Futures prices, trade and domestic supply of agricultural commodities

Maximiliano Méndez Parra

Thesis submitted for the Degree of Doctor of Philosophy

March 2015

Department of Economics

University of Sussex

WORK NOT SUBMITTED ELSEWHERE FOR EXAMINATION

I hereby declare that this thesis has not been, and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:

UNIVERSITY OF SUSSEX

Maximiliano Méndez Parra

DEGREE OF DOCTOR OF PHILOSOPHY

FUTURES PRICES, TRADE AND DOMESTIC SUPPLY OF AGRICULTURAL COMMODITIES

SUMMARY

Commodity markets display substantial volatility both in prices and in the quantities traded. This has led to the development of different instruments designed to address this volatility. Processors and traders, who are actively involved in the international market, participate in these commodity markets using cross-hedging strategies by their export and domestic supply decisions. Spot and future prices, as well as the cross-hedging strategies, affect export and the domestic supply decisions. Understanding this complex interaction calls for further and newer insights and this research contributes to this.

The primary objective of Chapter 1 of this thesis is to develop a model which explains the export and domestic supply decisions when traders, producers and speculators participate in a futures market for a primary commodity, which can be stored and for which future markets operate. As a result, exports and domestic supply are affected by the prices of the primary product, and jointly by the prices in the external and domestic market. Chapter 2 provides the historical, political and economic context of the Argentine economy and the agricultural sector, specifically on the three agricultural commodities used in the empirical part of this research. In Chapter 3, we perform a comprehensive analysis of the seasonal unit roots of monthly series of exports and domestic supply, using time series that include zero values. In the past, this technique has mostly been applied to quarterly data but never to monthly series that display periods of inactivity. The results indicate that, in general, the seasonality observed in the series analysed can be sufficiently explained by a deterministic approach. The estimation and further analysis of the supply equations derived in Chapter 1 are undertaken in Chapter 4. A comprehensive analysis of seasonal cointegration using monthly data was conducted but, in light of the results obtained in Chapter 3, only the Engle-Granger cointegration is applied. The results

indicate weak cointegration relationships. This may indicate the need for improved data and/or alternative econometric techniques.

ACKNOWLEDGEMENTS

Many people have been of enormous importance in the completion of this document. I am particularly grateful to my supervisors, Sherman Robinson and Andy Newell, who assisted me particularly during the early and final stages of this research, and especially to Andy McKay for his continuous support, guidance, stimulating insights, understanding and encouragement throughout. None of this would have been possible without their permanent support.

I also want to express my gratitude to the Alban Programme of the European Commission for providing financial support in the early stages of this research. The Economics Department at the University of Sussex has been an auspicious and academically stimulating place for the development of my work. I would also like to thank Barry Reilly for his guidance on some technical issues in this research. I am indebted to the faculty for providing a stimulating academic environment and support throughout. I also thank Chiara, Stefania, Goshia, Marta, Paola, Outi, Alvaro, Gonzalo, Martin, Jairo, Antonio, and the rest of the "younger and older" DPhil students for being part of an amenable and friendly space for research. In particular, I would like to thank Javier Lopez Gonzalez, who has walked with me through this process since the first day, and with whom I shared many worries, insights, professional and business endeavours, academic, linguistic (how to speak the Spanish language properly) and other fundamental discussions (football, mainly). I would also like to thank The Centre for the Analysis of Regional Integration at Sussex (CARIS) and its affiliates for giving me the opportunity to engage in consultancy activities which have funded part of this thesis.

I would like to express my gratitude to Jim Rollo, Peter Holmes and Michael Gasiorek for their continuous encouragement and support, particularly during the more complicated moments of this research.

I could not be more grateful to my parents and my sister Beta for their support and understanding, both at a distance and *in situ*; I would like to apologise if, in the final stages of this research, I have been distant, grumpy and annoying.

Finally, all my gratitude to my friend and wife Candela, for providing me with so much love and support in this, my personal endeavour. I know how much she has sacrificed and postponed in order to be there for me. I dedicate this thesis to her.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

LIST OF FIGURES

CHAPTER THREE

CHAPTER FOUR

LIST OF TABLES IN THE APPENDICES

LIST OF ABBREVIATIONS AND ACRONYMS

ADF	Augmented Dickey-Fuller
ACF	Autocorrelation Function
ADL	Autoregressive Distributed Lag
AIC	Akaike Information Criterion
AO	Additive outlier
AR	Autoregressive
ARMA	Autoregressive moving average
BG	Breusch-Godfrey
BIC	Bayesian Information Criterion
BJ	Bera-Jarque
CGE	Computable General Equilibrium
CPI	Consumer Price Index
DF	Dickey-Fuller
DW	Durbin-Watson
DGP	Data Generation Process
ECM	Error Correction Mechanism
EG	Engle-Granger
EGHL	Engle, Granger, Hylleberg, Lee
FAO	Food and Agriculture Organisation
FOB	Free on Board
GAMS	General Algebraic Modelling System
GDP	Gross Domestic Product
HEGY	Hylleberg, Engle, Granger and Yoo
IAPI	Instituto Argentino para el Intercambio

IMF	International Monetary Fund
INDEC	Instituto Nacional de Estadística y Censos
INTA	Instituto Nacional de Tecnología Agropecuaria
IO	Innovative outlier
ISI	Import Substitute Industrialisation
MA	Moving average
MATBA	Mercado a Termino de Buenos Aires
OLS	Ordinary Least Squares
PACF	Partial Autocorrelation Function
RER	Real Exchange Rate
SECM	Seasonal Error Correction Model
USD	United States' Dollar
VECM	Vector Error Correction Model

INTRODUCTION

The idea that terms of trade have a tendency to decline over time for developing countries, given their primary production basis, has been well understood and has received substantial attention (Prebisch, 1959). According to this "Structuralism" theory, due to developing countries' specialisation in primary goods, terms of trade can turn against them, as prices of primary products decline relative to those of industrial products. This can then create a centre–periphery relationship between developed and developing countries that hinders the development prospects of the latter when their production matrix does not change in favour of industrial products. This theory has served as a basis and inspiration for academic research and economic policy, primarily in developing countries, from the end of World War II to the beginning of the 1970s.

The increase in the international price of commodities during the 1970s presented an important challenge to this theory (Grilli and Yang, 1988). It was perceived that terms of trade were not effectively evolving against developing countries and, consequently, they became important destinations of real and financial investments. Not only was export income growing as a consequence of the rise in prices, but it was also perceived that further complications might arise (the so-called Dutch Disease) if this extraordinary income were not properly managed (Corden, 1984). This suggested that the problem of development would arise from a factor completely contrary to that which the "Structuralism" view had prescribed. The problem was not that primary product prices were low relative to manufactured products, but rather than they were high enough to discourage investment in other export activities.

However, commodity prices decreased during the 1980s. Developing countries that were heavily indebted because of the expenditure expansion brought about by the commodity boom experienced serious difficulties in the repayment of their obligations given the important fiscal deficits as well as imbalances in the balance of payments Fanelli *et al*.(1990). This "lost decade" was characterised by low growth, high inflation and recurrent balance-of-payments problems, generally tackled by devaluations that resulted in feedback issues, given the low prices of their export products and the impossibility of finding finance for those deficits (Gerchunoff and Llach, 1998).

Structural reforms were encouraged and practiced during the 1990s, with the idea of gaining efficiency by reducing the weight of the Government in the production of goods (private and public), through the limitation or elimination of economic regulations, and the promotion of openness to trade in view of enhancing competition in domestic markets (Williamson, 1990). Whilst perceived failures in these reforms led, in some cases, to changes in these practices (Rodrik, 2006) – particularly in some Latin-American countries – it is still (with some nuances) the paradigm in economic development.

The first decade of the new century brought increases in the price of commodities, similar to those observed in the 1970s. Adjusted by US prices, commodity prices doubled between the periods 2000–2004 and 2008–2012, oil increased even more, and edible commodities followed a similar evolution to the general index (International Monetary Fund, Index of Commodity Prices, base 2005). This has brought a revitalised dynamism to the exports of developing countries and increases in the purchasing power of imports.

Whilst it cannot be said that they are the main or unique explanatory elements, it is clear that issues surrounding the price of commodities have always been present in the discussion to explain the different economic phenomena that have affected developing countries in the last 50–60 years. However, it seems that the problem does not lie exclusively with the level (either high or low) or the tendency, but rather with the volatility of commodity prices. Boom periods have been accompanied, or caused, by movements in prices, and bust periods by incidences of very low commodity prices. Moreover, authors such as Braun and Joy (1968) have highlighted that booms contained the seeds of the immediate decline in economic activity, in a "stop–go" cycle frequently observed in some developing countries. This suggests that there is a strong link between the marked volatility of the economic cycle and the volatility of commodity products – generally the principal export products in developing countries.

