraph; (b) the language used was personalized and colloquial; (b) it was address to the taxpayer by her name and surname; (c) it was signed by and included the complete name of the Undersecretary of Public Revenues and Control; (d) it also included an image placed on the right of the written text; and, (e) it used bold letters to highlight some key words and phrases (Kaheman, 2011).

In the letter's header, a title with a text in caps and bold letters was also included aimed at emphasizing the local government's intensified monitoring efforts on TRSU's compliance. The key messages were also highlighted in the main text of the letter using also bold letters with the intention to attract taxpayer's attention. An image of a magnifying glass was also included in the letter with the purpose of reinforcing the monitoring message. Image 5 in the Annex 2 provides an example of a notification letter (in Spanish) and the text of the message in English. Like the deterrence message, the monitoring letter was also signed by the Undersecretary of Revenues and Control.

The goal of the notification letter was to influence the individual taxpayers' subjective perceptions about the monitoring efforts of the municipality on the TRSU's compliance, and therefore, her subjective beliefs about the probability of enforcement in the case of non-compliance. As it was already mentioned in the introduction of this chapter and in Section 3.2 above, a growing literature provides evidence on the effects of informational flows on influencing taxpayer's priors about the probability of being detected when tax cheating (Blumenthal et al, 1995; Kleven et al, 2011; Pomeranz, 2010; Fellner et al, 2011).

Interaction treatment

Taxpayers in this treatment group received simultaneously the deterrence message in the tax bill and the notification letter with the monitoring message. The aim of this treatment was to test the impact of simultaneously influencing the taxpayer's subjective perceptions about the consequences of non-compliance in terms of fines and other penalties, and intensified monitoring efforts on behalf of the local tax authorities, and the resulting increased probability of detection. The combination of treatments could allow us to evaluate the complementarity or substitutability of the determinants of tax evasion.

Placebo message

In order to make sure that any potential effects would come from the contents of the messages and not from receiving a message, the control group of taxpayers received a placebo message attached to the

tax bill. The message urged 9 de Julio's taxpayers to collaborate in maintaining the municipality clean and tidy. This message was the same that the one the municipality had been sending for the last four years to the entire population of TRSU's taxpayers. Image 6 in the Annex 2 provides a sample of a message (in Spanish) and the text translated to English.

3.3.3 Implementation and Summary Statistics

As in Chapter 2 (see Section 2.3), we implemented a stratified or block randomization strategy due to the presence of highly correlated observable characteristics for some groups of taxpayers[48]. Specifically, we use the number of properties (a proxy for taxpayer's income) and the outcome variable prior to the experiment to define four strata or blocks: (1) non-compliant taxpayers with one taxed property; (2) non-compliant taxpayers with multiple (more than one) taxed property; (3) compliant taxpayers with one taxed property; and, (4) compliant taxpayers with multiple taxed properties.

Within each block, the same proportion of taxpayers was randomly assigned to either the treatment groups or the control group. Thus, forty percent of the taxpayers were randomly assigned to the control group, and the remaining sixty percent was equally distributed to each of the treatment groups within each strata. The exclusion of the districts outside the capital city of the municipality from the experiment due to assignment concerns precluded us from using the geographical location of the taxpayers to define the blocks (See Section 2.3.1).

In order to ensure the right balance across groups, and following the same balancing strategy pursued in Chapter 2 (see Section 2.3), we ran the randomization 1000 times and selected the random draw that showed the best balance for all the pre-experimental covariates (including the pre-experimental outcome variable) controlling for the strata dummies. The final draw was selected according to two criteria: (a) the minimum statistical difference between treatment and control for draws statistically significant at a five percent or lower; and, (b) the minimum maximum t-stat (Bruhn and McKenzie, 2008).

The tax bills with the treatment message and the notification letter were sent in July 2012, affecting payments due in the August billing period of the same year. Table 23 presents summary statistics for the three treatment groups, compared to the control group in the pre-experimental August 2011-January 2012 billing period. As expected, given the randomized design of the experiment, average characteristics in the groups assigned to the different treatments look very similar: none of the differences between treatment and control groups are statistically significant at the 5% level. Most

[48] For instance, the calculation of a simple measure of intra-section correlation yields a coefficient of 0.5 for the one-month-lagged outcome (the taxpayer paid 100% or more of her tax liabilities) and 0.3 for the number of taxed properties.

importantly, there are no statistically significant differences among the groups in terms of the dependent variable.

Table 23 also indicates that previous to the experiment, around 65 percent of taxpayers had paid the liabilities for the period in full and made additional payments during the period, such as advanced payments or paying down their debt. Also, the typical individual taxpayer owned more than one property of 16 linear front meters. Finally, we also include in the Table 23 the balance of randomization for the public services variables in the period in which we conducted the experiment (August, 2012) to show that no public works affected the groups differentially.

As in the case of the experiment presented in Chapter 2, we carried a set of actions aimed at ensuring an appropriate implementation of the randomized design of the experiment. First, we sent to the local authorities, three weeks before the distribution of the tax bills, a file with a code assigning each taxpayer to either the treatment groups or the control. Based on that codification, the Municipality generated an administrative record that we compared against our files in order to detect and correct any mistakes with randomization. Second, we also conducted, one week before the bills were finally distributed, a physical check with a random sample of 1,000 printed bills at the Municipality in 9 de Julio. No assignment errors were found in this final control. Finally, we randomly assigned the tax bills with the messages to the municipal officials in charge of the distribution.

3.3.4 Empirical model

In this section, we present an empirical model of the effects of the informational treatments on tax compliance adapted from recent work on signalling games and electoral outcomes (Kendall et at, 2013). Consider an individual taxpayer *i* for which her expected utility is defined by her income $EU_i(Y_i)$. There is also a tax liability *T* and a fine *d* that is calculated as a percentage of the tax liability. Therefore, the penalty *F* the taxpayer pays in case of non-compliance is $T(1+d)$. Taxpayers have only imperfect information about the fines and other related penalties as well as the monitoring efforts of the tax authority *M*, and hence on the probability of detection *p* of tax cheating and enforcement of the tax code.

The experimental treatments described in Section 4.4. induce a randomized change in the information available to taxpayers about the salience or prominence of penalties and the subjective probability of detection and enforcement. Therefore, the randomized strategy creates four groups of taxpayers according to the informational treatment they receive:

Deterrence: Receiving a message about *F* (the penalty in case of non-compliance)

Monitoring: Receiving a message (a letter) about *C* (the monitoring efforts of the tax authority), and hence the probability of detection and enforcement of the tax code *p*.

Deterrence + Monitoring: Receiving a message about both *F* and *C*

Placebo: Receiving a placebo message not related to either *F* or *C*.

A message denotes having received a randomly assigned treatment with information on penalties, monitoring, penalties and monitoring or a placebo message, $M\{1,\ldots,4\}$. Therefore, we can define $T(m)$ as the group of taxpayers treated with information type $M=m$ and $b^i_{F,p}$ $(F, p|M=m)$ as the post-experimental subjective beliefs of taxpayer *i* treated with message *m* about the consequences of non-compliance *F* and the probability of detection of non-compliant behavior *p*.

Similarly to Chetty et al (2009), the taxpayer *i* only observes the tax liability *T* when she decides to comply with the tax code. The penalty *F* is computed by the tax authority once the taxpayer already evades her tax obligations and is detected (or alternatively decides to comply for any other reasons). Therefore, the message *m* can be seen as informational treatment aimed at raising the salience or prominence of fines and other related penalties.

We further assume that the potential outcome of taxpayers (e.g. paying their tax liabilities) that received the placebo message ($M=4$) is unaffected by the treatment assignment of the rest of the taxpayer's population. Hence, the post-experimental subjective beliefs in this group of taxpayers about penalties and the probability of detection and enforcement of the tax code are identical to their post-experimental beliefs (see Kendall et al, 2013).

The taxpayer's problem is dichotomous: whether or not to comply with the tax code. If the taxpayer complies with her tax obligations, her utility equals her gross income minus the tax liability or *Y-T*. On the contrary, if she decides to evade her liabilities there are two potential states of the world in terms of her expected income utility. First, the expected income utility of ***successful or undetected evasion*** *or* EU_{SE} for taxpayer *i*, treated by *m* is:

$$EU^i_{SE}\big(Y,T,F(m),p(m)\big)=\sum_F\sum_p \quad b^i_{F,p}(F,p|M=m)U[(1-p)(m)(Y_i)]$$

Instead, if the tax authority detects the non-compliant behavior and enforces the tax code or ***unsuccessful or detected evasion*** EU_{UE}, the expected income utility for taxpayer *i* treated by message *m* is:

$$EU^i_{UE}\big(Y,T,F(m),p(m)\big)=\sum_F\sum_p \quad b^i_{F,p}(F,p|M=m)U\big[p(m)\big(Y_i-F(m)\big)\big]$$

Similarly to the standard AS-Y model of tax compliance, the probability that taxpayer *i* with information *m* complies with her monies is given by $Pr\big[EU^i_C(Y_i-T)>EU^i_E\big(1-p(m)\big)(Y_i)+p(m)\big(Y_i-F(m)\big)\big]$ (Allingham and Sadmo, 1972; Sadmo, 2005). In other words, the taxpayer will evade her tax liabilities if the monetary gains of evading are larger than the benefits from complying with the tax code.

We empirically estimate the casual effects of the informational treatments *m* (e.g. the deterrence message, the monitoring letter and the interaction treatment) on the probability of complying with the tax code *Pr* using a standard probability or probit model. The randomized design of the field experiment ensures there is no selection bias in the construction of the treatment and the control groups (Dufflo et al, 2008; Angrist and Prischke, 2009). Formally, we estimate the following specification:

$$Prob(Y_i = 1) = f(\alpha_i + \beta_1 Deterrence_i + \beta_2 Monitoring_i + \beta_3 Deterrence_i * Monitoring_i + \gamma X_i + \varepsilon_i) \quad (1)$$

Where *Y* is the binary outcome variable equal to one if the individual taxpayer *i* meets her tax obligations and/or makes any additional payments (e.g. debts and advanced payments) in the August billing period; *Deterrence* is a binary variable equal to one if the taxpayer receives a deterrence message with her tax bill; *Monitoring* is a binary variable equal to one if the taxpayer receives a letter attached to the tax bill with information conveying intensified monitoring efforts on TRSU's compliance from the municipality; *Deterrence*Monitoring* is a binary variable equivalent to one if the taxpayer receives both the message and the letter; *X* is a vector of control variables comprising taxpayer´s observable characteristics (e.g. number of properties) and the amount and type of public services provided by the municipality (e.g. provision of paved roads, streetlights and street cleaning services); and ε is an individually and identically distributed (i.i.d) error term. As in the Chapter 2, we include the lagged outcome variable as an additional control to avoid potential serial correlation concerns, mainly because there is high persistence on payment behavior (Duflo et al (2008).

3.4 Results

In this section, we present the results of the empirical analysis. Table 24 presents the results of the average treatment effects (ATE) of receiving any message (e.g. a deterrence message and/or a monitoring letter), using the empirical model described in equation (1). The table shows marginal the effects computed at the mean and robust standard errors are displayed in parentheses. The evidence shows that receiving a treatment message and/or a letter increased the probability of complying with the TRSU by 3 percentage points with respect to the control group that received a placebo message at the 1% significance level.

The magnitude of the coefficient is even larger, and remains robustly significant at the 1% level, when we include the control variables in Column (2): the pre-experimental outcome variable, the first principal component of the public services provided by the municipality and the number of taxed properties. Interestingly, the probability of payment seems to be negatively correlated to the number of taxpayers' properties. In Chapter 2, we find similar results in a field experiment carried out in a relatively larger and more urbanized municipality than 9 de Julio. This result also relates to the standard AS-Y model that establishes a positive relation between taxpayer´s income and tax avoidance (Sandmo and Allingham, 1972; Sandmo, 2005). Contrastingly, the supply and quality of

locally provided public services does not seem to be correlated with tax compliance. Expectedly, the inclusion of block-fixed effects increases the precision of our estimates while maintains the direction and statistical significance of the message and interaction coefficients.

Table 25 explores the potential marginal effects of the deterrence message, the monitoring letter and the interaction treatment (the group of taxpayers that received simultaneously the message and the letter). The first set of regressions (baseline) includes the three treatment messages, the lagged outcome variable, and block-level fixed effects. The second set of regressions (controls) includes also the principal component of the variables for public service provision (paved roads, street cleaning and street lightning services during the period), the (log of the) number of properties that each taxpayer has, and the (log of the) average tax liability.

The first set of regressions show that the deterrence message had a larger effect on tax compliance than the ATE of receiving a message and/or a letter. The deterrence message increases the probability of complying with the tax code by a maximum of 5 percentage points at the 1% significance level in the second set of regressions that includes the control covariates as well as block fixed effects. In contrast, the monitoring letter does not seem to have a statistically significative effect. Still, tax compliance for taxpayers that received both the deterrence message and the monitoring letter is 2-3 percentage points higher than the counterfactual, depending on the specification, albeit the coefficient is imprecisely estimated at a lower significance level. Full regression tables are provided in Table 8 in the Appendix 4.

As in the field experiment presented in Chapter 2, a potential concern with the results presented in Table 25 is that some of the taxpayers may have not read the messages, and therefore were not treated, introducing a downward bias in our estimates. In particular, there is a group, the "unrecoverable debtors", defined as taxpayers that did not registered any payments on their TSRU liabilities between 2010 and 2012, who are more likely to ignore the messages included in the bill. To investigate this potential concern, we estimate in the specifications (6) to (8) in the Table 25 the same empirical model as before but excluding the taxpayers that were unrecoverable debtors by January, 2012 from the sample. The estimations show a slightly larger effect of the deterrence message treatment on tax compliance, suggesting that we could be in fact underestimating its impact due to the presence of a potential intended-to-treat (ITT) downward bias (Dufflo et al 2008; Angrist and Prischke, 2008). Table 26 shows that this change in the magnitude of the deterrence message coefficient is not driven by a potential bias in the underlining distribution of taxpayers amongst treatments and control groups. The balancing between the treatment messages and the control group remains virtually unchanged when unrecoverable debtors are excluded from the sample.

To provide some intuition for the findings in terms of revenues collected instead of individual level decisions, Figure 14 plots the tax compliance ratio –computed as tax payments over tax

liabilities-, for the three treatment groups and the control group in August. It also displays –in a solid line- the average tax compliance ratio prior to the experiment (January 2012) as well as –in a dotted line- the average tax compliance ratio in the same period (August) but one year before (2011). This figure suggests that tax compliance for taxpayers that received the deterrence message was 3 percentage points higher than for taxpayers in the control group (that received a placebo message). This difference is statistically significant according to a nonparametric test of proportions. Further, while the monitoring letter by itself does not seem to have any effect, tax compliance for taxpayers that simultaneously received both the deterrence message and the monitoring letter is around 3 percentage points higher than for the control group that received the placebo message.

Table 27 benefits from the block or stratified strategy of the field experiment described in Section 4.4 to investigate the potential presence of heterogeneous treatment effects across different groups of taxpayers defined by (the log) of the number of properties and compliance status with the tax code prior to the experiment (January, 2012). We include in all the regressions the same c ontrols as before and block-fixed effects. Columns (1) and (2) estimate the probit model described in equation (1) for taxpayers with one registered taxed property and that complied and violated the tax code one year before the experiment, respectively. Columns (3) and (4), in turn, show the results for taxpayers with multiple properties that complied and violated the tax code prior to the experiment, respectively.

The estimations presented in Column (1) show that tax compliance for non-compliant taxpayers with one taxed property and that received a deterrence treatment message was 4 percentage points higher than the control group that received the placebo message. Still, the coefficient is imprecisely estimated at the 10% significance level. The monitoring-based letter, in turn, had a larger effect than the deterrence-based message on the probability of complying with the tax code (6 percentage points) at the 5% significance level. Taxpayers in this group that received both the letter and the message increased the probability of payment by 5 percentage points relative to the counterfactual. In contrast, there are no significative effects of the message, the letter or the interaction treatment on compliance for compliant taxpayers with a single taxed property.

We find even stronger effects for non-compliant taxpayers with more than one taxed property. Tax compliance for those taxpayers in this group that received the deterrence treatment message was 6 percentage points higher than for the control group, albeit at the 10% significance level. The monitoring letter, in turn, had a lightly larger and more robustly estimated effect at the 5% significance level. There are no significative effects of the interaction treatment for these taxpayers. There were also no effects of neither of the treatments on compliant taxpayers with multiple properties.

3.5 Conclusions and further research directions

The AS-Y standard model of tax compliance postulates two main drivers behind the individual's taxpayer decision to comply with the tax code: on the one hand, her subjective perception of the monitoring efforts, and the resultant probability of detection of non-compliant behavior by the tax authorities, and on the other hand, the fines and other legal and administrative punishments that she may receive in the case of not paying her monies.

In this chapter, we present the results of a large randomized field experiment designed to empirically evaluate the relative importance of the deterrence information with respect to informational flows related to intensified monitoring efforts on behalf of the tax authority on determining compliance of a locally levied tax in a municipality of Argentina. We find that a deterrence-based message had a large average treatment effect on tax compliance, whereas the monitoring-based letter did not have a statistically significative impact. Only combined with the deterrence message, the letter with information on the intensified monitoring efforts of the local authorities on non-compliant behavior posit a positive effect on compliance with the tax code.

We also find heterogeneous treatment effects for different groups of taxpayers classified by compliance status (e.g. whether or not the taxpayer complied with the tax code prior to the experiment) and a proxy for income (e.g. the log of the pre-experimental number of taxed properties). The deterrence-based message exerts a larger impact on the compliance behavior of relatively well-off and non-compliant taxpayers. In turn, the provision of information on intensified compliance monitoring and tax enforcement efforts aimed at influencing taxpayer's subjective beliefs about the probability of detection of tax cheating had also a slightly larger effect on this group of taxpayers.