The problem of volatility in commodity markets has received substantial attention, particularly since the 1970s; so, too, have the effects of volatility on the exports and income of developing countries. The basic idea is that a high variance in commodity prices can challenge the presumed sustainability of the current account and fiscal deficits, given their effects on the terms of trade, tax collection and the income of households, with consequent effects on consumption and investment decisions.

Research has focused not only on the effects of this volatility but also on the different tools available with which to address it. There were also important efforts made in the development of different theoretical frameworks or approaches on the effects of risk and the different hedging tools available with which agents could address the effects. Several developments have occurred in the analysis on the effects of storage, buffers and forward and futures markets (among other hedging tools), on the volatility of commodity markets and particularly on their prices.

In general, this literature has been developed in isolation from the particular international integration of commodity markets. More precisely, commodity markets have been studied in a context of no national borders, in which case it is hard to motivate a discussion about trade in commodities, since this assumption of no borders dilutes the concept. However, whilst it is true that certain characteristics of commodities allow them to be seen as operating in 'international markets', there are other elements that make it possible to distinguish between markets for commodities. Legal and institutional regulations to disconnect domestic markets from the volatility of international markets, differences in taste and, more importantly, transaction costs, introduce a clear distinction between domestic and export markets (Obstfeld and Rogoff, 2001).

If agents were only interested in maximising their profits, the differences in the level of prices between export and domestic markets should be sufficient to obtain a sensible answer to the supply decision. However, in a context where agricultural output is stochastic and decisions on production must be made before any uncertainty on prices is resolved, the volatility of profits also presents a problem that needs to be addressed.

It is in this framework of volatile prices that the motivation of a producer supplying either one or both markets begins to make sense and to demand proper analytical treatment. Given these differences in markets and the uncertainty present in output and in the market, supply decisions can be governed by more than just the level of prices. Consequently, it is possible that a particular market is supplied not only by its profit-maximising contributions, but also because its price tends to be more stable.

This cross-hedging strategy (Anderson and Danthine, 1981) is performed in a context where storage is available to physically transfer output from one period to another, and when futures markets are available for taking output decisions under certain prices. This suggests that the cross-hedging decision is a complement to, rather than a substitute for, these tools. The general aim of this research is therefore to contribute to this theme – how the decision of supplying exports and domestic markets for agricultural storable commodities is motivated in a context where futures markets are available to hedge against fluctuations in spot prices.

Originally, this research was designed to consider these elements in the context of general equilibrium, as applied through a computable general equilibrium (CGE) model. However, some core assumptions which were needed for the general equilibrium analysis, as well as the treatment itself, proved to be too limiting. Moreover, it was seen that the complexity of the situation described would not be appropriately handled under the general equilibrium framework, where market descriptions tend to be too stylised. Therefore, with some of the elements developed, a partial equilibrium approach was pursued in the development of the theoretical model. This is the aim of the first chapter of this thesis.

The developed model integrates all these highlighted aspects. It represents the decision to supply domestic and export markets with an agricultural storable product where output is subject to technology risk and spot prices in both markets are unknown at the time of making the production decision. The domestic and exported commodities are a "transformation" of a primary commodity whose output is uncertain and for which a futures market exists. Therefore, a trader must decide on the quantities to supply to the export and domestic markets, and the amount of futures of the primary product which he or she should trade to hedge against fluctuations in the price of the input. This suggests that participation in all the markets describes the hedging strategy of the trader. This latter is represented in an analogous way as a processor and the second stages of production are introduced by Anderson and Danthine (1983) and Hirshleifer (1988) in futures markets. Consequently, the model integrates the presence of these second stages of production and of multiple products in the hedging decision.

Additionally, the model described requires theoretical validation in the sense that it should verify or provide explanations for departures of the properties and results previously found in the literature on futures and spot markets. Therefore, an analysis of the equilibrium in the futures and the spot markets is performed. The interest is to analyse the behaviour of the bias in the futures price (if normal backwardation or contango prevails), as well as an analysis of the positions taken in the futures markets by agents (if they operate as buyers or sellers), and particularly traders, under general and restrictive conditions in terms of attitudes to risk, costs of production, etc.

The strategy followed in the development of this model tries to let the model be as free of simplifying assumptions as possible – particularly in terms of the relationships between the values of the expected variances and covariances of all the prices involved – and let the model speak. Therefore, the parameters of the supply equations are defined by their constituent elements, explained by the relationships between the expected variances and covariances of the prices, whose sign and size cannot be determined. Therefore, an empirical estimation was attempted, which introduces the second element of this research.

In Chapter 2, a motivation for the election of Argentina as a case study is justified via the presentation of the historical context in which the agriculture in this country has developed. Although the agricultural sector is neither the main employer nor represents an important share of the product, it is of extreme importance in the external sector. The three commodities analysed, in particular, represent a sizable share of Argentine exports. Moreover, it was also shown that the Argentine supply is important in the determination of the product world equilibrium.

On the other hand, the historical Argentine political and economic contexts have been presented with the objective of introducing and rationalising some of the analytical and methodological aspects of the analysis, as well as contextualising some of the techniques used. In this sense, a description of the production technology is introduced with the objective of characterising the type of seasonality and exploring how it is affected in the case of a very large country. Moreover, Argentina's unstable economic performance is analysed with the objective of introducing some of the assumptions, and economic and empirical treatments used later in the research.

Additionally, the importance of the futures markets in Argentina is analysed. It is shown that these markets have been operating in the country for more than a hundred years and, although there have been periods where these markets lost their importance in commodities trading, the Argentine futures markets are among the most important in the World. Moreover, it is shown that, given the influence of important domestic and international traders, the model developed in Chapter 1 can be easily accommodated to the reality of Argentine grains commercialisation.

Chapter 3 addresses the analysis of seasonality in the context of the domestic and export supply of agricultural commodities. Given the characteristics of the model, a time series approach was suggested, using cointegration techniques applied to the analysis of the markets for soybeans, maize and wheat in Argentina. Monthly data indicated the presence of a strong pattern of seasonality in series that demanded an analysis of stochastic seasonality using the Hylleberg *et al*. (1990) or HEGY approach. One of the key problems was the frequent presence of zero values in the series.

Time series based on agricultural processes tend to present a distinctive and strong pattern of seasonality given by the intrinsic seasonal nature of agricultural activities, particularly in the case of annual crops. Production, export and other agricultural series, particularly when monthly observations are considered, present periods or seasons where zero values are observed that cannot be explained by the lack of reporting or other measurement aspects. On the contrary, during these periods these are the actual values of the series that require adequate treatment. On the other hand, aggregation into a lower frequency in order to eliminate these features may smooth the series, eliminating relevant information on the cycle of the series (Rossana and Seater, 1995), and affecting the inference that can be made out of them (Ghysels, 1990). The monthly export and the domestic supply of some annual crops in Argentina present these features.

The extant literature and empirical applications of the HEGY test had not considered this possibility, since they tend to work with aggregates or indexes of different nature that do not display zero values. Since it was not clear whether the critical values used to make inferences about the presence of seasonal unit roots were or were not affected by this type of data generation process, the calculation – using Monte-Carlo techniques – of new critical values reflecting this peculiarity was performed.

In addition, it is generally observed that the presence of structural breaks might reduce the power of seasonal unit root tests. The locations of the breaks were unknown, so the HEGY test was performed, under the presence of unknown structural breaks, to confirm the results obtained. Unfortunately, the literature in this particular field was even scarcer, and no critical values had been tabulated for the case of monthly data. Therefore, it was necessary to tabulate critical values for this particular case that are now available for future reference and use. In addition, the possibility of addressing the problem of seasonality through a deterministic approach has been considered. This implies that this chapter presents a very comprehensive, deep and innovative application of the seasonality analysis.

Chapter 4 deals with the empirical validation of the developed model. The estimation or empirical validation of a theoretical model for the first time is particularly problematic. It is not possible to use previous empirical applications as guidelines or for reference and, particularly in this case, the theory does not help in identifying the signs of the parameters or coefficients under estimation. Therefore, this chapter is seen as a starting point for an on-going validation exercise.

The technique applied for the estimation of the export and domestic supply equations developed in Chapter 1 is a seasonal cointegration approach based on Engle *et al*. (1993) or EGHL. Very little work has been applied to testing seasonal cointegration techniques using monthly data, and none using this particular approach. In particular, the cointegration relationships and equations for monthly data had to be developed specifically for this case. Therefore, the first part of Chapter 4 is devoted to the presentation of the technique and the development of the relationships necessary for the seasonal cointegration tests.

Given the lack of matching seasonal unit roots between the variables of the model (quantities and prices), a standard cointegration test for the long-term relationship is performed for the Argentine case. This is done pairwise (quantities versus each price), with consideration of the proper definition of the supply equations developed in Chapter 1, taking care of violations to the statistical assumptions of the methodology employed.

The final part of Chapter 4 is devoted to the estimation of the export and domestic supply equations developed in Chapter 1 through an error correction mechanism (ECM) model. The models have been enhanced with the addition of variables in order to appropriately identify the supply functions. Under the assumption of stable expectations of the variances and covariances, this estimation is expected to shed some light on the signs and sizes of the coefficients defined in the equations, and to verify the empirical validity of the model. The results obtained are discussed in depth and the estimation problems are identified, providing potential solutions to the methodological and data problems of the estimation.

The thesis concludes by detailing and discussing the results obtained and their implications. The theoretical, methodological and technical contributions to the understanding of the trade decision of agricultural commodities under uncertainty are presented, as are the testing and estimation techniques applied during the process, indicating the potential applications they may have. The analytical and methodological challenges faced and the ways in which they have been addressed are discussed. Finally, alternative procedures and methodologies are suggested to address some of the problems encountered that constitute current and future research avenues on these topics.

CHAPTER ONE

TRADE DECISIONS AND RISK: FUTURES MARKETS AND PROCESSED PRODUCTS

Summary

International trade in agricultural commodities is concurrent with the existence of futures markets and traders who intervene in those markets. This chapter presents a theoretical model with which to characterise the trader's decision in the domestic and export supply of storable agricultural commodities under the presence of futures markets, which integrates the approaches on the second stages of production and multiple products. Output of the input is subject to technological risk, and agents' decisions must be made – before uncertainty is resolved – on the quantity and price of the input and the price of the supplied products. The separation result between output and hedging decisions cannot be sustained in the problem of the trader. In general, traders constitute a natural counterpart for the rest of the agents who want to sell futures contracts, and only under very restrictive conditions would the trader sell futures contracts. In the short run, the futures price is biased even when traders are neutral to risk. In the long run, the futures price is still biased but the bias disappears if traders are neutral to risk. The decision to supply the export or the domestic markets is explained not only by the price level, but also by the volatility reduction properties that the supply of a particular market may exhibit. In this sense, it is possible that a market is supplied because the variance of its price is lower or because it can help to hedge against fluctuations in the price of the input. Operation in the futures markets can help to hedge against fluctuations in the price of the input and against the effects this may have on the price of the supplied products.

1.1. INTRODUCTION

Volatility in commodity markets has received substantial attention, particularly since the recognition in the 1970s of the importance of commodities in the exports and income of developing countries. The effects on the path of income of volatility of commodity markets in developing countries are well understood. High variance in commodity prices may challenge the intertemporal consistency of the balance of payments given its effects on the terms of trade. It may also affect the Government's fiscal balance through the effects on the value of exports (typically, when export taxes are in place) and the effects on the income of taxpayers. These effects are consequences and/or causes of the microeconomic effects seen in the income of individuals and households directly and indirectly involved in supplying commodity products, particularly in developing countries.

Many agricultural commodity markets are inherently volatile due to the stochastic nature of agricultural production. Weather is the main source of volatility in production technology. Since agricultural markets are internationally integrated, heterogeneous weather effects in the world introduce additional volatility in domestic markets. This suggests that not only do domestic weather shocks matter, but also the effects that weather exerts in other parts of the world.

On the other hand, these effects not only affect the results but also the way economic decisions are made. In a context of high volatility of prices, agricultural producers need to decide in advance how much of a product to produce without knowing with certainty the price at which they will sell their output. Moreover, the quantity finally produced is outside the producer's control so their decisions on how much to produce are taken under uncertainty. This means that at the time of the output decision, uncertainty is a major component in the production process.

Given these characteristics of commodity production, particularly in agriculture, different instruments have been developed to address these issues. Weather insurance has been introduced to deal with the uncertainty in technology, for example. On the other hand, storage has also been present since time immemorial to transfer physically output from one period to another, storing commodities when prices are low to be sold when prices are high. Additionally, futures markets have been developed later in time to address the volatility of the market price at the time of decision for production.

Whilst futures markets have existed for a long time, their implications have been deeply studied during the twentieth century, particularly after the oil crisis in the 1970s. Not only futures markets have been analysed from the operational point of view, but also on the more general aspect about their effects on the volatility of prices. The debate about whether futures markets increase or decrease the volatility in commodity markets has been under the spotlight.

The general theoretical approach of futures markets considers the existence of a large number of producers that participate in future markets, not only in order to hedge themselves against fluctuations in the spot price but also try to obtain a gain from the speculation opportunity that futures markets provide. There is a standard demand schedule that may or may not operate in the futures market. Additionally, a speculator takes the complementary position to the producer in the futures market. Because of this, it is verified that future price is a biased estimator of the expected price; that backwardation tends to prevail under typical attitudes to risk and that, under additive risk for technology, the production and hedging decisions are taken separately.

This framework faced some limitations given its simplicity when a more realistic setting is considered. In reality, there are traders or processors that participate in agricultural commodities. Farmers, in general, are the first link in the chain of commercialisation of agricultural commodities. Traders or processors, at the second stage, use this product as an input. Thus, they also are interested in securing a price for their inputs using futures markets. Nevertheless, processors have been introduced previously in the literature to address this issue. In fact, the introduction of processors has helped to reconsider the functioning of future markets.

Nevertheless, it is true that processors can be, and often are, multiproduct firms. For example, processors of soybeans produce two completely different products with different demands: Oil and soybean meals. Whilst it is true in this case that they tend to use a fixed proportions technology, a more general treatment will reveal that there is scope to reduce the volatility in their profits by taking different positions in the output of their processed products. More importantly, it has been shown that the optimal positions in the market for processed products and the future market must be determined simultaneously. It will be determined by how each price involved (processed products, input and futures prices) moves with respect to the others. Consequently, the decision of supplying each processed product will depend on its own price, the prices of other processed products, and the price of the future prices of the input.

On the other hand, agricultural commodities are highly integrated internationally. Arbitrage and the homogeneity of the product traded should secure that differences between prices across the world are mainly explained by transaction costs and taxes. This means that the effects of changes in demand conditions or in the supply of the same product in other markets are transmitted almost immediately to other markets. This, in turn, affects not only the decision to supply but also the position taken in the future markets. Therefore, there is a marked link between trade, the domestic market and futures markets.

Nevertheless, the existence of different institutional barriers may affect the international integration of markets. Taxes and/or Government intervention (frequently with that precise intention) may disconnect domestic and international markets, implying that the decision in terms of which or to what extent to supply a particular market may not be trivial or simply governed by demand. Therefore, a model that integrates futures markets and the decision of supplying either domestically or internationally can help to visualise more comprehensively these interactions and the way commodity markets work.

This chapter aims to develop a model that can integrate these aspects of commodity markets. In this model, farmers sell their primary product to traders who are those that effectively supply exports and domestic products. In this sense, traders can be seen as processors that transform a single input into both an exported product and a domestically consumed product. In the interest of farmers as suppliers of the input and processors as buyers, a futures market for the primary product is open to have certainty on the price of the output/input. The treatment proposed combines the existence of two stages in production, previously found in literature, and the existence of multiple products in the decision of hedging, in order to represent more realistically the reality of commodity markets. This means that our focus will be on the decisions made by traders in terms of their participation in futures markets, export markets and the domestic supply.

Nevertheless, in order to determine whether the prescriptions of this model are valid in terms of domestic and export supply decision, it is necessary to analyse the implications on the futures markets. The model will lack theoretical validity if the futures market in this context cannot be appropriately identified and characterised. It is necessary to analyse, which results are in line with the fundamentals of futures markets and provide economically meaningful explanations for their departures. This implies a deep study, based on comparative statics, on how the model behaves under different circumstances, how the futures market is affected and why. In the process, we will find very interesting theoretical results that have not been identified before but are in line with other compatible theoretical findings.