While the results presented in this chapter shed new light on the relative importance of the determinants of tax compliance in the standard AS-Y model, more in depth research is required to understand how information influences the compliance choice of the individual taxpayer. A research avenue worth following is to explore the relative importance of images with respect to written messages at influencing compliance by randomly assigning treatment messages with and without images. While evidence suggests that the use of visual images in written communications can contribute to the reader´s remembering and understanding of the text, it also indicates that images can also distract the attention of the reader from the text (Glenberg and Langton, 1992). Another interesting research direction would be to analyze the effects on compliance with the tax code of different levels of enforcement threat throughout the use of alternative written communications and images in the detection-based treatment messages. While non-expected utility models suggest that subjective perceptions about the probability of detection and enforcement of the tax code may differ across taxpayers, there have not been systematic attempts in the literature to test this result empirically (Hashimzade et al, 2013). Finally, it would also be interesting to empirically analyze whether the

effects of the informational treatments on the compliance decision of the taxpayer persist over time (and for how long) once the messages have been removed from the bill.

References

Acton, J. (1982). "An Evaluation of Economists' Influence on Electric Utility Rate Reforms." American Economic Review, 72: 114-199.

Acton, J. and B. Mitchell (1980). "The Effect of Time-of-Use Rates in the Los Angeles Electricity Study." RAND Corporation Report No. N-1533-DWP/HF. Santa Monica, United States: RAND Corporation.

Aguirre, B. and F. F. Rocha (2010). Trust and Tax Morale in Latin American and Caribbean Countries. Presented at the 14th Annual Conference of The International Society for New Institutional Economics, University of Stirling, Scotland.

Allingham, M.G. and A. Sandmo (1972). Income Tax Evasion: A Theoretical Analysis. Journal of Public Economics, pp. 323-338

Alm, J. (2010). Testing Behavioral Public Economics Theories in the Laboratory. National Tax Journal, December 2010, 63 /4, Part 1), 635-658.

Alm, J. and B. Torgler (2004). Culture Differences and Tax Morale in the United States and Europe. Center for Research in Economics, Management and the Arts Working Paper 2004-14. Basel.

Alm, J. and J. Martínez-Vázquez (2007). Tax Morale and Tax Evasion in Latin America. International Studies Program Working Paper 07-04, Georgia State University.

Alm, J. Jackson, B. and M. McKee (1992). Institutional Uncertainty and Taxpayer Compliance. American Economic Review, Vol. 82, pp. 1018-1026

Alm, J., Jackson, B. and McKee, M. (1993) "Fiscal Exchange, Collective Decision Institutions, and Tax Compliance," Journal of Economic Behavior and Organization, vol. 22, no. 3, pp. 285-303.

Alm, J., McClelland, G. and Schulze, M. (1992). Why do People Pay Taxes? Journal of Public Economics 48: 21-38.

Andreoni, J., Erard, B. and Feinstein, J. (1998) Tax Compliance. Journal of Economic Literature, Vol. 36, No.2, pp. 818-860, June.

Angrist, J. and J. Pischke (2008). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Angrist, J. and V. Lavy (2009). "The Effect of High Stakes High School Achievement Awards: Evidence from a Group-Randomized Trial." American Economic Review, 99: 1384-1414.

Barone, G and S. Mocetti (2009). Tax Morale and Public Spending Inefficiency. Working Paper 732. November. Rome, Italy: Banca D'Italia.

Bartolome, C. A. de (1995), 'Which tax rate do people use: Average or marginal?', Journal of Public Economics 56, 79–96.

Bayer, R. (2006) Moral Constraints and Evasion of Income Tax. The ICFAI Journal of Public Finance 4: 7-31.

Bergman, M. and Nevarez, A. (2006) Do Audits Enhance Compliance? An Empirical Assessment of VAT Enforcement. National Tax Journal Vol. LIX, No. 4, December.

Bernheim, D. and Rangel, A. (2007). "Behavioral Public Economics: Welfare and Policy Analysis with Nonstandard Decision-makers" in Behavioral Economics and Its Applications, edited by Diamond, P. and Vartiagen, H. Princeton University Press.

Bernheim, D. and Rangel, A. (2009). Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics. Quarterly Journal of Economics. Vol. 124, no. 1, February, pp. 51-104.

Bernstein, M. and J. Griffin (2006). "Regional Differences in the Price-elasticity of Demand for Energy." National Renewable Energy Laboratory, Midwest Research Institute.

Blumenthal, M. Christian, C. and J. Slemrod (2001). Taxpayer Response to an Increased Probability of Audit: Evidence from a Control Experiment in Minnesota. Journal of Public Economics, Vol. 79, pp. 455-483

Blumenthal, M. Christian, C. and J. Slemrod (2001b). Do Normative Appeals Affect Tax Compliance? Evidence from a Control Experiment in Minnesota. National Tax Association, Vol. 54.

Bruhn, M. and MacKenzie, D. (2009) In Pursuit of Balance: Randomization in Practice in Development Field Experiments. American Economic Journal: Applied Economics, 1(4), pp.200-232

Bushnell, J. and E. Mansur (2005). "Consumption Under Noisy Price Signals: A Study of Electricity Retail Rate Deregulation in San Diego." Journal of Industrial Economics, 53: 493-513.

Cameron, C., J. Gelbach and D. Miller (2008). "Bootstrap-Based Improvements for Inference with Clustered Errors." Review of Economics and Statistics, 90: 414-427.

Caves, D. and L. Christensen (1980). "Econometric Analysis of Residential Time-of-Use Electricity Pricing Experiments." Journal of Econometrics, 14: 287-306.

Chetty, R. and Saez, E. (2010). "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients," American Economic Journal: Applied Economics, American Economic Association, vol. 5(1), pages 1-31, January.

Chetty, R., Looney, A. and Kroft, K. (2009). "Salience and Taxation: Theory and Evidence." American Economic Review, 99(4): 1145-77.

Coleman, S. (1996). The Minnesota Income Tax Compliance Experiment: State Tax Result. Minnesota Department of Revenue. November.

Coleman, S. (2007). The Minnesota Income Tax Compliance Experiment: Replication of the Equity Experiment. MPRA, Paper No. 5820. November.

Congdon, W.J., Kling, J.R. and Mullainathan, S. (2011). Public Policy and Choice. Public Finance Through the Lens of Behavioral Economics. Brookings Institute. Washington D.C.

Cowell, F.A. and Gordon, J.P.F (1988) Unwillignes to Pay. Tax Evasion and Public Good Provision. Journal of Public Economics 36 (1988) 305-321. North-Holland

Cummings, R.G., Martínez-Vázquez, J. McKee, M. and B. Torgler (2006). Effects of Tax Morale on Tax Compliance: Experimental and Survey Evidence. Center for Research in Economics, Management and the Arts Working Paper 2005-29. Basel.

Dargay, J. (1993). "Demand Elasticities: A Comment." Journal of Transport Economics and Policy, 27(1): 87-90.

Davis, L. and E. Muehlegger (2010). "Do Americans Consume too Little Natural Gas? An Empirical Test of Marginal Cost Pricing." RAND Journal of Economics, 41: 791-810.

Dell'Anno, R. (2009). Tax Evasion, Tax Morale and Policy Maker's Effectiveness. The Journal of Socio-Economics, Vol. 38 (2009), pp. 988–997.

Dhami, S. and al-Nowaihi, A. (2007). Why do people pay taxes? Prospect theory versus expected utility theory. Journal of Economic Behavior & Organization 64: 171–192

Dhami, S. and al-Nowaihi, A. (2010). Optimal Income Taxation in the Presence of Tax Evasion: Expected Utility versus Prospect Theory. Journal of Economic Behavior and Organization 75(2): 313–37.

Dubin, J. (1985). Consumer Durable Choice and Demand for Electricity. Amsterdam: North-Holland.

Dubin, J. and D. McFadden (1984). "An Econometric Analysis of Residential Appliance Holdings and Consumption." Econometrica, 52: 345-362.

Dubin, J., Graetz, M. and Wilde, L. (1987) Are We a Nation of Tax Cheaters? New Econometric Evidence on Tax Compliance. American Economic Review, 77(2) pp. 240-45, May.

Duflo, E. Glennester, R. and M. Kremer (2008). Using Randomization in Development Economic Research: A Toolkit. Handbook of Development Economics, Vol. 4, ch. 61, pp.3895-3962.

Eisenhauer, J.G. (2008) Ethical Preferences, Risk Aversion, and Taxpayer Behavior. Journal of Socio-Economics 37: 289-304.

Fellner, G., Sausguber, R. and Traxler, C. (2011) Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information. Forthcoming in the Journal of the European Economic Association.

Fezzi, C. and D. Bunn (2010). "Structural Analysis of Electricity Demand and Supply Interactions." Oxford Bulletin of Economics and Statistics, 72: 827-856.

Fortin, B., Lacroix, G. and Villeval, M. (2007) Tax Evasion and Social Interactions. Journal of Public Economics 91: 2089-2112.

Frey, B.S. and B. Torgler (2007). Tax Morale and Conditional Cooperation. Journal of Comparative Economics, Vol. 35 (2007), pp. 136–159

Fryer, R.G. (2013). Information and Student Achievement. Evidence from a Cellular Phone Experiment. NBER working paper 19113. June. Cambridge, MA: The National Bureau of Economic Research.

Glenberg AM and Langton WE (1992). Comprehension of illustrated text: pictures help to build mental models, Journal of Memory and Language 31(2): 129–51

Hand, M. (2002). "The Economists: On the Future of Energy Markets." Public Utilities Fortnightly, 140: 12-18.

Hashimzade, N. Myles, G.D. and B. Tran-Nam (2010). Applications of Behavioural Economics to Tax Evasion. Journal of Economic Surveys. Published online: 24 MAY 2012. DOI: 10.1111/j.1467-6419.2012.00733.x

Hashimzade, N., Myles, G. D. and Tran-Nam, B. (2013), Applications of Behavioral Economics to Tax Evasion. Journal of Economic Surveys, 27: 941–977. doi: 10.1111/j.1467-6419.2012.00733.x

Hausman, J., M. Kinnucan and D. McFadden (1979). "A Two-Level Electricity Demand Model: Evaluation of the Connecticut Time-of-Day Pricing Test." Journal of Econometrics, 10: 263-289.

Hsiao, C. and D. Mountain (1985). "Estimating the Short-run Income Elasticity of Demand for Electricity." Journal of the American Statistical Association, 80: 259-265.

IDB (2008). Beyond Facts. Understanding Quality of Life. Cambridge, MA: Harvard University Press.

IDB (2013). More than Revenue. Taxation as a Development Tool. New York, NY: Palgrave Macmillan.

Imbens, G. and T. Lemieux (2008). "Regression Discontinuity Designs: A Guide to Practice." Journal of Econometrics, 142: 615-635.

Ito, K. (2010). "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." UC-Berkeley, mimeo.

Jones, D. (2010). "Information, Preferences, and Public Benefit Participation: Experimental Evidence from the Advance EITC and 401(k) Savings." American Economic Journal: Applied Economics 2 (2): 147–63.

Kahneman, D. (2011), Thinking, Fast and Slow, Penguin.

Kendall, C., Nannicini, T. and Trebbi, F. (2013) How do Voters Respond to Information? Evidence from a Randomized Campaign. NBER Working Paper No. 18986. Issued in April.

Kleven, H.J. Knudsen, M.B. Kreiner, C.T. Pedersen, S. and M. Saez (2011). Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. Econometrica, Vol. 79, No. 3, pp. 651-692.

Krichene, N. (2002). "World Crude Oil and Natural Gas: A Demand and Supply Model." Energy Economics, 24: 557-576.

Lacetera, N., Macis, M. and R. Slonim (2013). Economic Rewards to Motivate Blood Donations. Science 340(May): 927-928.

Lee, D. (2008). "Randomized Experiments from Non-random Selection in U.S. House Elections." Journal of Econometrics, 142: 675-697.

Lee, D. and D. Card (2008). "Regression Discontinuity Inference with Specification Error." Journal of Econometrics, 142: 655-674.

Lee, D. and T. Lemieux (2010). "Regression Discontinuity Designs in Economics." Journal of Economic Literature, 48: 281–355.

List, J. Sadoff, S. and M. Wagner (2010). So you Want to Run an Experiment, Now What? Some Simple Rules of Thumb for Optimal Experimental Design. Economic Science Association.

Liu, L. and M. Lin (1991). "Forecasting Residential Consumption of Natural Gas using Monthly and Quarterly Time Series." International Journal of Forecasting, 7: 3-16.

Loewenstein, G. and P. Ubel (2010). "Economics Behaving Badly." The New York Times, 14 July. Available at: http://www.nytimes.com/2010/07/15/opinion/15loewenstein.html

McCrary, J. (2008). "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." Journal of Econometrics, 142: 698-714.

Murphy, K. (2004). The Role of Trust in Nurturing Compliance: A Study of Accused Tax Avoiders. Centre for Tax System Integrity, Working Paper 49. The Australian National University.

Myles, G.D. and Naylor, R.A. (1996) "A model of tax evasion with group conformity and social customs," European Journal of Political Economy, Elsevier, vol. 12(1), pages 49-66, April.

National Institute of Statistics and Censuses (2003). "? Qué es el Gran Buenos Aires?", Press Release.

Ortega, D. Ronconi, L. and P. Sanguinetti (2012). Reciprocity and Willingness to Pay Taxes: Evidence from a Survey Experiment in Latin America. CAF Working Papers.

Parks, R. and D. Weitzel (1984). "Measuring Consumer Welfare Effects of Time-Differentiated Prices." Journal of Econometrics, 26: 25-65.

Parti, M. and C. Parti (1980). "The Total and Appliance-Specific Conditional Demand for Electricity in the Household Sector." Bell Journal of Economics, 11: 309-321.

Pew Research Center (2012). "Assessing the Representativeness of Public Opinion Surveys." Washington DC.

Pomeranz, D. (2010). No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. Harvard University. Job Market Paper. January. Mimeo.

Pommerehne, W. and Frey, Bruno. (1992) The Effects of Tax Administration on Tax Morale. Unpublished manuscript. University of Saarbrucken.

Rawling, G. (2004). Cultural Narratives of Taxation and Citizenship: Fairness, Groups and Globalisation. Centre for Tax System Integrity, Working Paper No. 52.

Reiss, P. and M. White (2005). "Household Electricity Demand, Revisited." Review of Economic Studies, 72: 853-883.

Reiss, P. and M. White (2008). "What Changes Energy Consumption? Prices and Public Pressures." RAND Journal of Economics, 39: 636-663.

Rincke, J. and C. Traxler (2011). Enforcement Spillovers. The Review of Economics and Statistics 93(4): 1224–1234

Sandmo, A . (2005) “The theory of tax evasion: A retrospective view,” National Tax Journal 58, no. 4, December, 643-663.

Sanjit, D. and al-Nowaihi, A (2010), "Optimal Income Taxation in the Presence of Tax Evasion: Expected Utility versus Prospect Theory," Journal of Economic Behav- ior and Organization, vol. 75, no. 2 (2010), pp. 313-37.

Sanjit, D. and al-Nowaihi, A. (2007) "Why Do People Pay Taxes? Prospect Theory versus Expected Utility Theory," Journal of Economic Behavior and Organization, vol. 64, no. 1, pp. 171-92

Scartascini, C. (2011) "Tax Compliance in Latin America". Research Department Seminar Presentation. January.

Scholz, J.T. and M. Lubell (1998). Trust and Taxpaying: Testing the Heuristic Approach to Collective Action. American Journal of Political Science, Vol. 42, No. 2 (Apr., 1998), pp. 398-417

Scholz, J.T. and N. Pinney (1995). Duty, Fear, and Tax Compliance: The Heuristic Basis of Citizenship Behavior. American Journal of Political Science, Vol. 39, No. 2 (May, 1995), pp. 490-512

Schwarz, R. and S. Orleans (1967). "On Legal Sanctions", University of Chicago Law Review 34, 274-300.

Shin, J. (1985). "Perception of Price when Price Information is Costly: Evidence from Residential Electricity Demand." Review of Economics and Statistics, 67: 591-598.

Smith, Adam (1776), An Inquiry into the Nature and Causes of the Wealth of Nations, Cannan ed., London, Methuen.

Snow, A. and Warren, R. (2005) Tax Avoidance, Evasion and Administration. In Auerbach and Feldestein (eds) Handbook of Public Economics, Volume 3. Amsterdam, Holland: North Holland.

Thaler, R.H. and C. R. Sunstein (2008). Nudge. Improving Decisions about Health, Wealth and Happiness. Yale University Press.

Torgler, B. (2003). Tax Morale: Theory and Empirical Analysis of Tax Compliance. PhD Thesis, University of Bassel.

Torgler, B. (2003). Tax Morale: Theory and Empirical Analysis of Tax Compliance. PhD Thesis, University of Bassel.

Torgler, B. (2004). Moral suasion: An alternative tax policy strategy? Evidence from a controlled field experiment in Switzerland. Economics of Governance 5: 235–253

Torgler, B. (2005). Tax Morale in Latin America. Public Choice, Vol. 122, pp. 133-157

Torgler, B. Demir, A. Macintyre and M. Schaffner (2008). Causes and Consequences of Tax Morale: An Empirical Investigation. Economic Analysis and Policy, Vol. 38 No. 2

Traxler, C. (2010). Social norms and conditional cooperative taxpayers. European Journal of Political Economy 26: 89–103

Traxler, C. and J. Winter (2012). Survey evidence on conditional norm enforcement. European Journal of Political Economy 28: 390–398

Wenzel, M. (2001). Alternating Norm Perceptions to Increase Tax Compliance. Centre for Tax System Integrity, Working Paper 38. The Australian National University.

Wenzel, M. (2005). The multiplicity of taxpayer identities and their Implications for Tax Ethics. Centre for Tax System Integrity, Working Paper 73. The Australian National University.