The results in terms of export and domestic supply decisions reveal that both, as it is expected, are affected by their own price. In addition, a cross price effect is present in both decisions, suggesting a cross hedging strategy given that the definition of the coefficients in the supply equations are affected by variances and covariance of these prices. In addition, the supply is affected by the future price of the primary product, but this effect may be different between exports and domestic supply. This suggests an important role in the determination of supply between the expected covariances of the prices of the exports and the domestic product and the price of the primary product, respectively.

The task is far from being easy. The existence of second stages of production and multiproduct firms introduces complex mathematical features that make the model very difficult to analyse. The sign of several parameters, for example, cannot be determined without making assumptions about the value of others. Whilst these complications could have been addressed by making some assumptions about values or relationships between parameters and variables, it was decided to leave the model as free from simplification as possible, in order to gain understanding but paying the price in terms of the difficulty on the analysis.

Whilst the parametric approach followed presents these mathematical tractability problems, it allows the precise identification of the effect of the expected variances and covariances in each of the coefficients of the equations. Moreover, it is convenient for numerical exercises such as simulations or econometric modelling. These, particularly the econometric validation, will be a future task of this research.

This chapter presents the following structure. In the first section, we will present how the literature has addressed futures markets from the theoretical point of view as well as identify the most common elements that characterise futures markets. At the same time, we will discuss some important points on the tradability and the international integration of commodity markets. In the second section, we introduce all the agents that give life to this model presenting the equations that represent them and a more detailed discussion on the trader is attempted. Initially, we analyse what motivates traders to participate in futures markets and analyse the implications of their participation. After that, we discuss the supply decision of traders in more depth by looking into the supply equations that represent their behaviour. In the fifth section, we complete the model and analyse the equilibrium in the futures market. In the sixth section, we characterise the long run equilibrium for exports and domestic markets. Finally, we summarise the main findings.

1.2. NOTES ON THE ANALYSIS OF FUTURES MARKETS

The literature about volatility in commodity prices when storage and futures markets are present is vast and rich. Particularly, during the 1970s and early 1980s, there was important research in this field, fuelled by the high volatility in prices experienced at that time. Carter (1999) highlights that the Journal of Futures Markets was first published in 1981. On the other hand, the presence of this volatility also demanded important research on the possibilities of price stabilisation and its benefits. After that period of deep theoretical research, an empirical validation literature continued, trying to contrast the theoretical prescriptions with reality (Carter, 1999).

Part of the analysis has been based on identifying the effects and convenience of the development of futures markets. The issue about stabilising or destabilising properties of future markets on prices has been the subject of a long debate. In the nineteenth century, the lack of theoretical and empirical analysis, as well as some preconceptions, led to the prohibition of the operation on futures under the assumption that they, as well as speculators, were responsible for the instability in prices as Jacks (2007) highlights. Nevertheless, the theoretical and empirical analysis on futures markets (and commodity markets in general) received important attention, particularly in the second part of the twentieth century.

Although improved and better analytical instruments have been developed, the debate about the effects of futures markets on prices is still open. Chari, Jogannathan and Jones (1990) show, under a framework without uncertainty in production and for non-storable commodities, futures markets may increase the volatility of the spot price. Kawai (1983) and Turnovsky (1983) report similar findings in the context of production uncertainty and storable commodities. This suggests that futures markets, rather than decreasing the volatility of the spot price, may actually increase it. Providing an alternative view, Britto (1984) and Newbery (1987) suggest that futures markets provide an insurance that make producers take risker decisions, generating higher volatility on output and, consequently, on the spot price.

However, it is also stressed that futures markets provide better information to make production decisions. By providing known and certain prices, futures markets allow producers to make better production decisions as well as facilitating storage as suggested by Peck (1976). Particularly with respect to the improved information that futures markets provide has been the result that Turnovsky (1979) concluded when comparing the effects of futures, storage and other stabilisation mechanisms. Cox (1976), surveying the empirical evidence, provided similar conclusions. On the other hand, Garbade and Silber (1983) find that news is reflected faster in futures prices than in spot prices, adding more evidence to the idea that futures provide improved information. Moreover, Turnovsky and Campbell (1985), in an alternative proof of the stabilisation properties of futures markets, suggest that under rational expectations, futures markets always guarantee the existence of a unique long-run equilibrium and without them, and low elasticities of demand, the solution cannot be guaranteed. Moreover, Netz (1995) shows empirically that futures markets, via an increased effect on storage, generate more stable spot prices.

Whilst there seems to be some theoretical and empirical evidence about the properties of stabilisation of futures markets, the most interesting finding is that futures markets, by providing a known and certain price, appear to improve the way producers take their production decisions. Welfare implications about the existence of future markets suggest that welfare is at least not reduced by the presence of futures markets as suggested by Turnovsky and Campbell (1985), given the improved information available to agents. In fact, Kawai (1983) highlights the fact that futures prices tend to be more stable than spot prices. This suggests that the price used to make production decisions is more stable with futures markets than without them.

An important part of the discussion about futures markets has been based on the explanation for their existence. Earlier references to the role and characteristics of futures markets can be traced to Keynes (1930) who originally identified that futures prices are not accurate estimates of the spot prices. In fact, Keynes gives the name “normal backwardation” to the generally found result that the future price underestimates the expected spot price. However, futures markets existed long before these works identified their importance. Earlier forms of futures trading can be found in the 1850s. However, it is recognised that the use and importance of futures markets dramatically increased during the 1970s as Carlton (1984) suggests.

The original idea that futures markets provide a return for storage services was originally developed by Working (1949). This means that the futures market is, in the end, a market for storage. It is also seen that storage can be most efficiently made if futures markets are available and the joint action of both may have an important stabilisation effect as suggested by Peck (1976). However, the idea of futures markets naturally providing returns for storage lacks some impetus when the possibility of the existence of futures markets for non-storable commodities is considered. This means that whilst storage and futures markets may be complements rather than substitutes for stabilisation, the explanation of their existence cannot be based solely on their relationship with storage.

The alternative explanation for the existence of futures markets is that they provide a hedging tool against fluctuations in prices. This has been the most widely accepted interpretation of futures market. In addition, different authors such as Danthine (1978), Holthausen (1979), Turnovsky (1983) among many others have shared it. Under risk-averse agents, futures markets provide a certain price that not only affects the output decision but also reduces the volatility of profits. However, it should be stressed that futures markets are not the only way producers can reduce their risk exposure. Share contracts, where the output is shared between the farmer and another agent at a specified price, are also available and even can be seen as complementary to futures trading, as suggested by Hirshleifer and Subrahmanyam (1993). Moreover, the development of buffer stocks has also been discussed and analysed among others by Newbery and Stiglitz

(1979) and Gemmill (1985). Nevertheless, the market-based tools for price stabilisation or hedging, such as futures trading, have become widely available.

However, futures markets also provide agents with the possibility of speculation. If the futures price for delivery at a particular time were higher than the expected price at that time, a speculator would find profitable to sell a futures contract (even though he does not have the product physically); and when the delivery time arrives, buy in the spot market and make the effective delivery of the product. The speculator would gain the difference between the price paid in the spot market and the price received by the futures contract sold.

Consequently, the decision of trading in futures is fuelled not only by hedging against fluctuations in prices but also by the possibility of exploiting a difference between the futures prices and the expected spot price as Gray and Rutledge (1971) highlighted. It is important to note that the speculative behaviour is not only a matter of "outsiders" that want to make a profit out of the speculation. In the standard analytical framework, all agents include a speculative component in their decisions. This means that speculative and hedging motives are concurrent in the decision to participate in futures markets (Kamara, 1982).

These analytical frameworks tend to present a simple and clear setting of the futures markets. Farmers seeking to maximise profits and reduce their volatility, and facing uncertainty about the price they will receive on their also uncertain output, have the opportunity to operate in the futures markets by selling commitments of delivery of the product at a given period for a certain price. The farmer has to decide how much he wants to produce and how much he wants to sell in the spot (for an uncertain price) and on the futures markets. Therefore, two decisions are involved: an output and a hedging decision. These decisions, depending on how technology risks affect production, can be either separated or taken simultaneously.

This separation between the output and hedging decisions was originally discovered by Danthine (1978) in a context without output risk. When this separation result holds, the output decision is based exclusively on the future price, which means that the output decision is made under certainty. The futures price is available to the producer regardless of its participation in futures markets. It is a price that the farmer, for example, can find in the newspaper and that can be used to decide how much to produce. The hedging decision (more appropriately the decision on how much to assign to the futures market) presents a component similar to the output component (the proper hedging component), and a speculative component that is affected by the futures price and the expectation of the spot price the farmer has formed. Therefore, depending on the relative sizes of the hedging and speculative components, the farmer may under or over hedge his output.

When non-storable products are considered, some of the results generally found on storable goods do not hold. One of these results is related to the bias in the future price. In storable goods, it is observed that the future price is a biased estimator of the expected cash price. The direction of the bias will depend on several factors, as we will see. However, this result does not hold when non-storable goods are considered, generating the result that the future price is an unbiased estimator of the expected cash price as Kawai (1983b) has shown. Additionally, Leuthold (1974) indicates that futures prices may not be accurate predictors of the futures cash prices, and decisions based on them, at least for non-storable commodities, may generate instability. Nevertheless, the storable and non-storable cases can be seen as extreme cases since nothing is completely storable or perishable. Furthermore, the reason for the existence of futures markets has been based on the hedging and speculation possibilities they provide rather than their relationship with storage.

In these models, there is a speculator who does not participate in production but who forms his own expectations on the spot price and, depending on the level of the future price, he buys or sells futures. It is generally considered that, in order to adopt a complementary position in the future market (i.e. being long if the farmer has being short) the speculator will require a premium that ultimately explains the bias in the future price as an estimator of the expected price. These frameworks are generally closed with a consumer who may or may not participate in the future markets, using standard linear demand function that can also have a stochastic component.

The economic setting presented by this framework has been extremely helpful to conceptualise and characterise the presence of future markets. Nevertheless, this setting may be seen to be unrealistic when we look into how commodity markets are organised. Other relevant agents determine the behaviour of the spot and the futures markets. These agents generally appear when a more realistic representation of the demand is considered.

Processors, for example, as buyers of the primary product to be used in production in a second stage, are also interested in securing a certain price for their inputs that may lead them also to participate in future markets.

The presence of production stages in commodity markets has been analysed by Stein (1979), Anderson and Danthine (1983) and Hirshleifer (1988) among others. Hirshleifer in particular includes processors of a single input that produce a single output. In his paper, the sources of bias in the futures price and the positions taken by the different agents in the futures markets are analysed and confronted with previous theoretical findings in models without second stages of production. Since processors are interested in reducing the volatility of the price of the input they require, they have a tendency to buy futures (to secure a price for the input) or go long in the jargon, providing a “natural” complementary position of the producer, who has a tendency to sell futures or go short. However, as we will see, this does not guarantee that the future price will be free from bias. On the other hand, Anderson and Danthine (1983) working with one processed product, analysed the separation result previously found by Danthine (1978) and Holthausen (1979) and found that the presence of a second stage is irrelevant in the determination of the bias in the futures price. However, in their approaches, the marginal cost of production tends to zero, almost eliminating the existence of the second stage.

On the other hand, traders also participate in commodity markets. Traders, having superior logistics and commercialisation skills, found it profitable to be a link between the production of the farmer and the demand. A trader can be seen as a processor since a cultural or idiosyncratic transformation is performed on the commodity. Traders, as the processor, are also affected by the volatility on the price of the commodity or input and, therefore, their interest in participating in futures markets to secure a certain price for the input is manifested. Therefore, the analogy between traders and processors facilitates the analysis of the behaviour of traders. In the following, we will refer to traders and processors interchangeably.

The basic framework can be also enriched if we consider the effects other commodities markets and their futures markets can have. Participants in futures markets, as we have seen, are also interested in making a speculative profit. This implies that they are not limited to speculate only in this market. The farmer’s profit, for example, might not only be affected by the income generated by selling his output, but also on the speculation made in the futures market of his output and the futures markets of other products.

Nevertheless, pure speculation may not be the only reason for the operation in other future markets. A genuine hedging motive may be behind this decision. If we consider goods delivered at different times and places as different products (such as wheat delivered in August can be considered as a different product to wheat delivered in November), it is possible to form a hedging portfolio considering positions taken in these different future markets. This means that other futures markets can provide hedging for the fluctuation in the price of the product by taking positions in the different futures markets that are open for that product, creating an overlapping structure of futures contracts that close at different points in time, as highlighted by Anderson and Danthine (1981). However, on the other hand, if we amplify this definition of distinctive products, it is possible not to limit the possibilities of hedging to a subset of futures markets. This suggests that it might be convenient to take positions in different futures markets for completely different products. This is particularly the case when futures markets are not available for the product in production. This, of course, will depend on how the futures market for that product behaves in respect to the market of the product under hedging.

At the same time, it is very common that producers' profits may be affected by the production of more than one product. Multiproduct firms are generally the rule rather than the exception in this context. Positions taken in the spot markets for each of the products will affect profits, not only in their levels, but also in their volatility. This means that profits are affected not only by the level of the products' prices, but also by the volatility of the price of all products in production, not only in terms of their respective variance but also in respect to the covariances between these prices. In this sense, a product may be produced, not for its contribution to the level of profits, but given its low volatility or lower volatility with respect to the rest of production mix. Thus, it can reduce the volatility in profits.

This possibility has received some attention in the literature. The idea that a producer of a commodity can use several futures and multiple cash positions has been analysed by Anderson and Danthine (1980; 1981). In their framework, farmers can take positions in different futures markets (distinguished by time and location) as well as different cash positions. They found that optimal cash and futures positions must be determined simultaneously, and when dealing with a good for which no futures exist, a cross hedging strategy may be convenient (taking positions in cash and futures for other products). The optimal cash and futures positions will depend eventually on the matrix of covariances between all prices in the model.

1.2.1. International Commodity Markets

Commodities, on the other hand, are generally considered as the archetype of an internationally tradable good. Relatively low transport costs make it possible for commodities to be traded worldwide. Whilst commodities are in general tradable, the opposite is not true (generally, tradable goods are not commodities). However, the tradability of commodities is particularly deep. The level of international integration in commodities markets is difficult to see in other tradable goods. The general homogeneity of commodities renders output from different parts of the world are perceived as perfect substitutes by the demand, and transport costs are the only elements that may explain differences in their prices. Given transport costs are relatively low, arbitrage ensures that any change in demand or supply is almost immediately transmitted to all markets. Because of this, it is very common that the price of a product in a particular place can be taken, without further considerations, as the reference price for the complete world market. This also holds in futures markets, where the evolution of the futures prices in the different markets for a given commodity is not only affected by events in those markets but also in the rest of the world.

Nevertheless, if commodity markets are effectively perfectly integrated, the distinction between export and domestic markets may lack sense. If arbitrage ensures that prices across locations follow similar paths, there is no space for differences before or after customs. In fact, if international commodity markets are integrated, the “Law of One Price” should hold at least in the long run. However, Ardeni (1989), applying cointegration techniques, suggests that international arbitrage cannot guarantee, even in the long run, equality between prices. Different formal and informal barriers may prevent this to happen.

The existence of taxes on trade as well as Government intervention can have important effects on the link between domestic and international markets. Export taxes, for example, are sometimes justified to disconnect the evolution of the export price from the domestic price in order to favour local consumers when international prices are too high. 'Grains boards', on the other hand, have been introduced in the past (and still exist today[1]), intervening between supply or demand with the idea of achieving stable levels of prices and production. Consequently, local legislation and Government policy can prevent integration between markets by introducing a wedge that can explain different paths and volatility of these prices.

Moreover, differences in legislations and efforts across countries with respect to contract enforcement as well as the costs of trading (particularly on the enforcement of payments), may introduce another wedge between futures markets and commodity markets across the world. In this sense, local producers or traders may be more inclined to trade in the domestic market given the perceived lower costs of securing payment or litigation. Moreover, additional transaction costs involved in the export activity may be seen as too high for individual producers, thus making them more biased to trade in the domestic market. This "home bias" in trade, is a phenomenon that was originally identified by McCallum (1995). It was further analysed in Obstfeld and Rogoff (2001) in more general trade contexts that could be extended to the case analysed here.

Therefore, the presence of these institutional and political aspects prevent conceiving commodity markets as unique worldwide. When these elements are considered, it is possible to identify different commodity markets with respective demands and different prices. Moreover, the existence of different volatility of the prices in each of these markets adds an extra factor in the consideration of having different supply decisions for each of them. It is here a clear distinction between export and domestic markets emerges.