Wenzel, M. and Taylor, N. (2004). "An Experimental Evaluation of Tax-reporting schedules: a case of evidence-based tax administration". Journal of Public Economics 88 2785'2799

Witte, A. and Woodbury, D. (1985) "The Effect of Tax Laws and Tax Administration on Tax Compliance: The Case of the U.S. Individual Income Tax", National Tax Journal, 38(1) pp. 1-13.

Yaniv, G. (1999) Tax Compliance and Advance Tax Payments: A Prospect Theory Analysis. National Tax Journal 52: 753-764

Yitzhaki, S. (1974). A Note on Income Tax Evasion: A Theoretical Analysis. Journal of Public Economics, Vol. 3, pp. 201-202.

Tables and Figures

Table 1: Tariff schedule valid since November 1st, 2008

Category	Accumulated consumption (m^3/year)		Unit price per m^3
	From	To	
R1	0	500	0.144
R21	501	650	
R22	651	800	
R23	801	1000	0.156
R31	1001	1250	0.247
R32	1251	1500	0.332
R33	1501	1800	0.437
R34	1801	$\infty\rightarrow$	0.517

Notes: The econometric analysis focuses on the discontinuity between the categories R32 and R33.

Table 2: Socioeconomic characteristics of households in the estimation sample

Variable	Estimation sample survey	Household Survey
Female	75.4	52.0
Age	52.5	32.1
Married	60.8	38.6
Tertiary education	38.0	19.4
Home owner	88.9	71.2
Number of rooms	3.3	3.1
Number of families	1.1	1.0

Notes: Individual characteristics in the estimation sample survey are from the member of the household that is responsible for paying the gas bills. Statistics on households from Argentina's Household Survey (*Encuesta Permanente de Hogares*) refer to households living in Buenos Aires. "Number of families" refers to the number of different families living in the household. "Number of rooms" refers to the total number of rooms in the household.

Table 3: Knowledge on bill amount and price determination mechanism

Last amount billed		
Question		**% Yes**
Do you remember the amount of your last bill?		91.8
Price determination mechanism – Perceived knowledge		
Question		**% Yes**
Do you know how the total amount of the bill is computed?		30.7
Price determination mechanism – Objective knowledge		
Question	**Correct Answer**	**% Correct Answer**
How often does the company re-categorize consumers?	Every billing period	14.4
Re-categorization is calculated based on…	Last year's consumption	38.9
What is the consumption level that divides categories R32 and R33	1500 m^3	3.7

Notes: Results from a phone survey of 353 customers that had an annual accumulated consumption between 1480 and 1520 m^3 in the bill issued in May 2009. For the questions about objective knowledge, four alternatives were presented. In the question "How often the company re-categorize consumers?" the options were: a) every billing period, b) every two billing periods, c) every six billing periods, d) other. In the question "Re-categorization is calculated based on …" the options were: a) difference in consumption between the current bill and the previous; b) last year's consumption; c) last semester consumption; d) other. In the question "What is the consumption level that divides categories R32 and R33" the options were: a) 1000 m^3; b) 2000 m^3; c) 1,500 m^3; d) Does not know.

Table 4: Dates and periods by treatment status
(Days normalized: May 1st, 2009 = Day 1)

	Treatment	Control	Raw Difference	Adjusted Difference
Panel A: Dates				
Final Measurement Period 0	10.93	10.77	0.17	0.27
	(0.14)	(0.11)	(0.17)	(0.29)
Bill Period 0 (*Treatment Date*)	17.07	16.91	0.16	0.30
	(0.13)	(0.12)	(0.18)	(0.30)
Final Measurement Period 1	71.62	71.45	0.17	0.31
	(0.13)	(0.12)	(0.17)	(0.29)
Panel B: Days Between Bill Period 0 and Final Measurement Period 1				
Days	54.55	54.54	0.01	0.01
	(0.02)	(0.03)	(0.03)	(0.03)
Panel C: Days Between Final Measurement in Period 0 and Period 1				
Days	60.69	60.68	0.00	0.04
	(0.02)	(0.02)	(0.03)	(0.06)

Notes: Means and standard errors in parentheses. Sample includes all customers with accumulated annual consumption between 1480 and 1520 m^3 in the bill that was issued in May, 2009. Period 0 corresponds to the one whose bill was issued in May 2009. Dates in the table are normalized so May 1st, 2009 corresponds to day 1. Day 11 = May 11th, 2009. Day 17 = 17th May, 2009. Day 72 = July 12th, 2009. Standard errors clustered by accumulated consumption in Period 0.

Table 5: Region of residence by treatment status

	Treatment	Control	Raw Difference	Adjusted Difference
Quilmes	0.172	0.171	0.001	0.000
	(0.004)	(0.009)	(0.010)	(0.014)
Avellaneda	0.119	0.134	-0.015*	-0.016
	(0.006)	(0.006)	(0.008)	(0.014)
Ate. Brown	0.126	0.113	0.013*	0.005
	(0.005)	(0.004)	(0.006)	(0.011)
Flores	0.092	0.090	0.002	0.003
	(0.004)	(0.004)	(0.006)	(0.008)
E. Echeverría	0.092	0.087	0.005	0.016*
	(0.004)	(0.005)	(0.006)	(0.009)
Belgrano	0.069	0.077	-0.007	-0.009
	(0.004)	(0.005)	(0.006)	(0.013)
Floresta	0.060	0.066	-0.005	0.001
	(0.004)	(0.004)	(0.006)	(0.010)
Devoto	0.064	0.054	0.010*	-0.005
	(0.003)	(0.004)	(0.005)	(0.011)
Norte	0.045	0.054	-0.009*	-0.014
	(0.004)	(0.003)	(0.005)	(0.011)
Other	0.160	0.154	0.006	0.019
	(0.006)	(0.005)	(0.008)	(0.014)

Notes: Means and standard errors in parenthesis. The Raw Difference column reports mean difference between the Treatment and Control groups. The Adjusted Difference column presents the coefficient of regressing the respective variable on a dummy for treatment and a linear term for annual accumulated consumption by Period 0. Standard errors clustered by accumulated consumption in Period 0. *** p<0.01, ** p<0.05, * p<0.1. The Other category includes Almagro, Mataderos, Centro, Lomas de Zamora, Barracas, Lanús, San Vicente and Berazategui.

Table 6: Consumption levels and ratios to accumulated consumption by treatment status

	Treatment	Control	Raw Difference	Adjusted Difference
Consumption Levels (m^3)				
Period -5	456.378	450.914	5.464**	3.922
	(1.469)	(1.714)	(2.230)	(4.423)
Period -4	464.935	456.953	7.982***	-1.592
	(1.927)	(1.414)	(2.360)	(5.062)
Period -3	243.048	242.274	0.774	-7.062*
	(1.576)	(1.391)	(2.076)	(3.526)
Period -2	110.743	107.277	3.466***	2.760
	(0.970)	(0.809)	(1.247)	(2.933)
Period -1	96.448	96.001	0.447	6.835
	(1.293)	(0.940)	(1.579)	(4.075)
Period 0	140.970	138.733	2.237	-0.630
	(1.046)	(1.008)	(1.435)	(3.080)
Ratio of Consumption in a Period and Annual Accumulated Consumption by Period 0				
Period -5	0.302	0.303	0.000	0.003
	(0.001)	(0.001)	(0.002)	(0.003)
Period -4	0.308	0.307	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.003)
Period -3	0.161	0.163	-0.002	-0.004
	(0.001)	(0.001)	(0.001)	(0.002)
Period -2	0.073	0.072	0.001	0.002
	(0.001)	(0.001)	(0.001)	(.002)
Period -1	0.064	0.064	-0.001	0.005
	(0.001)	(0.001)	(0.001)	(0.003)
Period 0	0.093	0.093	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.002)

Notes: Means and standard errors in parenthesis. The Raw Difference column reports mean difference between the Treatment and Control groups. The Adjusted Difference column presents the coefficient of regressing the respective variable on a dummy for treatment and a linear term for annual accumulated consumption by Period 0. Standard errors clustered by accumulated consumption in Period 0. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Bill amount by treatment status

	Treatment	Control	Raw Difference	Adjusted Difference
Period -5	101.380	100.558	0.822*	0.678
	(0.305)	(0.288)	(0.414)	(0.731)
Period -4	138.959	134.761	4.198***	3.658
	(1.127)	(0.679)	(1.384)	(3.034)
Period -3	51.406	52.004	-0.598	-2.473***
	(0.350)	(0.363)	(0.498)	(0.891)
Period -2	42.233	41.255	0.978	-0.560
	(0.376)	(0.277)	(0.461)	(0.795)
Period -1	82.449	79.348	3.101	1.299
	(0.512)	(0.403)	(0.643)	(1.816)
Period 0	91.728	72.336	19.393***	17.965***
	(0.565)	(0.436)	(0.704)	(1.474)

Notes: Means and standard errors in parenthesis. Bill amounts are in Argentine Pesos. The Raw Difference column reports mean difference between the Treatment and Control groups. The Adjusted Difference column presents the coefficient of regressing the respective variable on a dummy for treatment and a linear term for annual accumulated consumption by Period 0. Standard errors clustered by accumulated consumption in Period 0. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Impacts of price increase in bill 0 on consumption in period 1, alternative functional forms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	-15.901**	-15.266**	-15.577**	-15.597*	-16.539*	-14.595	-16.516*	-17.126
	(6.491)	(5.929)	(6.153)	(9.194)	(9.201)	(9.917)	(9.174)	(12.408)
AAC	0.720**	1.004**	0.707**	2.201	0.812	7.515**	0.811	2.352
	(0.290)	(0.412)	(0.272)	(1.441)	(0.702)	(2.872)	(0.696)	(5.143)
AAC * Treatment		-0.619		-2.777		-12.721***		-2.244
		(0.531)		(1.968)		(3.936)		(8.023)
AAC^2			-0.011	0.059	-0.011	0.738**	-0.012	-0.487
			(0.012)	(0.069)	(0.012)	(0.356)	(0.044)	(1.014)
AAC^2 *Treatment				-0.014		-0.149		-0.009
				(0.090)		(0.465)		(1.619)
AAC^3					-0.000	0.023*	-0.000	-0.074
					(0.002)	(0.012)	(0.001)	(0.079)
AAC^3 *Treatment						-0.040**		0.136
						(0.015)		(0.121)
AAC^4							0.000	-0.002
							(0.000)	(0.002)
AAC^4 *Treatment								0.001
								(0.003)
Constant	433.271***	436.134***	434.620***	439.962***	435.079***	447.792***	435.135***	443.718***
	(3.299)	(3.992)	(3.539)	(6.161)	(4.700)	(6.506)	(5.037)	(8.874)
N	7190	7190	7190	7190	7190	7190	7190	7190
R-squared	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.002

Notes: The dependent variable is consumption in Period 1. Average of the dependent variable is 425.49. The estimation method is OLS. Standard errors clustered by accumulated consumption in Period 0.*** p<0.01, ** p<0.05, * p<0.1.

Table 9: Impacts of price increase in bill 0 on consumption in Period 1, alternative bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	-15.901**	-15.266**	-16.296*	-14.780**	-18.863*	-15.872*	-16.091	-13.176
	(6.491)	(5.929)	(8.105)	(6.882)	(10.386)	(8.484)	(13.525)	(11.124)
AAC	0.720**	1.004**	0.759	1.443**	1.060	2.403**	0.703	1.898
	(0.290)	(0.412)	(0.451)	(0.612)	(0.832)	(0.863)	(1.774)	(2.309)
AAC * Treatment		-0.619		-1.522*		-3.134**		-3.220
		(0.531)		(0.782)		(1.333)		(2.821)
Constant	433.271***	436.134***	433.000***	438.133***	435.086***	441.855***	438.154***	441.205***
	(3.299)	(3.992)	(4.042)	(4.647)	(5.287)	(5.386)	(7.297)	(8.397)
Bandwidth	20	20	15	15	10	10	5	5
N	7190	7190	5417	5417	3679	3679	1946	1946
R-squared	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.002

Notes: The dependent variable is consumption in Period 1. The bandwidths determine the range of normalized annual accumulated consumption used to select observations for a particular regression. For example, in Columns (1) and (2) only individuals with normalized annual accumulated consumption between [-20, 20] are included. Average of the dependent variable is 425.49. The estimation method is OLS. Standard errors clustered by accumulated consumption in Period 0.*** p<0.01, ** p<0.05, * p<0.1.

Table 10: Exploiting placebo AAC thresholds to test for mean reversion

	(1)
AAC ≥ -10	-1.646
	(6.250)
AAC ≥ 0	-16.969**
	(7.209)
AAC ≥ 10	-3.129
	(6.383)
AAC	0.890
	(0.531)
Constant	435.767***
	(7.941)
N	7190
R-squared	0.001

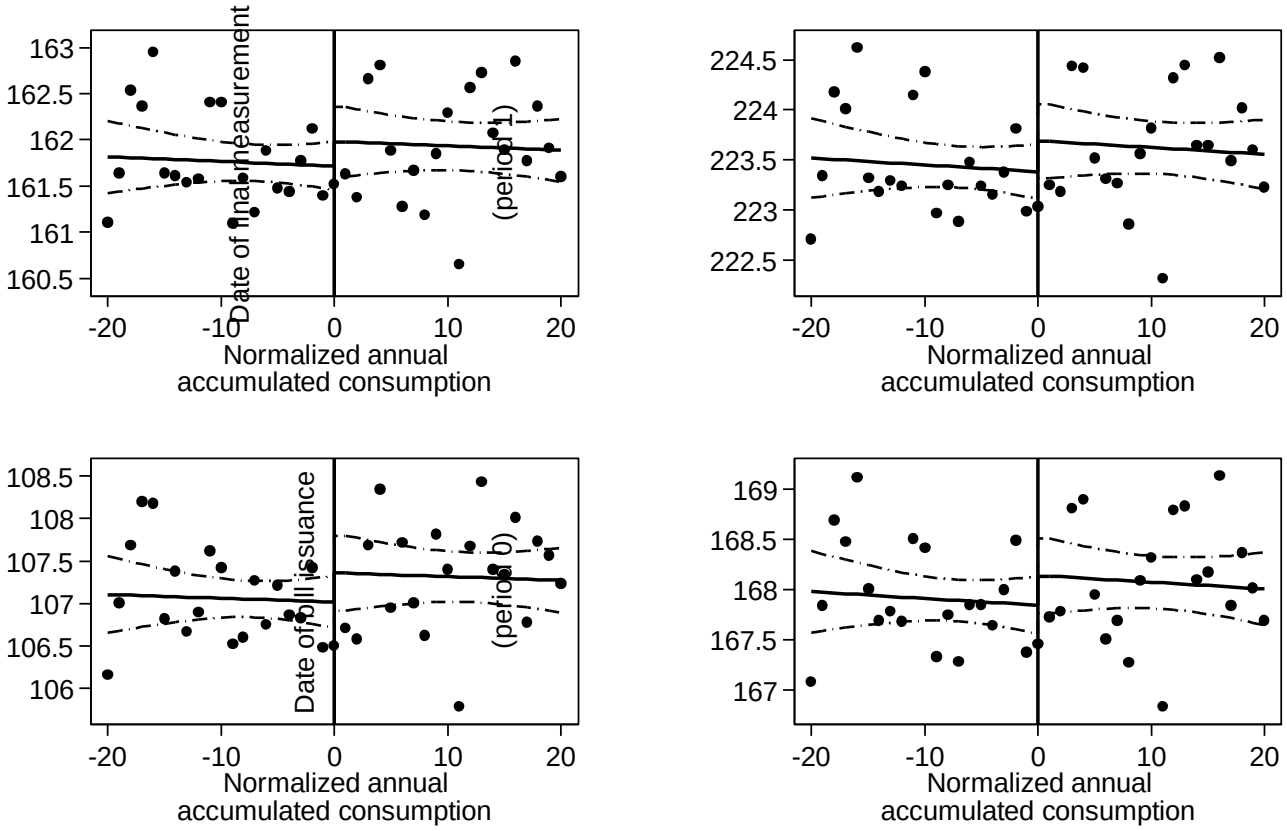
Notes: The dependent variable is consumption in Period 1. Explanatory variables include indicators for annual accumulated consumption (AAC) higher than -10, 0 and 10 and also AAC as a continuous variable. Individuals with normalized annual accumulated consumption between [-20, 20] are included. Average of the dependent variable is 425.49. The estimation method is OLS. Standard errors clustered by accumulated consumption in Period 0.*** p<0.01, ** p<0.05, * p<0.1.

Table 11: Differential effects across consumers in the treatment group by distance to the threshold

	(1)
Close	-16.339**
	(6.567)
Far	-18.500*
	(9.458)
AAC	0.794**
	(0.374)
Constant	434.015***
	(3.835)
N	7190
R-squared	0.001

Notes: The dependent variable is consumption in Period 1. Explanatory variables include indicators for normalized annual accumulated between 1 and 10 (Close) and between 11 and 20 (Far) and 10 and also AAC as a continuous variable. Individuals with normalized annual accumulated consumption between [-20, 20] are included. Average of the dependent variable is 425.49. The estimation method is OLS. Standard errors clustered by accumulated consumption in Period 0.*** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Dates by annual accumulated consumption



Dates in the figure are normalized so December 1st, 2008 corresponds to day 0.