However, very little has been analysed in this sense and the research on how futures markets affect the export of supply decisions has received very little attention. The relationship between international trade and futures markets has been analysed only in the context of exchange rate risk by Kawai and Zilcha (1986) and Viaene and Zilcha (1998). In these models, an exporting firm faces risk on prices, output and on the exchange rate. Therefore, a futures market for the foreign exchange is introduced and their implications are analysed. However, this treatment does not consider the possibility

[1] The Canadian Wheat Board has been functioning since 1935.

of a firm that must choose between supplying two different markets (a domestic and export market), or the effects of this decision on the volatility of profits. The firms involved are exporting without assessing the reasons or the motivation for either that or the effects that this decision may have. Therefore, the trade decision is exogenous and independent of the existence of future markets. On the other hand, Hueth and Schmitz (1972) analyse the implications of international trade in intermediate and final goods on prices following the framework developed earlier by Massell (1969). However, this approach considers again the situation of goods that are only traded internationally and not domestically. Moreover, their analysis is clearly deterministic in the sense that no role is allowed for expectations on prices and, consequently, no role for futures markets.

An analytical framework that considers multiple processed products, to represent the export and the domestic supply decision; and second stages in production to represent the role of traders in commodity markets helps to represent more realistically and improve the quality of the analysis of commodity markets. This is the task of the following sections.

1.3. EXPORT AND DOMESTIC SUPPLY WITH FUTURES MARKETS

1.3.1. General Setting

In this section, we will introduce the main elements of the model under development. In the first part, we will just present the farmer, the storage and the speculator problems. This follows a standard treatment already presented in the literature. However, we will not ignore some important elements that will help to understand and motivate the rest of this presentation. Then, we will present the problem of the trader or processor and will devote space to analyse some of the implications of the existence of this agent in terms of its hedging and, more importantly, its decision of export and domestic supply. We will then analyse the equilibrium in both, the futures and the spot market. The analysis of the equilibrium in the futures markets is necessary to give the model some theoretical validation by verifying some results already found in the literature.

The setting of this model follows closely the parametric approach adopted by Kawai (1983), Turnovsky (1983), and Turnovsky and Campbell (1985). This approach is convenient for its analytical properties and the possibility of being adopted in mathematical software such as Mathematica or GAMS in order to perform some comparative static exercises and simulations. Moreover, this approach will prove to be convenient for the econometric estimation in the subsequent part of this research. The alternative and more general approaches, whilst more economical in terms of notation, are harder to implement in this way.

There are four agents in this model: a farmer or producer, a storage company, a processor and a speculator. The farmer produces a primary good under an additive stochastic production function. At time *t-1*, the farmer must decide how much he will supply of the primary product at time *t*; however, the final production and supply will be affected by exogenous factors outside his control. In a real context, the farmer decides how much area he will devote to the production of this good and we could consider a "level of effort" (given by the amount of fertilisers and other practices) also involved in this decision. However, external factors outside its control (mainly weather) determine the final quantity produced.

However, the farmer has knowledge about the statistical distribution of these external shocks. At time *t-1*, on the other hand, the farmer faces uncertainty about the price he will receive because of his efforts. The output decision must be made, in this context, under the uncertainty given by the unknown price he will receive for that output. However, he has knowledge, available also to the rest of the farmers, about the mean and the variance of the primary product price. On the other hand, he has the possibility at time *t-1* of trading in futures for delivery of the primary product at time *t*, as such he can make a commitment for delivery at time *t* under a known price at time *t-1.*

As we will see, the assumption of an additive stochastic production function is relevant. This implies that production shocks affect the level of the output implying that weather, for example, adds to or subtracts from the planned output. A more general stochastic production function, such as a multiplicative one, will affect the marginal productivity. The implication of using a multiplicative stochastic production function is that the separation result found by Danthine (1978) for example, will not hold. This means that both output and hedging decisions must be taken simultaneously. Concerning how relevant is the description of this type of risk to the current analysis is a matter for discussion. Weather seems not to have an effect on marginal productivity, which implies that an additive treatment may be accurate. However, we cannot discard the possibility that other types of shocks that may affect output may have a more general form. Nevertheless, the decision on a particular type of shock is generally based on analytical simplicity.

The storage company is an agent that carries a stock of the primary product from one period to another. This company must decide how much it will store at time *t-1* to carry into period *t*. He does not face any technological uncertainty but he does face uncertainty about the price that will prevail at time *t*. The storage company also has knowledge about the mean and the variance of the price and this knowledge is shared by the farmers and the rest of the agents. The idea behind this assumption is that there is no advantage in the availability or processing of the information by any agent in the model. The storage company, on the other hand, also has access to the operation in the futures market for the primary product will be stored. Therefore, at time *t-1*, the company must decide how much it will store at time *t-1* to be sold at time *t*, and how much it will trade in futures at that time to be delivered in the following period. The activity of storage can be performed (in fact, it is) by farmers. We could have simplified our model by assigning this activity to the farmer and make the appropriate adjustments to his problem. However, we prefer to keep this agent as an additional agent in order to gain insight and simplicity at the time of the exposition.

The speculator, on the other hand, is an agent that does not produce or store either the primary product or any processed product. However, the possibility of trading in futures is open to him at time *t-1* to deliver the product at time *t*. Consequently, if he sells a futures contract (or goes short) at time *t-1*, at the time of delivery, he will need to buy the physical primary product at the spot price that will prevail at that time in order to make the effective delivery of the goods. The operation of futures, in reality, does not involve the delivery of the product. According to Stein (1979), only 2% of all futures contracts are actually settled by delivery. In fact, operators tend to re-buy the futures contract at the time of delivery. The reason is that there are costs associated with the delivery of the good (transportation, inspection costs, etc.). The speculator is not necessary in this model. As we will see, all the agents in the model will have elements in their behaviour that will be similar to those of the speculator. Nevertheless, it is important clearly to identify the speculation component in their decisions and the introduction of a pure speculator will help in this task. However, the model could be simplified by excluding this agent.

The processor or trader transforms the primary product into two different processed goods: an exported good and a domestic good. In essence, both goods are similar, but we assume that the buyers of these products are different and they cannot trade any of these products between them. The processor does not face any uncertainty about the technology of production but at time *t-1*, the processor faces uncertainty about the prices for both processed products (supply price uncertainty) at time *t*, and also about the price he will pay for the input (primary product) he will use in the production of both processed products. Therefore, he is subject to uncertainty on all prices involved in his decisions. Nevertheless, he has information about the means and variances of the prices of the processed products he produces and the primary product he uses. There is no futures market for any of the processed products but the processor can trade in the futures market for the primary product, for example, he can buy a futures contract at time *t-1*, for delivery at time *t* paying a known price at time *t-1*. We will return to this when we formally present this agent.

In the following, we will formally present the problem for each agent. The farmer, storage company and speculator problems will be presented briefly since they have been already introduced in the literature and, whilst illustrative, the focus of this research is on the processor or trader rather than on the rest of the agents. The farmer's profit function at time *t*, π_t^f, can be represented by

$$\pi_t^f = P_t\left(q_t - x_{t-1}^f\right) + P_{t-1}^x x_{t-1}^f - \frac{1}{2}c\bar{q}_t^2, \qquad c > 0 \tag{1.1}$$

where P_t is the spot price of the primary product at time *t*; q_t is the quantity produced of the primary product at time *t;* x_{t-1}^f is the quantity of futures contracts the farmer trades at time *t-1* to be delivered at time *t;* P_{t-1}^x is the price of the future contract at time *t-1* for a product to be delivered at time *t;* and $\bar{q}_t$ is the quantity planned by the farmer at time *t-1* to be produced at time *t.* Therefore, at time *t-1* our farmer must decide how much he will supply at time *t* under uncertainty since, at that time, the spot price that will prevail for time *t* is unknown for the farmer as well as the final output. The effective and the planned output are related by the following expression:

$$q_t = \bar{q}_t + \varphi_t$$

where φ_t is a disturbance term with mean equal to zero and known variance. A multiplicative type of risk will affect output by $q_t = \bar{q}_t \mu_t$, where μ_t will have a mean of 1 and constant variance. For the farmer even though the spot price and the quantity finally produced are unknown, he has certainty about the cost of production and there is an open market for futures with a known price that he can use in his decisions. The quadratic cost function has been extensively used in this kind of approach and it has the advantage that generates linear first order conditions. The producer's utility follows the usual mean-variance approach, where the level of profits increases utility but the volatility decreases it.