Figure 1 (cont.): Dates by annual accumulated consumption

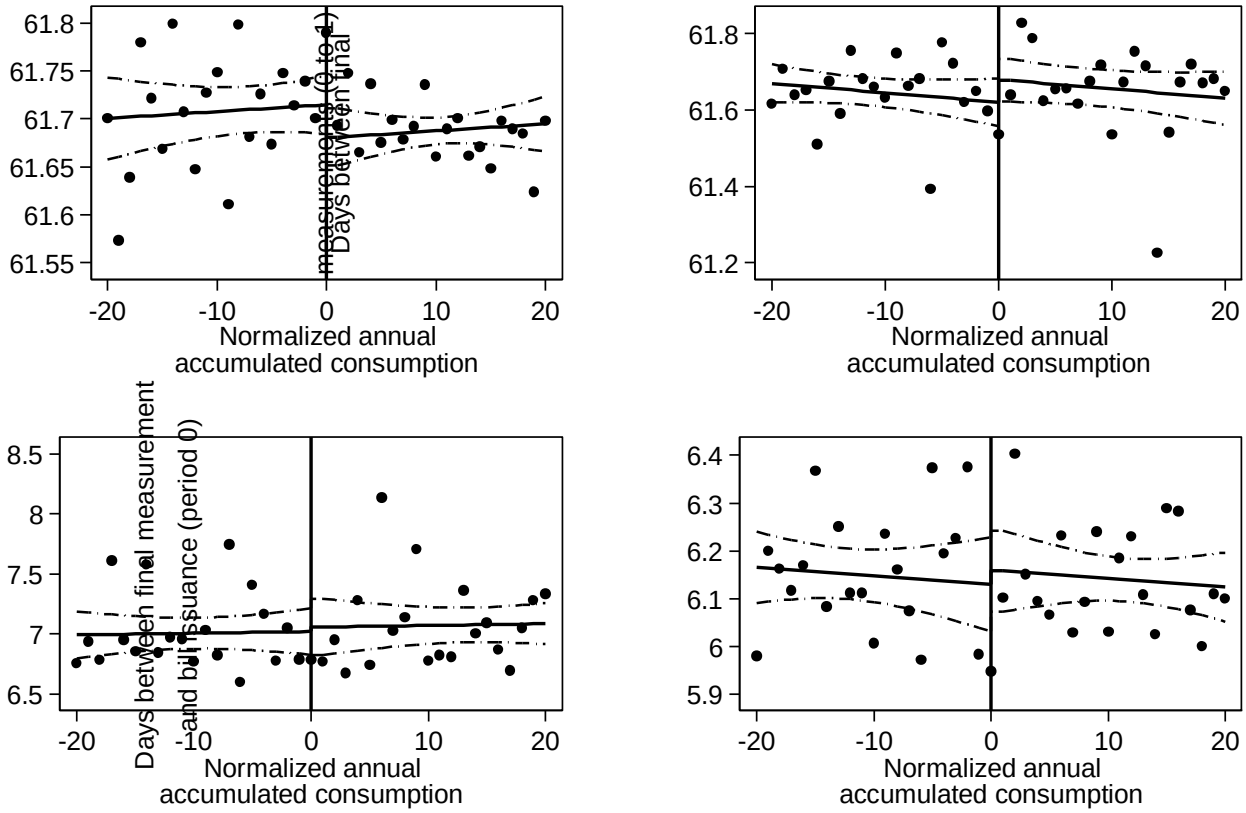


Figure 2: Region of residence by annual accumulated consumption (by six largest regions in terms of number of users)

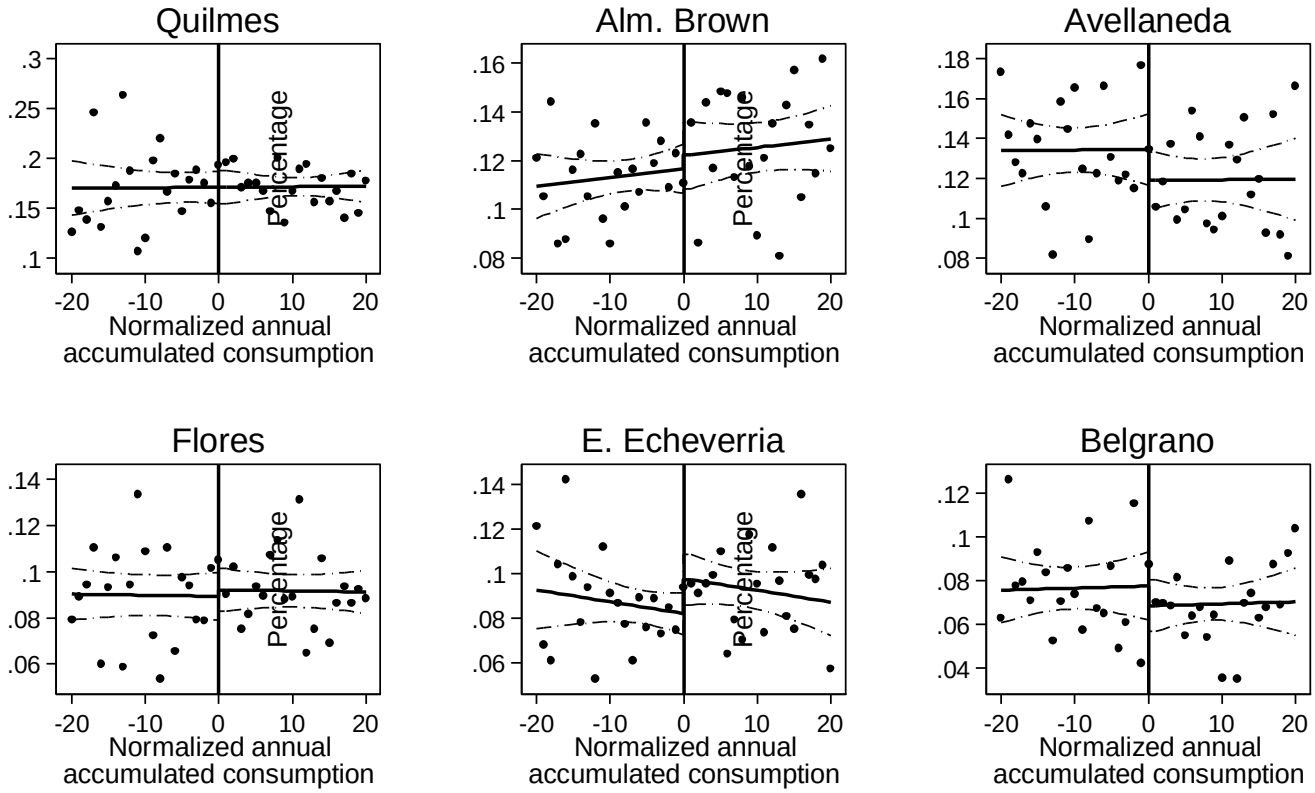


Figure 3: Ratio of consumption to annual accumulated consumption in period 0

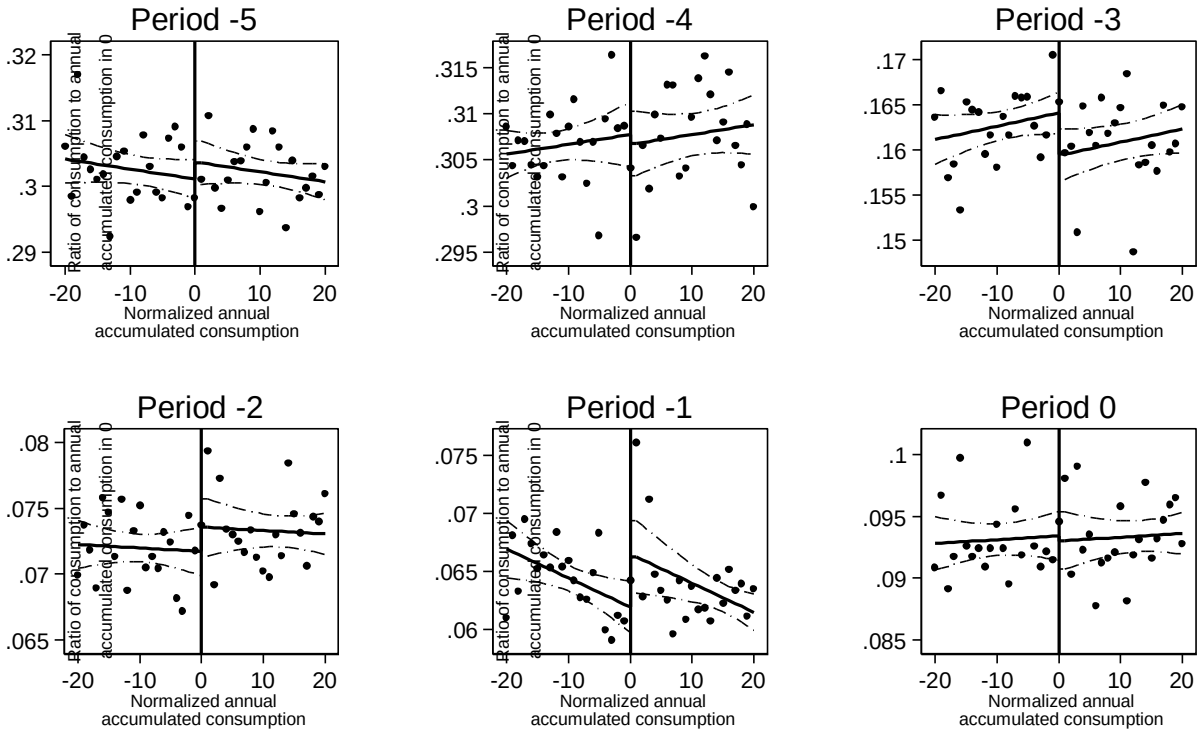


Figure 4: Average bill by annual accumulated consumption

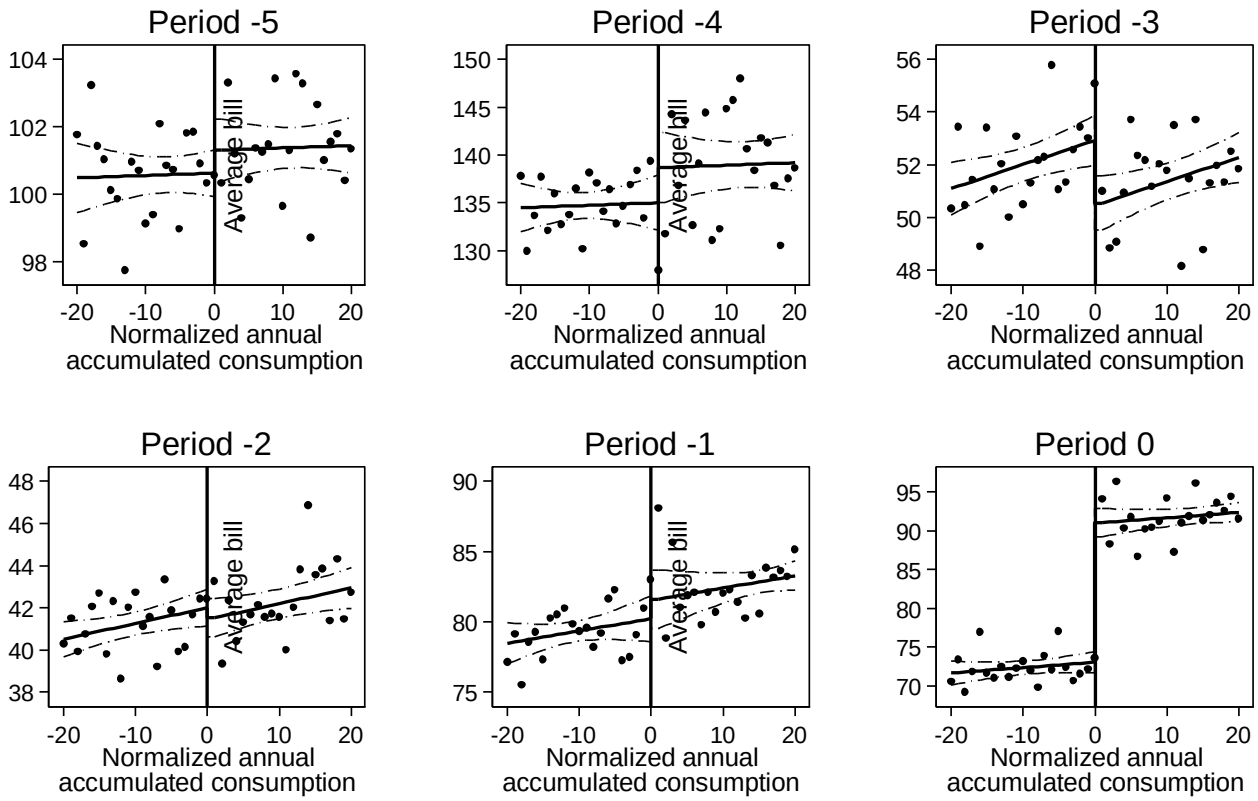


Figure 5: Number of observations by annual accumulated consumption in period 0

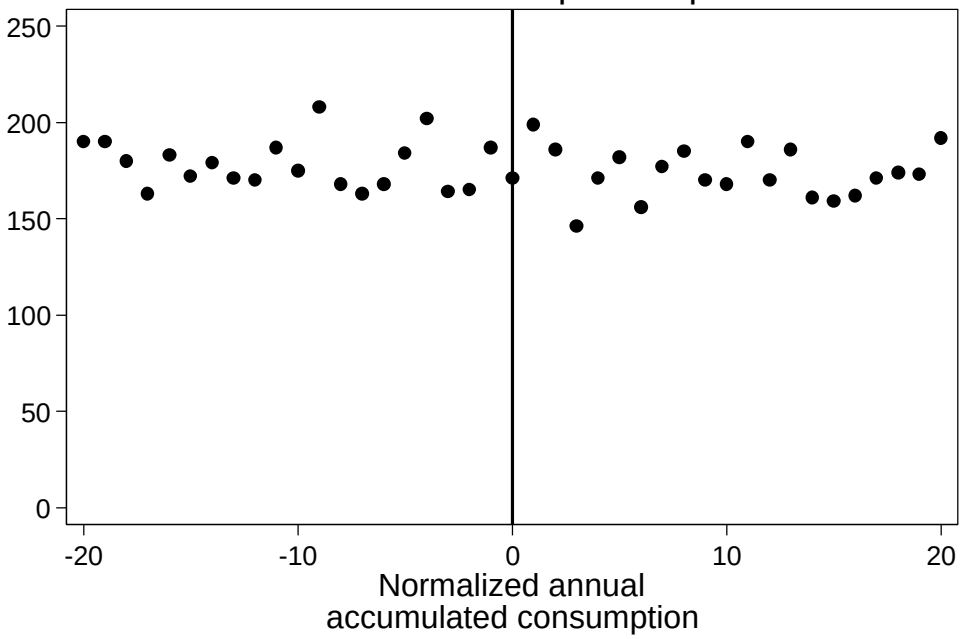


Figure 6: Average consumption in July 2009
by annual accumulated consumption (in m3)

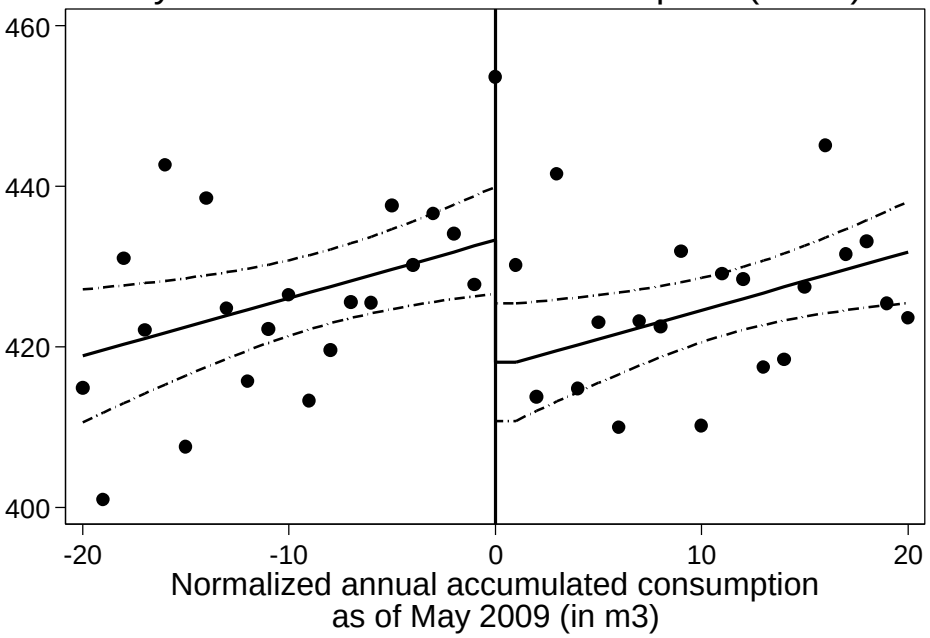


Table 12: Socio-economic Characteristics of Junin and the Province of Buenos Aires

Jurisdiction	Surface (km2)	Population (2010)	Urbanization rate (2010)	School attendance (2001) [1]	GDP in thousands of PPP US$ (2003)	GDP p.c. in PPP US$ (2003)	GDP as percent of Buenos Aires (2003)
Junin	2,260	94,926	90.70%	92.60%	974,615	10,539	0.88%
Average Municipality (Buenos Aires)	2,278	115,514	94.90%	91.90%	817,025	7,654	0.74%

Notes: [1] Percentage of the population between 5 and 18 years old, attending any type of school
Source: own elaboration based on Ministry of Interior of Argentina, National Censuses of 2001 and 2010, Provincial Bureau of Statistics and *Tribunal de Cuentas de la Provincia* de Buenos Aires.