$$\Omega_t^f = \bar{\pi}^f(t,t-1) - \frac{1}{2} a_f \sigma_{\pi^f}^2(t,t-1) \tag{1.2}$$

where a_f reflects the farmer's risk aversion, $\bar{\pi}^f(t,t-1) = E_{t-1}\pi_t^f$ is the expectation of the profit for time t conditional to the information available at t-1; and $\sigma_{\pi^f}^2(t,t-1) = E_{t-1}\left(\pi_t^f - E_{t-1}\pi_t^f\right)^2$ is the expected variance of the profit for time t conditional to the information available at time *t-1.* The expected profit and variance can be expressed as

$$E_{t-1}\pi_t^f = \bar{\pi}^f(t,t-1) = \bar{P}_t \bar{q}_t - \bar{P}_t x_{t-1}^f + P_{t-1}^x x_{t-1}^f - \frac{1}{2} c \bar{q}_t^2 \tag{1.3}$$

$$\sigma_{\pi^f}^2(t,t-1) = E_{t-1}\left(\pi_t^f - E_{t-1}\pi_t^f\right) = \left(\bar{q}_t - x_{t-1}^f\right)^2 \sigma_P^2(t,t-1) \tag{1.4}$$

where $\bar{P}_t$ is the expectation of the primary product price for time t conditional on the information available at time *t-1* and $\sigma_P^2(t,t-1)$ is the expected variance of the primary product spot price for time t conditional on the available information at *t-1.* Since there is a large enough number of farmers, the covariance between the output price and the output shock tends to zero since the individual producer only has information about his own output shock and not about the effect on other farmers, as highlighted by Turnovsky (1983), eliminating these terms from equation (1.4). Substituting equations (1.3) and (1.4) into equation (1.2), and maximising respect to the planned output and the quantity of futures at *t-1,* both variables for which the producer has control, we obtain:

$$\bar{q}_t = \frac{P_{t-1}^x}{c} \tag{1.5}$$

$$x_{t-1}^f = \frac{P_{t-1}^x}{c} + \frac{P_{t-1}^x - \bar{P}_t}{a_f \sigma_P^2(t,t-1)} \tag{1.6}$$

Equations (1.5) and (1.6) reflect the well-known separation result found originally by Danthine (1978) and Holthausen (1979). The hedging decision does not interfere in the output decisions. Output decisions are made under certainty since neither the attitude to risk nor the risk itself affects the output decision. Since the future price for delivery at time *t* is known at *t-1*, the output decision is made under a certain price and known marginal costs. It is important to remark that the future price is used in the output decision regardless of participation in the futures market. However, if futures markets were not available, the decision would need to be based on expectations on the price received. In contrast, in our case, the farmer uses a known price he can find by listening to the radio in the morning or by reading a newspaper.

The hedging decision, expressed in equation (1.6), determines the quantity to trade in the futures markets. It is here where the farmer will decide whether to operate in the futures market. This decision has two components: a pure hedge component, the first term in equation (1.6), analogous to the output decision, and a speculative component, given by the second part. The similarity between the hedge component and the output decision reflects the hedging nature or motive on the operation in the futures market. The second component, on the other hand, reflects the extra gains that the farmer could make by operating in the futures market.

This implies that the total amount to trade in futures is finally affected by the difference between the current future price for delivery at *t* and the expected spot price that will prevail at *t*. If the farmer considers that the futures price for delivery at *t* is greater than the expected spot price, he will sell futures by an amount larger than necessary to hedge its production. On the other hand, if he expects that the expected spot price will be higher than the futures price for delivery at that time, he will buy futures (or will be long) leading him to a hedging position smaller than necessary to hedge its output. Eventually, if the speculative component is negative and large enough and outweighs the hedge component, the farmer will actually buy futures rather than sell them. Therefore, hedging is not the only motive for participation in the futures market.

There has been an important discussion in the literature about the sources of differences between futures and the expected price as it can be seen in Hirshleifer (1989). Keynes (1930) considers that, in general, it should be expected that the futures price will be a downward estimator of the expected price, or that "normal backwardation" should prevail, since a speculator will require a premium to hedge producers. This suggests that the main motivation in the operation in the futures markets is just the possibility of insurance or hedging.

However, Anderson and Danthine (1983) suggest that another source of backwardation or contango (the expected price is an upward estimator of the spot price) is that producers and speculators could be subjected to different types of risk and, furthermore, their positions could not be compatible. Danthine (1978) also highlights that speculators could have better information than farmers and this could explain differences between these two prices. On the other hand, Kawai (1983b) finds that the bias is zero in a model with non-storable goods when the demand is stochastic but the supply is certain. Moreover, the possibility that over the production cycle, periods of contango and backwardation could appear was analysed by Hirshleifer (1989). The final effect of backwardation or contango is, as we have seen, that the producer under or over hedges its output. We will analyse these elements later in this chapter when we have introduced all the agents in the model.

We also consider the possibility of storage by assuming a sufficiently large number of storage companies that at *t-1* buy an amount i_{t-1} of the primary product with the intention of selling it at *t*. Basically, the main activity of the storage company is to transform a product available at *t-1* into a product available a *t.* Therefore, the storage company's profit function is given by

$$\pi_t^c = P_t i_{t-1} - \rho P_{t-1} i_{t-1} - \frac{1}{2} h\left(i_{t-1} - i^*\right)^2 + x_{t-1}^c \left(P_{t-1}^x - P_t\right)$$

where ρ reflects different type of costs of the operation with stocks, the most important being the interest rate or the opportunity cost of having idle capital for a certain period, and x_{t-1}^c is the quantity traded in futures by the storage company. On the other hand, there is a desired level of stocks given by i^* and there is a quadratic cost function that penalises any deviation from that desired level of stock where h>0. The rationale behind this specification is that there is a desired level of stock, driven by the convenience of having the product ready to be delivered when facing an unexpected demand. Furthermore, the holding cost is the difference between the actual cost of holding products and the benefit of carrying a large stock that reduces the probability of being out of stock. If the stock is very large, operational costs of storing will make the operation of the stock less profitable. Eventually, the stock can be carried without operational costs if the stock is equal to the optimal or desired level of stocks. This function is identical to the one used by Kawai (1983) and prevents, to some extent, the non-linearity in the storage rule as well as reducing the possibility that stocks could go below zero, as suggested by Newbery and Stiglitz (1981), and Wright and Williams (1982). Other types of specifications will generate a discontinuity in the storage rule since it is impossible for the market, in this context, to borrow from the future or carry negative stocks (Deaton and Laroque, 1992). The discontinuity in the storage rule can be solved using an appropriate storage rules under optimal control techniques. Its treatment is outside the scope of this chapter. As with the farmer, the storage company's utility follows the mean-variance approach and maximises the following utility function.

$$\Omega_t^c = \overline{\pi}^c(t,t-1) - \frac{1}{2} a_c \sigma_{\pi^c}^2(t,t-1) \tag{1.7}$$

where a_c is the storage company's coefficient of risk aversion, $\overline{\pi}^c(t,t-1) = E_{t-1}\pi_t^c$ is the storage company's expected profit conditional to the information available at *t-1* and $\sigma_{\pi^c}^2(t,t-1) = E_{t-1}(\pi_t^c - E_{t-1}\pi_t^c)^2$ is the variance of the storage company's profit conditional on information available at *t-1*. The last two expressions can be described by

$$\overline{\pi}^c(t,t-1) = \overline{P}_t i_{t-1} - \rho P_{t-1} i_{t-1} - \frac{1}{2} h(i_{t-1} - i^*)^2 + x_{t-1}^c(P_{t-1}^x - \overline{P}_t) \tag{1.8}$$

$$\sigma_{\pi^c}^2(t,t-1) = (i_{t-1} - x_{t-1}^c)^2 \sigma_P^2(t,t-1) \tag{1.9}$$

At time *t-1*, the storage company must decide how much it will store to sell at time *t* and how much it will trade in the futures market. Substituting expressions (1.8) and (1.9) into equation (1.7) and maximising with respect to the quantity to store and the quantity to trade in futures of the primary good yields:

$$i_{t-1} = i^* + \frac{P_{t-1}^x - \rho P_{t-1}}{h} \tag{1.10}$$

$$x_{t-1}^c = i^* + \frac{P_{t-1}^x - \rho P_{t-1}}{h} + \frac{P_{t-1}^x - \overline{P}_t}{a_c \sigma_P^2 (t, t-1)} \tag{1.11}$$

As we have seen in the case of the farmer, the separation result prevails. The storage decision is independent of the aversion to risk and expectations with respect to prices and variances. Again, a pure hedging and a speculative component motivate the decision on the operation in the futures market. The hedging component is given by the two first terms in equation (1.11) that again are similar to the quantity produced or in this case, the quantity stored. The third term in expression (1.11) is the speculative component with a similar interpretation to that for the farmer. Note also that if *h*, which can be considered a physical deterioration parameter of the store, is very high the storage company will only keep a constant storage, but it will continue to operate in futures and in this case, the storage company will behave as a pure speculator. The same applies for the farmer is *c* tends to infinity. The farmer will stop producing and will speculate in the futures market.