Table 13: Description of Variables in Junin

Variable	Description
Trash collection	Categorical variable equal to 0 if neither trash collection nor street cleaning services are provided; 1 if only trash collection services are provided; and 2 if trash collection and street cleaning services are both provided.
Streetlights	Number of streetlights around the properties associated to the taxpayer
Linear front meters	Average linear front meters of all the properties associated to the taxpayer
Mean valuation of properties	Fiscal assesment of property value. Average for all the properties belonging to an individual taxpayers
Properties	Number of registered properties associated to the taxpayer
Monthly payments	Dummy = 1 if the taxpayer has at least one property for which the bill is paid monthly.
Tax liability	Total amount billed to the taxpayer
Tax payment	Total amount paid by the taxpayer in the period
Tax compliance ratio	Tax payments over tax liabilities ratio (in percent)
Paid_by1D	Dummy = 1 if the payment took place before the 1st due date
Paid_by2D	Dummy = 1 if the payment took place before the 2nd due date
Paid	Dummy = 1 only if the taxpayer has paid in full the total tax liabilities for the period
Overpaid	Dummy = 1 if the taxpayer has paid not only her liabilities for the period in full but she has also made some additional payments
Additionalpayments	Dummy = 1 when the taxpayer has made advanced payments or paid down part of her debt, regardless of whether they have paid their current liabilities in full.
Debtor	If taxpayer has at least one remaining unpaid bill
Debt stock	Accumulated stock of debt
Unrecoverable Debtor	Dummy=1 if the taxpayer did not pay any tax bill between 2007 and 2011

Source: All data collected from the administrative records of the Municipality of Junín

Table 14: Messages included in the Tax Bill

#	Message	Text of the message	Image
1	Deterrence	Did you know that if you do not pay the CVP on time for a debt of AR$ 1,000 you will have to disburse AR$ 268 in arrears at the end of the year and the Municipality can take administrative and legal actions?	
2	Fairness	In the first 6 months of this year, CVP's collection contributed to place 28 new streetlights, water connections in 29 streets and sewerage networks in 21 blocks.	
3	Equity	Did you know that only 30% of taxpayers do not pay the CVP? What about you?	
4		Control group	No message/Image

Table 15: Baseline Summary Statistics and Balance of Randomization (Bim 3)

	(1) Control Group	(2) Difference: Deterrence	(3) Difference: Fairness	(4) Difference: Equity
Trash collection	1.585***	-0.013	0.008	0.009
	(0.004)	(0.011)	(0.011)	(0.011)
Streetlights	2.755***	-0.025	-0.012	-0.041
	(0.010)	(0.025)	(0.025)	(0.025)
Lineal front meters	15.629***	-0.393	-0.200	-0.018
	(0.197)	(0.519)	(0.519)	(0.518)
Mean valuation of properties	15,485.426***	7.038	329.102	506.534
	(182.102)	(480.677)	(481.076)	(480.439)
Properties	1.395***	0.015	-0.056	-0.040
	(0.013)	(0.035)	(0.035)	(0.035)
Monthly payments	0.010***	0.002	0.006**	-0.002
	(0.001)	(0.002)	(0.002)	(0.002)
Tax liability	121.612***	-0.320	-2.898	-3.070
	(1.804)	(4.762)	(4.766)	(4.760)
Tax payments	56.129***	0.661	-1.404	-0.891
	(1.081)	(2.862)	(2.864)	(2.861)
Paid_by1D	0.213***	-0.011	0.014	-0.003
	(0.003)	(0.009)	(0.009)	(0.009)
Paid_by2D	0.332***	-0.014	-0.003	-0.008
	(0.004)	(0.010)	(0.010)	(0.010)
Paid	0.404***	-0.016	0.002	-0.004
	(0.004)	(0.010)	(0.010)	(0.010)
Overpaid	0.031***	-0.002	-0.000	0.004
	(0.001)	(0.004)	(0.004)	(0.004)
Addpay	0.045***	0.002	-0.002	0.000
	(0.002)	(0.004)	(0.004)	(0.004)
Trash collection (Bim 5)	1.586***	-0.012	0.008	0.007
	(0.004)	(0.011)	(0.011)	(0.011)
Streetlights (Bim 5)	2.752***	-0.026	-0.013	-0.037
	(0.010)	(0.025)	(0.026)	(0.025)

Notes: Each row shows a regression of the pre-treatment variable in question on treatment dummies and a constant term. Observations are presented for the bi-monthly period prior to treatment (May-June). The constant captures the value for the control group (no message). Columns (2)-(4) show the difference between the treatment groups and the control group. Monetary amounts are in Argentine Pesos (AR$). Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 16: Average Treatment Effects (Bim 5)

VARIABLES	(1) Paid_by1D	(2) Paid_by2D	(3) Paid	(4) OverPaid	(5) Additional Payments	(6) Paid_1D	(7) Paid_2D	(8) Paid	(9) Overpaid	(10) Additional Payments
T1: Deterrence	0.017*	0.030**	0.047***	0.007**	0.011***	0.017*	0.031**	0.048***	0.007*	0.010**
	(0.010)	(0.012)	(0.015)	(0.004)	(0.004)	(0.010)	(0.012)	(0.015)	(0.004)	(0.004)
T2: Fairness	-0.005	-0.002	-0.001	-0.004	-0.004	-0.006	-0.004	-0.003	-0.004	-0.003
	(0.010)	(0.012)	(0.015)	(0.004)	(0.004)	(0.010)	(0.013)	(0.015)	(0.004)	(0.004)
T3: Equity	0.005	0.004	-0.009	0.002	0.006	0.004	0.004	-0.008	0.002	0.007
	(0.010)	(0.013)	(0.015)	(0.004)	(0.004)	(0.010)	(0.013)	(0.015)	(0.004)	(0.004)
Observations	23,195	23,195	23,176	23,073	23,211	23,186	23,186	23,168	23,065	23,081
Model	Baseline	Baseline	Baseline	Baseline	Baseline	Controls	Controls	Controls	Controls	Controls
Sample	Whole	Whole	Whole	Whole	Whole	Whole	Whole	Whole	Whole	Whole
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable used in each regression identified in the header. The first set of regressions (baseline) include the three treatment messages, the lagged outcome variable, and block-level fixed effects. The second set of regressions (controls) include also the variables for public service provision (trash collection and street lightning services during the period), the (log of the) number of properties that each taxpayer has, the (log of the) average linear front size of the properties, and a dummy that controls for those taxpayers who selected to pay monthly. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 17: Average Treatment Effects (Bim 5) – Taxpayers living outside the city

VARIABLES	(1) Paid_1D	(2) Paid_2D	(3) Paid	(4) OverPaid	(5) Additional Payments	(6) Paid_1D	(7) Paid_2D	(8) Paid	(9) Overpaid	(10) Additional Payments
T1: Deterrence	0.034	0.115*	0.133*	0.008	0.033	0.042	0.127*	0.140*	0.005	0.031
	(0.053)	(0.063)	(0.073)	(0.037)	(0.040)	(0.054)	(0.068)	(0.074)	(0.033)	(0.040)
T2: Fairness	0.006	0.117*	0.149*	0.000	0.009	0.004	0.119*	0.136*	0.002	0.013
	(0.055)	(0.067)	(0.078)	(0.036)	(0.041)	(0.055)	(0.068)	(0.077)	(0.033)	(0.042)
T3: Equity	0.031	0.053	-0.031	0.030	0.051	0.025	0.050	-0.035	0.025	0.049
	(0.052)	(0.068)	(0.087)	(0.032)	(0.036)	(0.052)	(0.065)	(0.086)	(0.029)	(0.036)
Observations	471	471	470	469	476	468	468	467	466	467
Model	Baseline	Baseline	Baseline	Baseline	Baseline	Controls	Controls	Controls	Controls	Controls
Sample	Whole	Whole	Whole	Whole	Whole	Whole	Whole	Whole	Whole	Whole
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable used in each regression identified in the header. The first set of regressions (baseline) include the three treatment messages, the lagged outcome variable, and block-level fixed effects. The second set of regressions (controls) include also the variables for public service provision (trash collection and street lightning services during the period), the (log of the) number of properties that each taxpayer has, the (log of the) average linear front size of the properties, and a dummy that controls for those taxpayers who selected to pay monthly. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 18: Excluding Unrecoverable Debtors (Bim 5)

VARIABLES	(1) Paid_1D	(2) Paid_2D	(3) Paid	(4) OverPaid	(5) Additional Payments	(6) Paid_1D	(7) Paid_2D	(8) Paid	(9) Overpaid	(10) Additional Payments
T1: Deterrence	0.022	0.038**	0.043***	0.010*	0.015**	0.023	0.039**	0.044***	0.009*	0.014**
	(0.014)	(0.016)	(0.016)	(0.005)	(0.006)	(0.014)	(0.016)	(0.016)	(0.005)	(0.006)
T2: Fairness	-0.004	-0.002	0.001	-0.005	-0.006	-0.005	-0.004	-0.000	-0.005	-0.004
	(0.014)	(0.016)	(0.016)	(0.006)	(0.006)	(0.014)	(0.016)	(0.016)	(0.006)	(0.006)
T3: Equity	0.003	0.003	-0.017	0.002	0.009	0.002	0.002	-0.017	0.002	0.009
	(0.014)	(0.016)	(0.016)	(0.005)	(0.006)	(0.014)	(0.016)	(0.016)	(0.005)	(0.006)
Observations	15,845	15,845	15,831	15,794	15,915	15,837	15,837	15,823	15,786	15,800
Model	Baseline	Baseline	Baseline	Baseline	Baseline	Controls	Controls	Controls	Controls	Controls
Sample	Non-defaulters	Non-defaulters	Non-defaulters	Non-defaulters	Non-defaulters	Non-defaulters	Non-defaulters	Non-defaulters	Non-defaulters	Non-defaulters
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

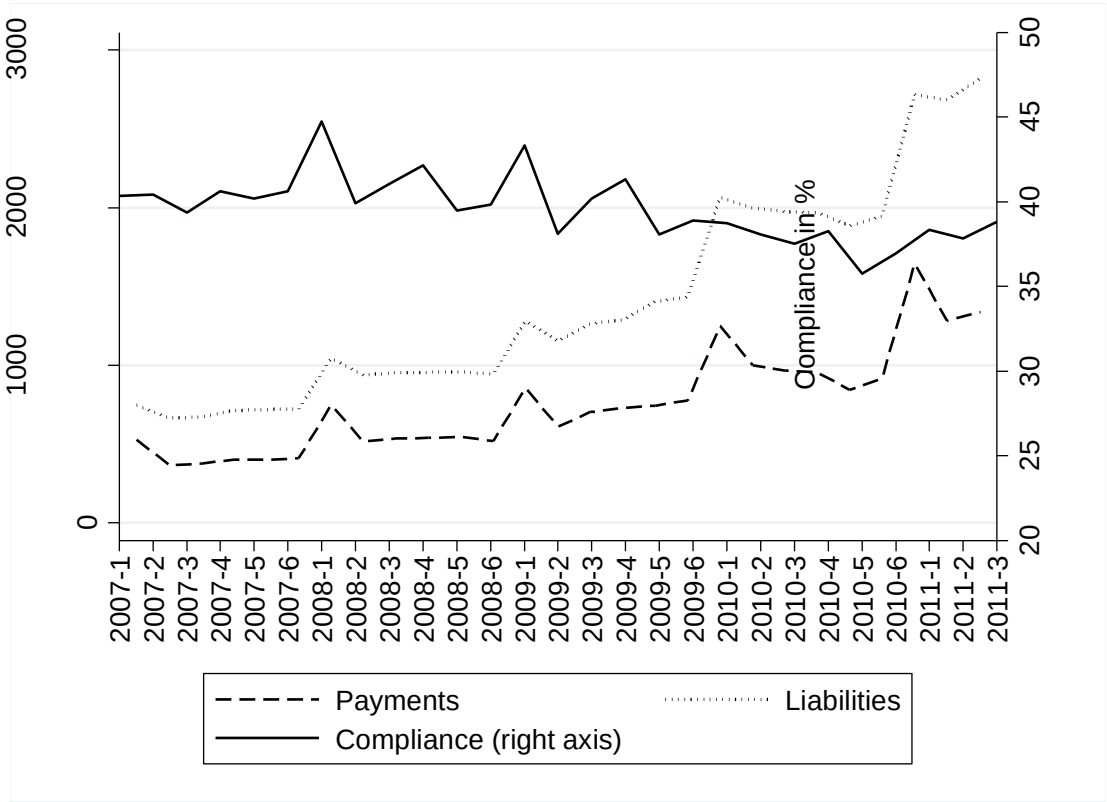
Notes: Dependent variable used in each regression identified in the header. The first set of regressions (baseline) include the three treatment messages, the lagged outcome variable, and block-level fixed effects. The second set of regressions (controls) include also the variables for public service provision (trash collection and street lightning services during the period), the (log of the) number of properties that each taxpayer has, the (log of the) average linear front size of the properties, and a dummy that controls for those taxpayers who selected to pay monthly. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 19: Balancing of the Sample Excluding Unrecoverable Debtors

	(1) Control Group	(2) Difference: Deterrence	(3) Difference: Fairness	(4) Difference: Equity
Trash collection	1.667***	-0.006	0.011	0.000
	(0.005)	(0.012)	(0.012)	(0.012)
Streetlights	2.866***	-0.012	-0.009	-0.050*
	(0.011)	(0.030)	(0.030)	(0.030)
Lineal front meters	14.447***	0.178	-0.182	0.508
	(0.174)	(0.459)	(0.458)	(0.456)
Mean valuation of properties	17,409.872***	-31.546	-192.915	489.637
	(199.751)	(528.298)	(526.250)	(524.859)
Properties	1.445***	0.013	-0.048	-0.045
	(0.016)	(0.043)	(0.043)	(0.043)
Monthly payments=1	0.003***	0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Tax liability	128.839***	0.314	-0.869	-1.142
	(1.766)	(4.670)	(4.652)	(4.639)
Payment for current bim	77.044***	-0.488	-2.326	-2.133
	(1.336)	(3.546)	(3.533)	(3.524)
Paid 100% of the total tax liability before first expiration	0.303***	-0.018	0.013	-0.007
	(0.004)	(0.012)	(0.012)	(0.012)
Paid 100% of the total tax liability before second expiration	0.471***	-0.021*	-0.010	-0.015
	(0.005)	(0.013)	(0.013)	(0.013)
Paid	0.571***	-0.022*	-0.005	-0.013
	(0.005)	(0.013)	(0.013)	(0.013)
Overpaid	0.045***	-0.004	-0.001	0.003
	(0.002)	(0.005)	(0.005)	(0.005)
Addpay	0.063***	0.003	-0.004	-0.002
	(0.002)	(0.006)	(0.006)	(0.006)
Trash collection (Bim 5)	1.669***	-0.004	0.012	-0.002
	(0.005)	(0.012)	(0.012)	(0.012)
Streetlights (Bim 5)	2.862***	-0.011	-0.008	-0.044
	(0.012)	(0.031)	(0.030)	(0.030)

Notes: Each row shows a regression of the pre-treatment variable in question on treatment dummies and a constant term. Observations are presented for the bi-monthly period prior to treatment (May-June). The constant captures the value for the control group (no message). Columns (2)-(4) show the difference between the treatment groups and the control group. Monetary amounts are in Argentine Pesos (AR$). Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 7. Tax liabilities, payments and compliance ratio (AR$ millions and percent), 2007-2011



Source: own elaboration based on administrative information provided by the Municipality of Junín.

Figure 8. Interaction effects of the treatments with the lagged dependent variable.

Figure 8a. Marginal effect of deterrence message according to past payment behavior

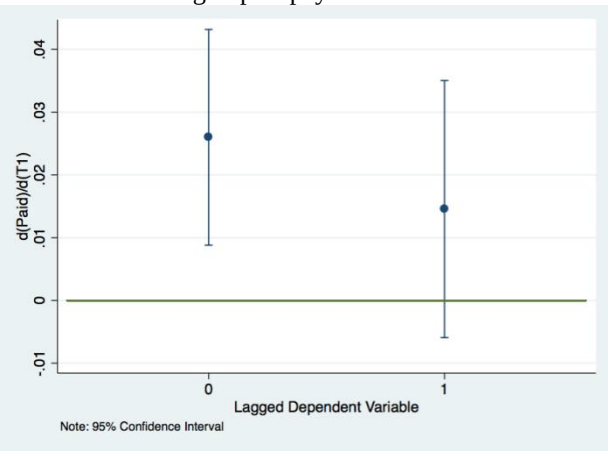


Figure 8b. Marginal effect of equity message according to past payment behavior

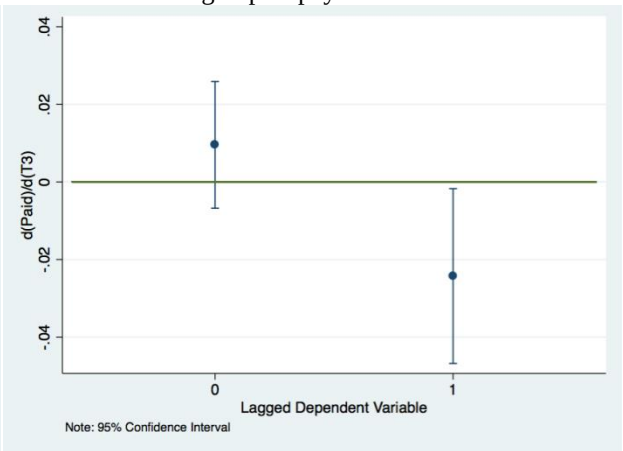


Figure 9. Interaction effects of the treatments with having debts

Figure 9a. Marginal effect of deterrence message according to having debts or not

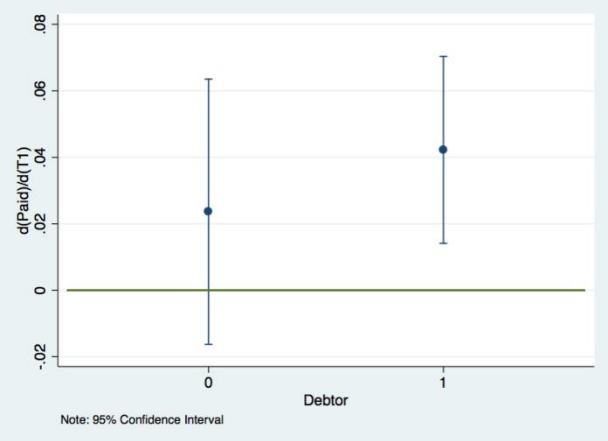


Figure 9b. Marginal effect of equity message according to having debts or not

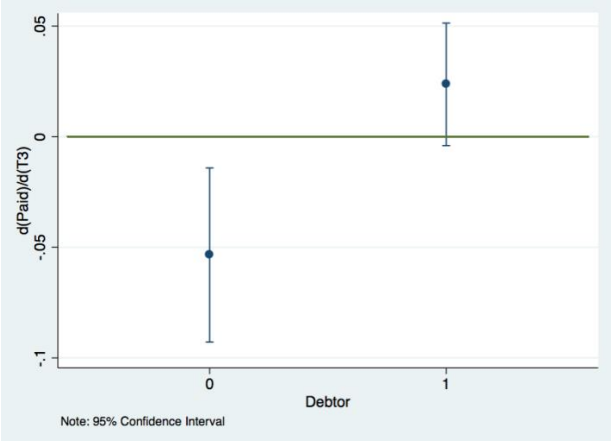


Figure 10. Interaction effects of the equity message with debt stock

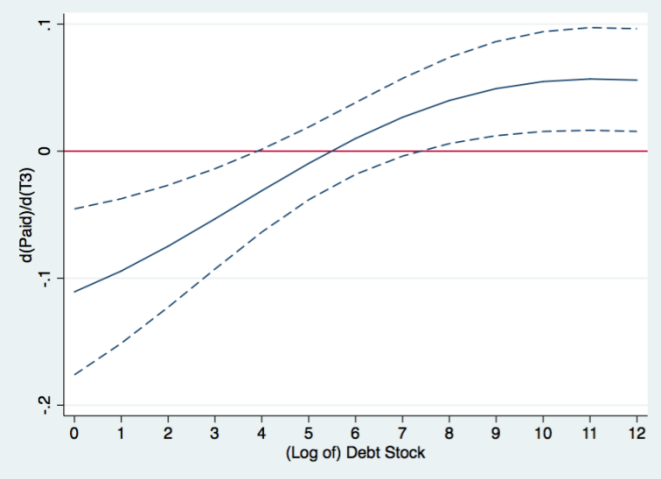


Figure 11. Interaction effects of the message with size of the property

Figure 10a. Marginal effect of deterrence message according to property size

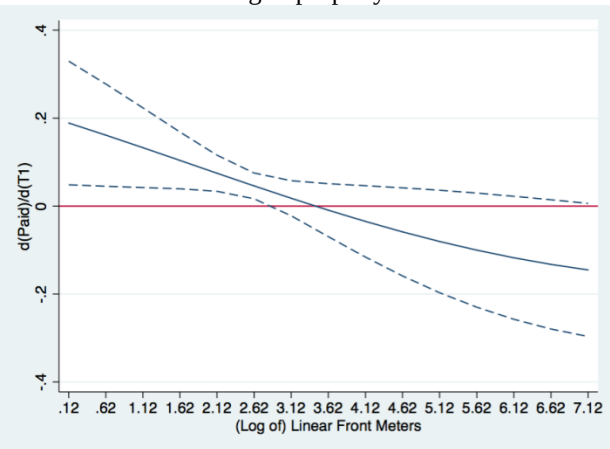


Figure 10b. Marginal effect of fairness message according to property size

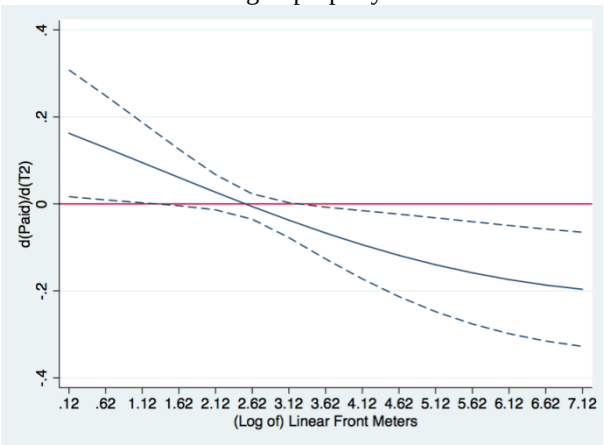
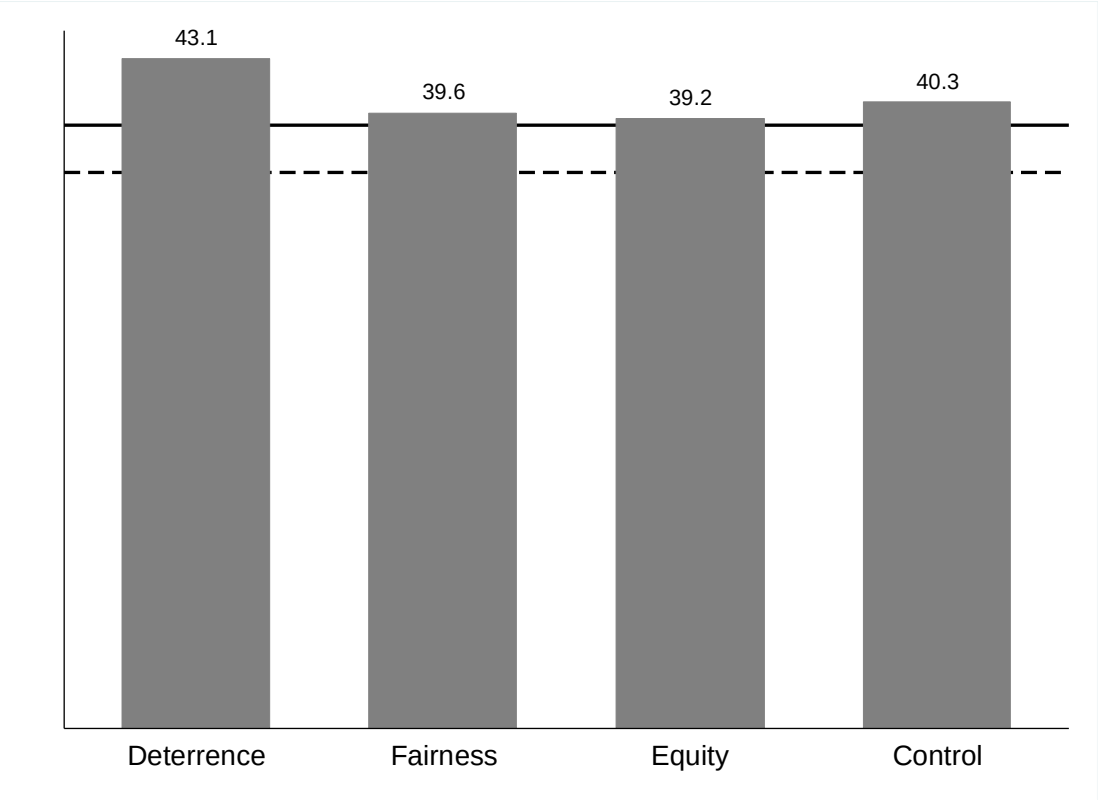


Figure 12. Tax Compliance by Treatment Groups and Control Group (in percent) (Bim 5)



Notes: This figure plots the tax compliance ratio (tax payments over tax liabilities) for the treatment groups and the control. It also displays –in a solid line- the average tax compliance ratio prior to the experiment (Period 3 – May-June) as well as –in a dotted line- the average tax compliance ratio in the same period 5 (September-October) but one year before (2010). The difference in compliance between the group that received the deterrence message and the control is statistically significant at the 1% level according to a proportions test. Differences are not statistically significant for the other two groups.

Table 20. Socioeconomic characteristics of 9 de Julio and the "average" municipality in the province of Buenos Aires

Jurisdiction	Surface km2	Population (2010)	Urbanization rate (2010)	School attendance (2001)	GDP pc in PPP US$
9 de Julio	4,230	47,733	75.7%	92.5%	9,628
Average Municipality (Buenos Aires)	2,278	115,514	94.9%	91.9%	7,654

Notes: School attendance is the percentage of the population between 5 and 18 years old, attending any type of school.

Sources: own elaboration based on Ministry of Interior of Argentina, National Censuses of 2001 and 2010, Provincial Bureau of Statistics and *Tribunal de Cuentas de la Provincia* de Buenos Aires.

Table 21. Description of variables in 9 de Julio

Variable	Description
Paid total tax bill	Dummy = 1 if the taxpayer has not only paid in full the total tax liabilities for the period but she has also made some additional payments
Local public services	First principal component of the public services provided by the municipality of 9 de Julio (e.g. streetlights, street cleaning and paved roads)
Surface	Average square meters of all properties associated to the taxpayer
Linear front meters	Average linear front meters of all the properties associated to the taxpayer
Properties	Number of registered properties associated to the taxpayer
Tax liability	Total amount billed to the taxpayer in August 2012
Tax payments	Total amount paid by the taxpayer in the period
Tax compliance ratio	Tax payments over tax liabilities ratio (in percent)
Debtor	Dummy =1 if the taxpayer has at least one remaining unpaid bill
Debt stock	Accumulated stock of debt
Unrecoverable Debtor	Dummy=1 if the taxpayer did not pay any tax bill between 2010 and January, 2012

Table 22. Messages included in the Tax Bill in 9 de Julio

#	Message	Text of the message	Images
1	Deterrence	Did you know that if you do not pay on time the TRSU for a debt of, for example, 1,000 pesos you will have to pay 269 additional pesos in arrears by the end of the year, and we can take administrative actions?	
2	Monitoring	I wanted to communicate you that from the Office of Public Revenues of 9 de Julio we are monitoring more closely compliance and payment delays with the TRSU.	
3	Interaction	Deterrence message + Monitoring letter	 
4	Placebo message	A clean city is the least soiled. Collaborate, keep clean 9 de Julio	No image

Table 23. Baseline Summary Statistics and Balance of Randomization in Pre-experimental Period

	(1) Control Group	(2) Difference: Deterrence	(3) Difference: Deterrence + Detection	(4) Difference: Detection
Paid total tax bill	0.651***	-0.001	0.007	-0.001
	(0.007)	(0.013)	(0.013)	(0.013)
Log (Tax payments)	4.283***	-0.004	-0.001	0.015
	(0.016)	(0.027)	(0.027)	(0.027)
Log (Tax liabilities)	4.390***	0.006	-0.008	0.004
	(0.010)	(0.018)	(0.018)	(0.018)
Log (Number of properties)	0.845***	-0.002	-0.006	-0.000
	(0.005)	(0.008)	(0.008)	(0.008)
Paved roads (August)	0.993***	-0.001	-0.000	0.001
	(0.001)	(0.002)	(0.002)	(0.002)
Streetlights (August)	0.857***	-0.006	-0.007	-0.012
	(0.005)	(0.008)	(0.008)	(0.008)
Street cleaning (August)	0.662***	0.010	0.012	0.015
	(0.007)	(0.012)	(0.012)	(0.012)
Debtor	0.484***	0.004	-0.009	-0.001
	(0.008)	(0.013)	(0.013)	(0.013)

Notes: Each row shows a regression of the pre-treatment variable in question on treatment dummies and a constant term. Observations are presented for the average six-month period prior to treatment (August 2011-January 2012). The constant captures the average value for the control group (no message).. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 24. Average Treatment Effects of Receiving a Message and/or a Letter (August)

	(1)	(2)
T0: Taxpayer received a message/letter	0.029***	0.035**
	(0.011)	(0.015)
Lagged outcome		0.449***
		(0.044)
Local public services		0.001
		(0.004)
Log (Number of properties)		-0.010***
		(0.001)
Fixed effects	Yes	Yes
Observations	10,343	10,165

Notes: Regressions using a linear probability model of the probability of paying 100% or more of total tax liabilities in the August billing period. Colum (1) includes the three treatment messages, the lagged outcome variable, and block-level fixed effects. Colum (2) includes the principal component of the variables for public service provision (street cleaning, street lightning and paved roads services during the period), and the (log of the) number of properties that each taxpayer has. Robust standard errors in parentheses. .*** p<0.01, ** p<0.05, * p<0.1.

Table 25. Average Effects of the Treatment Messages (August)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1: Deterrence Message	0.039**	0.041**	0.045**	0.049***	0.048***	0.044***	0.049***	0.048***
	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)	(0.016)	(0.017)	(0.017)
T2: T1 x T3	0.022	0.025	0.025	0.030*	0.030*	0.016	0.020	0.020
	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)	(0.016)	(0.017)	(0.017)
T3: Detection Letter	0.034**	0.035**	0.030*	0.031*	0.032*	0.035**	0.030*	0.030*
	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)
Observations	10,230	10,230	10,225	10,225	10,225	8,674	8,670	8,670
Model	Baseline	Baseline	Controls	Controls	Controls	Baseline	Controls	Controls
Sample	Whole	Whole	Whole	Whole	Whole	Non-defaulters	Non-defaulters	Non-defaulters
Fix effects	No	Yes	No	Yes	Yes	Yes	Yes	Yes

Notes: Regressions using a probit model of the probability of paying 100% or more of tax liabilities in the August billing period. The first set of regressions (baseline) include the three treatment messages, the lagged outcome variable, and block-level fixed effects. The second set of regressions (controls) include also the principal component of the variables for public service provision (trash collection, paved roads and street lightning services during the period), the (log of the) number of properties that each taxpayer has, and the (log of the) average tax liability. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 26. Baseline Summary Statistics and Balance of Randomization in Pre-experimental Period, Excluding Unrecoverable Debtors

	(1) Control Group	(2) Difference: Deterrence	(3) Difference: Deterrence + Detection	(4) Difference: Detection
Paid total tax bill	0.593***	0.010	0.000	0.000
	(0.008)	(0.012)	(0.012)	(0.012)
Log (Tax payments)	4.283***	-0.003	0.000	0.013
	(0.015)	(0.027)	(0.027)	(0.027)
Log (Tax liabilities)	4.435***	0.014	-0.022	-0.002
	(0.012)	(0.021)	(0.020)	(0.020)
Log (Surface)	5.594***	0.017	0.005	0.012
	(0.020)	(0.034)	(0.034)	(0.034)
Log (number of properties)	0.865***	0.000	-0.012	-0.003
	(0.006)	(0.010)	(0.010)	(0.010)
Paved roads	0.994***	-0.001	-0.001	0.000
	(0.001)	(0.002)	(0.002)	(0.002)
Streetlights	0.854***	-0.007	-0.007	-0.012
	(0.005)	(0.009)	(0.009)	(0.009)
Street cleaning	0.700***	0.018	0.014	0.018
	(0.007)	(0.012)	(0.012)	(0.012)
Debtor	0.395***	-0.007	-0.008	-0.007
	(0.008)	(0.014)	(0.014)	(0.014)

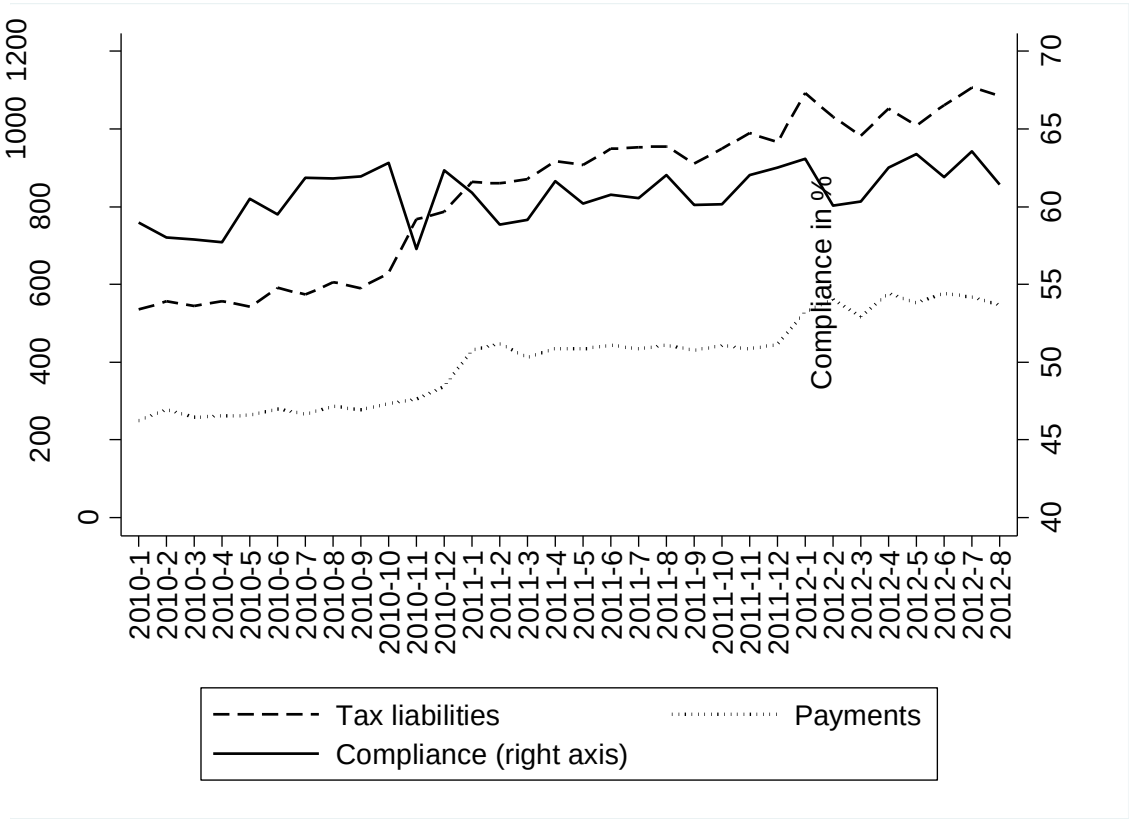
Notes: Each row shows a regression of the pre-treatment variable in question on treatment dummies and a constant term. Observations are presented for the average six-month period prior to treatment (August 2011-Janurary 2012), excluding unrecoverable debtors. The constant captures the average value for the control group (no message). Monetary amounts are in Argentine Pesos (AR$). Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 27. Treatment Effects by Strata (August)

	(1) One property, non-complier	(2) One property, complier	(3) Multiple properties, non-complier	(4) Multiple properties, complier
T1: Deterrence Message	0.037*	0.018	0.058*	0.032
	(0.022)	(0.015)	(0.030)	(0.035)
T2: T1 X T3	0.047**	0.009	0.031	-0.018
	(0.023)	(0.016)	(0.029)	(0.037)
T3: Detection Letter	0.056**	0.004	0.061**	-0.029
	(0.023)	(0.016)	(0.030)	(0.036)
Observations	3,215	4,180	1,664	1,111
Model	Controls	Controls	Controls	Controls
Sample	Whole	Whole	Whole	Whole
Fix effects	Yes	Yes	Yes	Yes

Notes: Regressions using a probit model of the probability of paying 100% or more of tax liabilities in the August billing period. All the regressions include the three treatment messages, the lagged outcome variable, and block-level fixed effects, the variables for public service provision (trash collection, paved roads and street lightning services during the period), the (log of the) number of properties that each taxpayer has, the (log of the) average linear front size of the properties, and a dummy that controls for those taxpayers who selected to be in a repayment plan.

Figure 13. Tax liabilities, payments and compliance ratio (AR$ millions and percent), 2010-2012



Source: own elaboration based on administrative information provided by the Municipality of 9 de Julio

Figure 14. Tax Compliance by Treatment Groups and Control Group (in percent) (August)

Notes: This figure plots the tax compliance ratio (tax payments over tax liabilities) for the treatment groups and the control. It also displays –in a solid line- the average tax compliance ratio prior to the experiment (January 2012) as well as –in a dotted line- the average tax compliance ratio in the same period (August) but one year before (2011). The difference in compliance between the group that received the deterrence message and the control is statistically significant at the 1% level according to a proportions test. Differences are not statistically significant for the other two groups.
Source: own elaboration based on administrative information provided by the Municipality of 9 de Julio

Appendix 1: Supporting documentation

THE JOURNAL OF
INDUSTRIAL
ECONOMICS

Journal of Industrial Economics
Department of Economics
University of Essex
Wivenhoe Park
Colchester, Essex CO4 3SQ
UK

E-mail: jindec@essex.ac.uk

JIE 2012033-3 October 3, 2013

Dear Julian

I am happy to accept your paper "Does energy consumption respond to price shocks? Evidence from a regression-discontinuity design" for publications in the Journal of Industrial Economics.

There are just a few typos and omissions that you should address:

1. Add note to Figures 1- 4 and 6 with labels for the lines, dashed lines and dots.
2. Correct second "Place Figure 4 ..." to Figure 5.
3. In page 16, explain how you got elasticity of -0.15 (-0.038/.25) and delete a dot after the word "United States".

Please send me the revised version of the paper as soon as possible. I completed my tenure as editor of the Journal last July and would like to see the acceptance process for this paper finalized.

Looking forward to receiving the final version of the paper,

Best regards,

Saul Lach

Journal of Industrial Economics

Banco Interamericano de Desarrollo
Banco Interamericano de Desenvolvimento

Inter-American Development Bank
Banque Interaméricaine de développement

Departamento de Investigación

Research Department

Washington, DC, April 19th 2013

Dear Professor Andrew Newell,

I'm writing to you to certify that Mr. Lucio Castro did a significant and relevant proportion of work as co-author of our paper "*Does Energy Consumption Respond to Price Shocks? Evidence from a Regression-Discontinuity Design*" that has been conditionally accepted for publication at the Journal of Industrial Economics (JIE).

Please do not hesitate to contact me for any further reference.

Best regards,

Carlos G. Scartascini
Research Department
Inter-American Development Bank
1300 New York Ave. NW
Washington, DC 20577
www.cscartascini.org

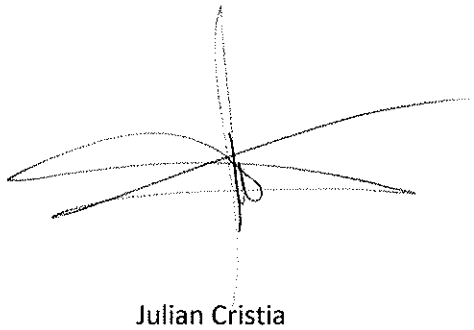
Washington, DC 19 April, 2013

Dear Andrew Newell,

I'm writing to you to certify that Mr. Lucio Castro did a significant and relevant proportion of work as co-author of our paper "*Does Energy Consumption Respond to Price Shocks? Evidence from a Regression-Discontinuity Design*" that has been conditionally accepted for publication at the Journal of Industrial Economics (JIE).

Please do not hesitate to contact me for any further reference.

Best regards,



Julian Cristia
Senior Research Economist
Research Department
Inter-American Development Bank

The World Bank Group

Washington DC, 22 April 2013

Dear Professor Andrew Newell,

I am writing to certify that Mr. Lucio Castro did a significant and relevant proportion of work as co-author of our paper "Does energy consumption respond to price shocks? Evidence from a regression discontinuity design", conditionally accepted for publication at the *Journal of Industrial Economics* (JIE).

Please do not hesitate to contact me for any further reference.

Best regards,

Paulo Bastos
Economist
Development Research Group
The World Bank

Appendix 2. Additional regressions and figures

Table 1: Robustness check - Impacts of price increase in bill 0 on consumption in period 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	-15.901**	-17.415***	-15.266**	-16.637***	-15.577**	-17.013***	-15.597*	-16.670*	-16.296*	-18.000**	-14.780**	-16.371**
	(6.491)	(6.342)	(5.929)	(5.656)	(6.153)	(5.934)	(9.194)	(8.996)	(8.105)	(7.922)	(6.882)	(6.583)
AAC	0.720**	0.822***	1.004**	1.169***	0.707**	0.806***	2.201	2.547*	0.759	0.876*	1.443**	1.603**
	(0.290)	(0.272)	(0.412)	(0.368)	(0.272)	(0.251)	(1.441)	(1.282)	(0.451)	(0.450)	(0.612)	(0.604)
AAC * Treatment			-0.619	-0.753			-2.777	-3.335*			-1.522*	-1.620**
			(0.531)	(0.493)			(1.968)	(1.823)			(0.782)	(0.783)
AAC^2					-0.011	-0.013	0.059	0.068				
					(0.012)	(0.011)	(0.069)	(0.061)				
AAC^2 * Treatment							-0.014	-0.011				
							(0.090)	(0.086)				
Constant	433.271***	461.555***	436.134***	464.847***	434.620***	463.133***	439.962***	468.975***	433.000***	465.119***	438.133***	470.180***
	(3.299)	(33.802)	(3.992)	(33.767)	(3.539)	(33.715)	(6.161)	(35.078)	(4.042)	(43.858)	(4.647)	(44.095)
Bandwidth	20	20	20	20	20	20	20	20	15	15	15	15
Areas Effects	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
N	7190	7190	7190	7190	7190	7190	7190	7190	5417	5417	5417	5417
R-squared	0.001	0.056	0.001	0.056	0.001	0.056	0.001	0.057	0.001	0.057	0.001	0.058

Notes: The dependent variable is consumption in Period 1. The bandwidths determine the range of normalized annual accumulated consumption used to select observations for a particular regression. For example, in Columns (1) and (2) only individuals with normalized annual accumulated consumption between [-20, 20] are included. Average of the dependent variable is 425.49. The estimation method is OLS. Standard errors clustered by accumulated consumption in Period 0. Even columns present regressions with areas (neighborhoods) dummies. *** p<0.01, ** p<0.05, * p<0.1.

Table 2. Dependent variable Paid_by1D - Junin

VARIABLES	(1) Paid_by1D	(2) Paid_by1D	(3) Paid_by1D	(4) Paid_by1D	(5) Paid_by1D	(6) Paid_by1D	(7) Paid_by1D	(8) Paid_by1D
T1: Deterrence	0.017*	0.017*	0.017*	0.017*	0.017*	0.022	0.023	0.023
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)
T2: Fairness	-0.004	-0.005	-0.006	-0.006	-0.005	-0.004	-0.005	-0.004
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)
T3: Equity	0.006	0.005	0.005	0.004	0.004	0.003	0.002	0.002
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)
Paid_by1D (t-1)	0.600***	0.563***	0.562***	0.550***	0.549***	0.614***	0.601***	0.599***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)	(0.010)
Streetlights			0.015***	0.003	0.003		0.001	0.000
			(0.003)	(0.004)	(0.004)		(0.005)	(0.005)
Trash Collection			0.120***	0.096***	0.100***		0.101***	0.116***
			(0.007)	(0.009)	(0.009)		(0.013)	(0.013)
Monthly payments			0.038	0.028	0.029		-0.051	-0.048
			(0.030)	(0.030)	(0.030)		(0.081)	(0.082)
Log(Number of Properties)			-0.036***	-0.062***	-0.047***		-0.143***	-0.075***
			(0.012)	(0.012)	(0.015)		(0.016)	(0.024)
Log(Lineal Front Meters)			-0.036***	-0.018**	-0.011		-0.022*	0.010
			(0.008)	(0.008)	(0.009)		(0.012)	(0.015)
Log(Tax Liability)					-0.010			-0.042***
					(0.006)			(0.011)
Observations	23,195	23,195	23,186	23,186	23,186	15,845	15,837	15,837
Model	Baseline	Baseline	Controls	Controls	Controls	Baseline	Controls	Controls
Sample	Whole	Whole	Whole	Whole	Whole	Non-defaulters	Non-defaulters	Whole
Fix effects	No	Yes	No	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3. Dependent variable Paid_by2D - Junin

VARIABLES	(1) Paid_by2D	(2) Paid_by2D	(3) Paid_by2D	(4) Paid_by2D	(5) Paid_by2D	(6) Paid_by2D	(7) Paid_by2D	(8) Paid_by2D
T1: Deterrence	0.031**	0.030**	0.031**	0.031**	0.031**	0.038**	0.039**	0.039**
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.016)	(0.016)	(0.016)
T2: Fairness	-0.002	-0.002	-0.004	-0.004	-0.004	-0.002	-0.004	-0.003
	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)	(0.016)
T3: Equity	0.005	0.004	0.004	0.004	0.004	0.003	0.002	0.002
	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)	(0.016)
Paid_by2D (t-1)	0.840***	0.809***	0.809***	0.797***	0.797***	0.806***	0.794***	0.792***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
Streetlights			0.014***	0.008*	0.008		0.008	0.007
			(0.004)	(0.005)	(0.005)		(0.006)	(0.006)
Trash Collection			0.092***	0.081***	0.088***		0.079***	0.101***
			(0.009)	(0.011)	(0.012)		(0.015)	(0.016)
Monthly payments			0.108***	0.104***	0.105***		0.213***	0.218***
			(0.040)	(0.040)	(0.040)		(0.082)	(0.083)
Log(Number of Properties)			-0.050***	-0.070***	-0.034*		-0.157***	-0.053**
			(0.015)	(0.015)	(0.019)		(0.017)	(0.027)
Log(Lineal Front Meters)			-0.030***	-0.021**	-0.005		-0.029**	0.020
			(0.009)	(0.010)	(0.011)		(0.013)	(0.016)
Log(Tax Liability)					-0.023***			-0.064***
					(0.008)			(0.013)
Observations	23,195	23,195	23,186	23,186	23,186	15,845	15,837	15,837
Model	Baseline	Baseline	Controls	Controls	Controls	Baseline	Controls	Controls
Sample	Whole	Whole	Whole	Whole	Whole	Non-defaulters	Non-defaulters	Whole
Fix effects	No	Yes	No	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Dependent variable Paid - Junin

VARIABLES	(1) Paid	(2) Paid	(3) Paid	(4) Paid	(5) Paid	(6) Paid	(7) Paid	(8) Paid
T1: Deterrence	0.046***	0.047***	0.047***	0.048***	0.047***	0.043***	0.044***	0.044***
	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)
T2: Fairness	-0.000	-0.001	-0.002	-0.003	-0.003	0.001	-0.000	0.001
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)
T3: Equity	-0.009	-0.009	-0.008	-0.008	-0.009	-0.017	-0.017	-0.016
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)
Paid (t-1)	1.041***	1.011***	1.014***	1.001***	1.007***	0.876***	0.864***	0.861***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Streetlights			0.010**	0.003	0.003		0.000	-0.000
			(0.005)	(0.005)	(0.005)		(0.006)	(0.006)
Trash Collection			0.092***	0.071***	0.058***		0.064***	0.080***
			(0.011)	(0.013)	(0.014)		(0.014)	(0.016)
Monthly payments			0.159***	0.153***	0.150***		0.339***	0.341***
			(0.053)	(0.053)	(0.053)		(0.087)	(0.086)
Log(Number of Properties)			-0.033*	-0.061***	-0.121***		-0.136***	-0.066**
			(0.017)	(0.018)	(0.024)		(0.017)	(0.030)
Log(Lineal Front Meters)			-0.039***	-0.028**	-0.053***		-0.022*	0.013
			(0.010)	(0.012)	(0.013)		(0.013)	(0.018)
Log(Tax Liability)					0.039***			-0.045***
					(0.010)			(0.016)
Observations	23,176	23,176	23,168	23,168	23,168	15,831	15,823	15,823
Model	Baseline	Baseline	Controls	Controls	Controls	Baseline	Controls	Controls
Sample	Whole	Whole	Whole	Whole	Whole	Non-defaulters	Non-defaulters	Whole
Fix effects	No	Yes	No	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Dependent variable Overpaid - Junin

VARIABLES	(1) Overpaid	(2) Overpaid	(3) Overpaid	(4) Overpaid	(5) Overpaid	(6) Overpaid	(7) Overpaid	(8) Overpaid
T1: Deterrence	0.007*	0.007**	0.007*	0.007*	0.007*	0.010*	0.009*	0.009*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
T2: Fairness	-0.004	-0.004	-0.004	-0.004	-0.003	-0.005	-0.005	-0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
T3: Equity	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Overpaid (t-1)	0.035***	0.032***	0.032***	0.031***	0.031***	0.031***	0.031***	0.031***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)
Streetlights			0.001	0.001	0.001		0.003	0.002
			(0.001)	(0.001)	(0.001)		(0.002)	(0.002)
Trash Collection			0.008***	0.007**	0.009***		-0.000	0.005
			(0.003)	(0.003)	(0.003)		(0.005)	(0.005)
Monthly payments			0.011	0.010	0.011		0.030	0.032
			(0.010)	(0.010)	(0.010)		(0.026)	(0.026)
Log(Number of Properties)			-0.003	-0.005	0.005		-0.015**	0.009
			(0.004)	(0.004)	(0.005)		(0.006)	(0.009)
Log(Lineal Front Meters)			-0.014***	-0.015***	-0.010***		-0.025***	-0.013**
			(0.003)	(0.003)	(0.004)		(0.005)	(0.006)
Log(Tax Liability)					-0.006***			-0.015***
					(0.002)			(0.004)
Observations	23,174	23,073	23,166	23,065	23,065	15,794	15,786	15,786
Model	Baseline	Baseline	Controls	Controls	Controls	Baseline	Controls	Controls
Sample	Whole	Whole	Whole	Whole	Whole	Non-defaulters	Non-defaulters	Whole
Fix effects	No	Yes	No	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6. Dependent variable Additionalpayments - Junin

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Additional Payments	Additional Payments	Additional Payments	Additional Payments	Additional Payments	Additional Payments	Additional Payments	Additional Payments
T1: Deterrence	0.011**	0.011***	0.010**	0.010**	0.010**	0.015**	0.014**	0.014**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
T2: Fairness	-0.004	-0.004	-0.003	-0.003	-0.003	-0.006	-0.004	-0.004
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
T3: Equity	0.007	0.006	0.007	0.007	0.007	0.009	0.009	0.009
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
Additional Payments (t-1)	0.061***	0.056***	0.054***	0.053***	0.053***	0.060***	0.058***	0.058***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Streetlights			0.001	0.001	0.001		0.001	0.001
			(0.001)	(0.002)	(0.002)		(0.002)	(0.002)
Trash Collection			0.008**	0.004	0.006		-0.006	-0.001
			(0.003)	(0.004)	(0.004)		(0.005)	(0.006)
Monthly payments			0.019*	0.019*	0.019*		0.037	0.038
			(0.010)	(0.010)	(0.010)		(0.026)	(0.026)
Log(Number of Properties)			0.029***	0.026***	0.033***		0.029***	0.050***
			(0.004)	(0.004)	(0.005)		(0.006)	(0.010)
Log(Lineal Front Meters)			-0.013***	-0.012***	-0.009**		-0.019***	-0.009
			(0.003)	(0.004)	(0.004)		(0.006)	(0.007)
Log(Tax Liability)					-0.004*			-0.013***
					(0.003)			(0.005)
Observations	23,313	23,211	23,182	23,081	23,081	15,915	15,800	15,800
Model	Baseline	Baseline	Controls	Controls	Controls	Baseline	Controls	Controls
Sample	Whole	Whole	Whole	Whole	Whole	Non-defaulters	Non-defaulters	Whole
Fix effects	No	Yes	No	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1. Marginal effects for T1 at different levels of covariates – Junin

Figure 1a. Marginal effects for deterrence at different levels of trash collection

Figure 1b. Marginal effects for deterrence at different property sizes

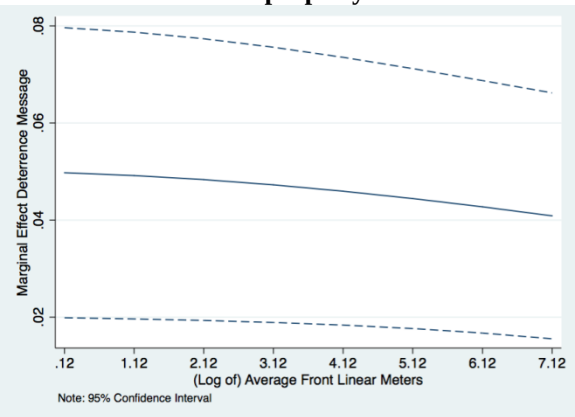


Figure 1c. Marginal effects for deterrence at different number of properties

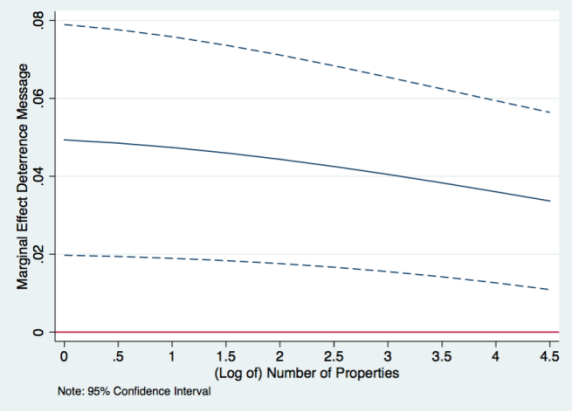


Table 7. Probit regressions. Interaction models. Junin.

VARIABLES	(1) Paid	(2) Paid	(3) Paid	(4) Paid
T1: Deterrence	0.152***	0.069	0.066	0.497**
	(0.047)	(0.060)	(0.109)	(0.197)
T2: Fairness	0.028	-0.019	-0.061	0.431**
	(0.050)	(0.058)	(0.098)	(0.201)
T3: Equity	0.059	-0.149***	-0.297***	-0.079
	(0.049)	(0.055)	(0.087)	(0.205)
Paid (t-1)	2.667***	1.355***	0.853***	2.629***
	(0.030)	(0.047)	(0.051)	(0.025)
Streetlights	0.008	-0.001	0.028**	0.009
	(0.014)	(0.015)	(0.013)	(0.014)
Trash Collection	0.187***	0.165***	0.165***	0.185***
	(0.034)	(0.036)	(0.035)	(0.034)
Monthly payments	0.404***	1.241***	0.974***	0.402***
	(0.139)	(0.092)	(0.098)	(0.139)
Log(Number of Properties)	-0.159***	-0.135***	0.032	-0.159***
	(0.046)	(0.048)	(0.053)	(0.046)
Log (Lineal Front Meters)	-0.074**	-0.018		-0.045
	(0.030)	(0.031)		(0.034)
Debtor		-1.543***	-1.321***	
		(0.051)	(0.049)	
Debtor Stock			-0.173***	
			(0.008)	
T1: Deterrence x paid(t-1)	-0.066			
	(0.080)			
T2: Fairness x paid(t-1)	-0.086			
	(0.078)			
T3: Equity x paid(t-1)	-0.187**			
	(0.076)			
T1: Deterrence x debtor		0.093		
		(0.079)		
T2: Fairness x debtor		0.034		
		(0.081)		
T3: Equity x debtor		0.242***		
		(0.077)		
T1: Deterrence x debt stock			0.010	
			(0.018)	
T2: Fairness x debt stock			0.012	
			(0.016)	
T3: Equity x debt stock			0.054***	
			(0.015)	
T1: Deterrence x meters				-0.144*
				(0.074)
T2: Fairness x meters				-0.170**
				(0.076)
T3: Equity x meters				0.022
				(0.078)
Constant	-1.220***	0.055	0.865***	-1.265***
	(0.199)	(0.217)	(0.198)	(0.202)
Observations	23,168	23,168	23,168	23,168
Interaction	Lagged Dep. Var.	Debtor	Debt Stock	Linear Front Meters
Model	Controls	Controls	Controls	Controls
Sample	Whole	Whole	Whole	Whole
Fix effects	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

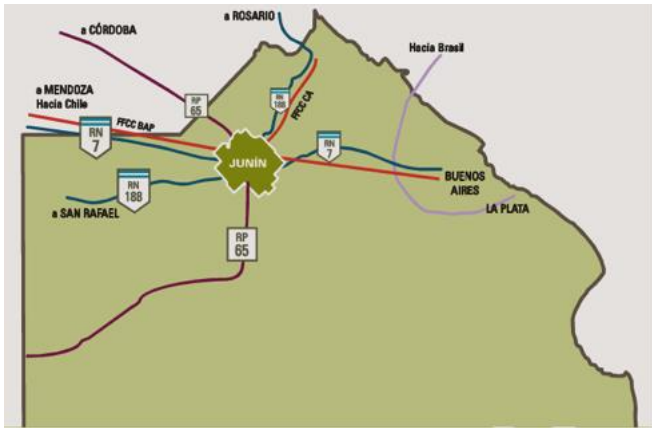
Table 8. Average treatment effects for the whole sample and non-defaulters in August - 9 de Julio

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1: Deterrence Message	0.039**	0.041**	0.045**	0.049***	0.048***	0.044***	0.049***	0.048***
	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)	(0.016)	(0.017)	(0.017)
T2: T1 x T3	0.022	0.025	0.025	0.030*	0.030*	0.016	0.020	0.020
	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)	(0.016)	(0.017)	(0.017)
T3: Detection Letter	0.034**	0.035**	0.030*	0.031*	0.032*	0.035**	0.030*	0.030*
	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)
Lagged outcome	0.759***	0.545***	0.612***	0.466***	0.463***	0.402***	0.342***	0.344***
	(0.013)	(0.016)	(0.015)	(0.018)	(0.018)	(0.017)	(0.018)	(0.018)
Locally provided services			0.283***	0.272***	0.274***		0.248***	0.249***
			(0.017)	(0.017)	(0.017)		(0.015)	(0.016)
Log(Number of Properties)			-0.069***	-0.007	-0.121***		-0.016	-0.095**
			(0.023)	(0.038)	(0.045)		(0.033)	(0.040)
Log(Tax Liability)					0.076***			0.054***
					(0.015)			(0.013)
Observations	10,230	10,230	10,225	10,225	10,225	8,674	8,670	8,670
Model	Baseline	Baseline	Controls	Controls	Controls	Baseline	Controls	Controls
Sample	Whole	Whole	Whole	Whole	Whole	Non-defaulters	Non-defaulters	Non-defaulters
Fix effects	No	Yes	No	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

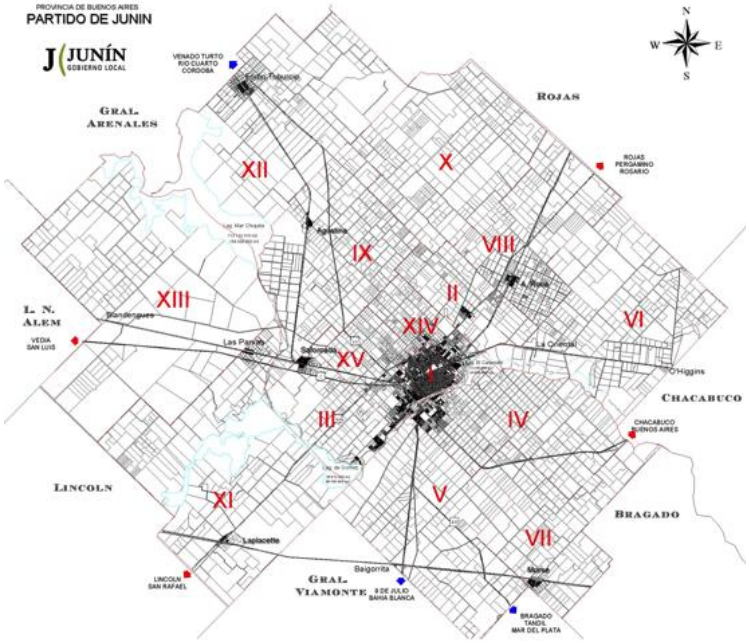
Appendix 3. Maps

Map 1. Locational map of Junín within the Province of Buenos Aires



Source: http://www.junin.gov.ar/archivos/mapas.html

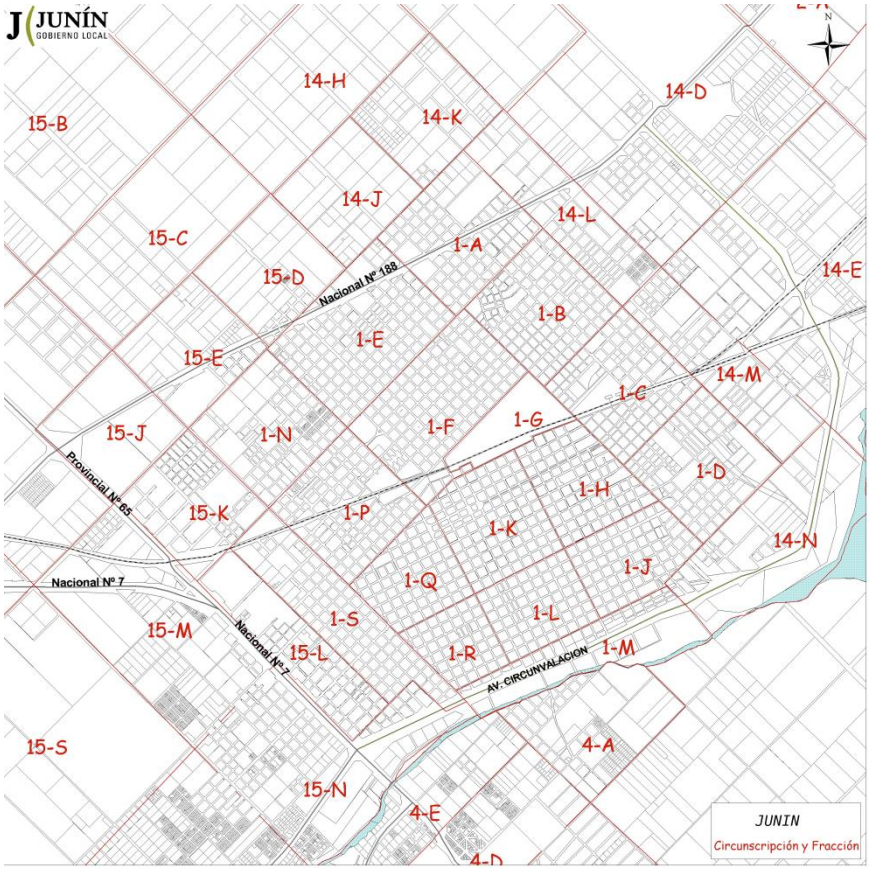
Map 2. Map of Junín by District



Note: this figure presents the map of the 15 administrative districts in which Junín is divided. District I represents the largest district and capital of the municipality, the city of Junín. Names of adjoining municipalities are also display in caps.

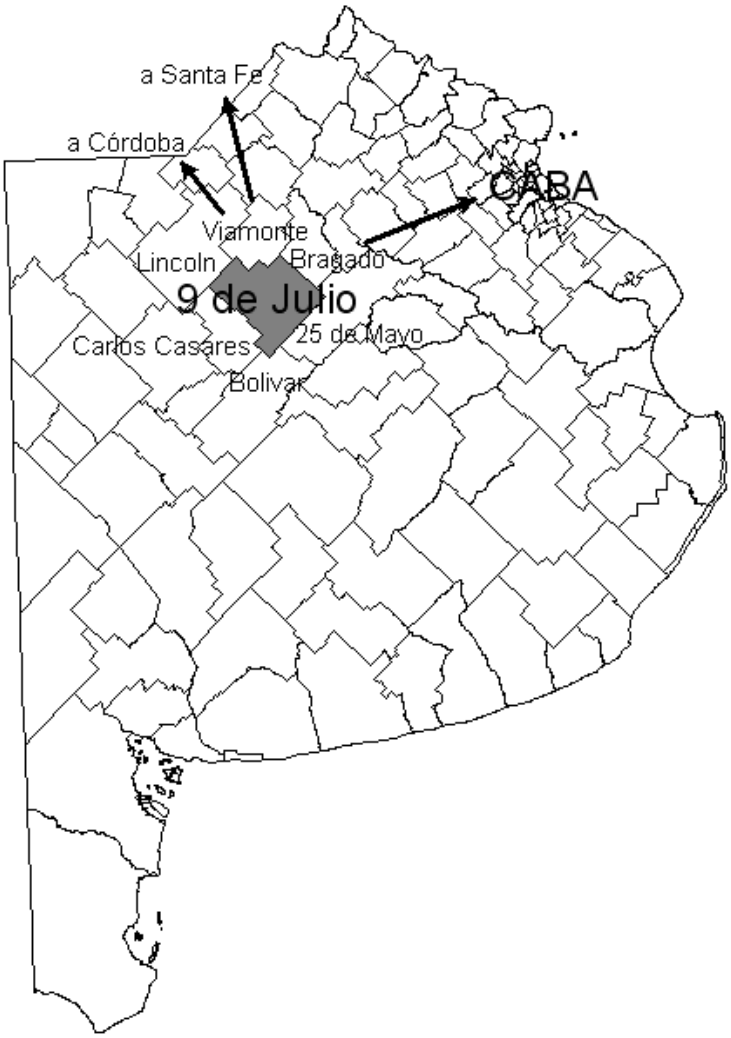
Source: http://www.junin.gov.ar/archivos/mapas.html

Map 3. A closer look into some of the districts of Junin



Source: http://www.junin.gov.ar/archivos/mapas.html

Map 4. Locational map of 9 de Julio within the Province of Buenos Aires



Notes: the map shows 9 de Julio (in grey) and its surrounding districts in the province of Buenos Aires. The arrows point to the direction of the three largest cities of Argentina: City of Buenos Aires (CABA by its acronym in Spanish), Córdoba and Rosario.

Source: own elaboration based on maps provided by the Municipality of 9 de Julio.

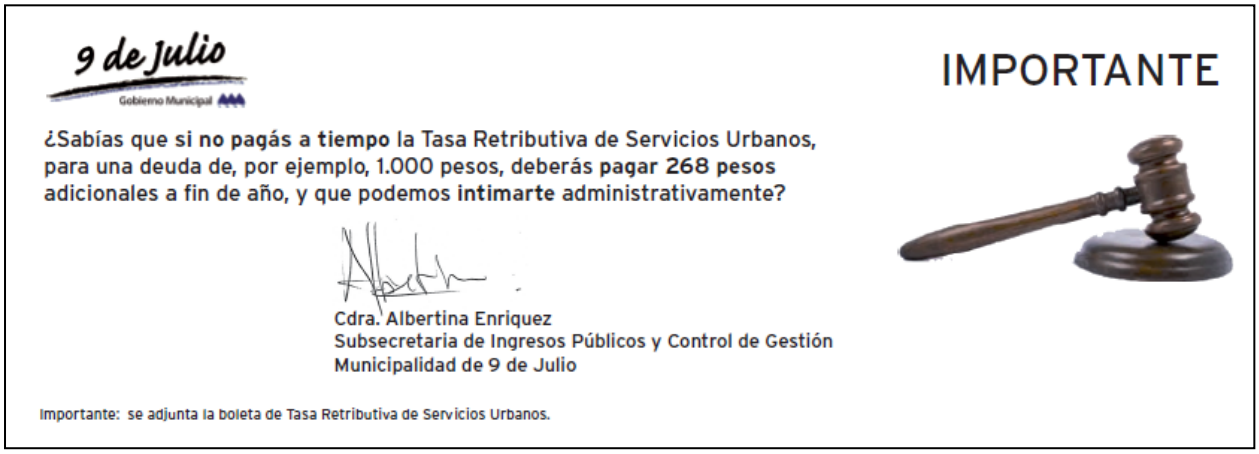
Appendix 4. Sample Tax Bills with Treatment Messages (in Spanish)

Image 1. Sample Tax Bill – Deterrence Treatment - Junin

Image 2. Sample Tax Bill – Fairness Treatment - Junin

Image 3. Sample Tax Bill – Equity Treatment - Junin

Image 4. Sample Deterrence Message – 9 de Julio



English Translation

IMPORTANT

Did you know that **if you do not pay on time** the TRSU for a debt of, for example, 1,000 pesos **you will have to pay 269 additional pesos in arrears by the end of the year, and we can take administrative actions**?

Accountant Albertina Enriquez

Under-secretary of Revenue and Control

Municipality of 9 de Julio

Important: the TRSU bill is attached.

Image 5. Sample Monitoring Letter (in Spanish) – 9 de Julio

9 de Julio
Gobierno Municipal

Notificación:

Estamos monitoreando el pago de TRSU

Contribuyente: [redacted]
Partida: [redacted]
Domicilio: [redacted]
Localidad: [redacted]
Código postal: [redacted]

Estimado/a [redacted]:

Quería comunicarte que en la Oficina de Ingresos Públicos de 9 de Julio, estamos **monitoreando más de cerca la morosidad y el pago de la Tasa Retributiva de Servicios Urbanos (TRSU)**.

Esperamos contar con tu colaboración.

Saludos cordiales,

Cdra. Albertina Enriquez
Subsecretaria de Ingresos
Públicos y Control de Gestión
Municipalidad de 9 de Julio

English Translation

Notification

We are monitoring compliance with the TRSU

Name:

Property:

Address:

Town/City:

Zip code:

Dear Mr/Ms/Miss

I want to communicate you that from the Office of Public Revenues of 9 de Julio **we are monitoring more closely compliance and payment delays with the TRSU**.

We look forward your cooperation.

Best regards,

Accountant Albertina Enriquez

Under-secretary of Revenue and Control

Municipality of 9 de Julio

Image 7. Sample Placebo Message – 9 de Julio

English Translation

A clean city is the least soiled. Collaborate, keep clean 9 de Julio

Important: the TRSU bill is attached.