A speculator is an agent that neither produces nor stores any of the goods considered in the model but participates actively in the futures market. The introduction of a speculator is not essential in this model. As we have seen, every agent, in his hedging decision, also includes speculative component. However, it helps to understand the decisions of the rest of the agents. The speculator profit function can be characterised by

$$\pi_t^s = x_{t-1}^s \left(P_{t-1}^x - P_t \right)$$

Using the same procedure followed for the previous agents, we can obtain the optimal position in the future markets for the speculator.

$$x_{t-1}^s = \frac{P_{t-1}^x - \overline{P}_t}{a_s \sigma_P^2 (t, t-1)} \tag{1.12}$$

Where a_s is the speculator coefficient of risk aversion. In the case of the speculator, the futures contract is not used to hedge a cash position, so the optimal position will be long or short according to the expected sign of the numerator. It is interesting to see that in the absence of backwardation or contango, the speculator will not participate in the futures market. Thus, a bias is necessary in the futures market in order that the speculator participates and takes a complementary position to the rest of the agents.

1.3.2. The problem of the trader

The introduction of the trader/processor or a second stage in production constitutes an important contribution to this framework. Thanks to this treatment, we will model the behaviour of the export and the domestic supply. As we include the possibility of trading in futures, we need also to characterise clearly its implications in spot and futures markets. Therefore, we will devote some time to analyse the behaviour of the trader in depth.

The trader buys the primary product and transforms it, through some cultural processes, into an exported and/or a domestic good. As we see, this framework can also be used to accommodate the situation of a processor that produces two completely different products that uses a common input. For example, a vegetable oil company using soybeans produces soybean oil and soybean meals, with different uses and consequently, different demand functions. A dairy company using milk produces an even wider range of products. This highlights the flexibility of this framework and the wide range of applications it could have.

However, if we follow a restricted or traditional definition of production, this treatment encounters problems. In the common view, trading is generally not seen as a production activity, which means that we could not use the problem of the processor to represent a trader. However, if we keep a wider, but also stricter definition, where production can be seen as the domain and control of different biological, physical, chemical and cultural processes with the objective of obtaining a certain good, trading could fit in as a productive activity. Trading could involve moving the product from one place to another (a physical transformation) as well as intervening in the market by using their superior skills in collecting, gathering and processing information (a cultural transformation).

Therefore, according to this view, the treatment of the trader as a regular processor is adequate.

The processor or trader profit function can be described by

$$\pi_t^p = P_t^d q_t^e + P_t^d q_t^d - P_t\left(q_t^e + q_t^d\right) - \frac{1}{2} d\left(q_t^e + q_t^d\right)^2 + x_{t-1}^p\left(P_{t-1}^x - P_t\right), \qquad d > 0$$

The processor must decide at *t-1* how much it will supply at time *t* of the exported good q_t^e and the domestic good q_t^d under uncertain prices, P_t^e and P_t^d. However, the processor has information about the behaviour of these prices in the past and can form expectations about their mean values and variances. Neither uncertainty about the technology of production nor any other exogenous factors that could affect the final output of both processed products is present. In order to supply these quantities, the trader faces a quadratic cost function similar to the one faced by the farmer. However, in the definition of the cost function, we can clearly distinguish two components: On the hand, the processor must buy the primary product used in the production of the two processed products, and this component is just proportional to the price paid by the input (third term of the expression above). On the other hand, there are costs inherent in the processing activity that take a quadratic form. These costs could reflect different types of cost related to the dealing activity (marketing, research, etc.), but they could also reflect the cost of a real processing activity similar to those presented above. Eventually, the cost of production could also reflect the case of joint production considered by Anderson and Danthine (1980)

On the other hand, as the farmer does not have certainty at time *t-1* about the price he will receive for his product, the trader does not have certainty about the price he will pay for the input required to produce the two processed products. This means that the output decision must be made under uncertainty of the prices of both the export and the domestic product price, and the price of the input. However, he has information about the evolution of this price in the past, as well as the rest of the agents and can form similar expectations about the mean and variance of this price. It could be said that processors, as in the case of the speculators, given economies of scale in the processing of information, could have better information about the distributions of prices. We will not consider this case, but their implications can be seen in Danthine (1978). Therefore, the definition we are giving for the cost function implies that our processor, in contrast to the rest of the agents, faces uncertainty on the cost of production as well since the input price is uncertain.

Additionally, the trader, as well as the rest of the agents, has the possibility of trading in the futures market, to assure the delivery of the input product at a certain price at time *t*. Therefore, at *t-1* he must also decide how much of the primary product he will trade in futures, given by x_{t-1}^{p}.

In is important to highlight that Anderson and Danthine (1983) consider that the processor will have two different approaches depending on the ability to adjust the quantity of input after the uncertainty is revealed. In the case that the processor has input flexibility, at *t-1* he will just decide how much he will trade in futures. This is because the amount to produce of both processed products (and the consequent demand for the input) will be decided at time *t* given the known prices for the primary product and the known prices at *t* for the exported and domestic goods. Note that in this case, the decision on trading in futures can be seen as merely speculative as there is no need to hedge against the volatility in the price of the input.

In case the processor cannot easily adjust the input at *t*, the quantity to be produced of both products (that will determine the input demand) must be decided before the uncertainty is removed on prices. This situation could reflect a long-term supply commitment that impedes making later adjustments. Moreover, in the precise case of traders, there might be contracts made with farmers that impose a commitment on the trader to buy the output the farmer supplies. This is, for example, seen in the cases of dairy companies where they signed contracts with farmers require collection of their milk regardless of the level of prices.

The treatment that Anderson and Danthine (1983) give to input inflexibility considers explicitly the case that the output price is non-random such that the marginal product is known, leading them to a similar behaviour as the farmer or the storage company. However, the possibility of random prices was considered but with futures markets for the output. In fact, the separation result is restored when future markets exist for the processed products. We will reflect on the situation in which the processor has input inflexibility but where it faces price risk for its output since it tends to reflect more accurately the situation in the trading of commodities, as we have seen. Consequently, the processor or trader wants to maximise the following mean-variance utility function:

$$\Omega_t^p = \overline{\pi}^p(t,t-1) - \frac{1}{2} a_p \sigma_{\pi^p}^2(t,t-1) \tag{1.13}$$

where a_p reflects the risk aversion of the trader and $\overline{\pi}^p(t,t-1)$ and $\sigma_{\pi^p}^2(t,t-1)$ are the expected profits and expected variance of the profits of the trader. Their definitions are given by

$$\overline{\pi}^p(t,t-1) = \overline{P}_t^e q_t^e + \overline{P}_t^d q_t^d - \overline{P}_t\left(q_t^e + q_t^d\right) - \frac{1}{2} d\left(q_t^e + q_t^d\right)^2 + x_{t-1}^p\left(P_{t-1}^x - \overline{P}_t\right) \tag{1.14}$$

$$\sigma_{\pi^P}^2(t,t-1) = \sigma_P^2(t,t-1) * \left(q_t^e + q_t^d + x_{t-1}^p\right)^2 + \left(q_t^d\right)^2 \sigma_{Pd}^2 + 2q_t^d q_t^e \sigma_{Pd,Pe} + (q_t^e)^2 \sigma_{Pe}^2 - 2\left(q_t^e + q_t^d + x_{t-1}^p\right) * \left(q_t^e \sigma_{P,Pe} + q_t^d \sigma_{P,Pd}\right) \tag{1.15}$$

where σ_{Pd}^2 and σ_{Pe}^2 are the expected variances for time *t* of the domestic and the export price conditional on the information available at *t-1*, respectively; $\sigma_{Pd,Pe}$ is the expected covariance of the export and the domestic price conditional on the information available at *t-1*; and $\sigma_{P,Pe}$ and $\sigma_{P,Pd}$ are the expected covariances of the primary product price and the exported and the domestic good price respectively conditional again on the information available at *t-1*.

Since the processor has historical information about all prices in the model, he can form expectations of these variances and covariances. Here we stress the fact that the quantity to be produced must be decided before the uncertainty on prices is removed, reflecting the necessity of forming those expectations in the absence of futures markets for the processed products. Inserting equations (1.14) and (1.15) into expression (1.13) and differentiating respect to the quantity supplied of the exported and domestic good and the quantity to trade in futures will yield the following first order conditions. To simplify the notion, $\sigma_P^2(t,t-1) = \sigma_P^2$

$$\overline{P}_t^e - \overline{P}_t - d\left(q_t^e + q_t^d\right) - a_p \begin{bmatrix} \sigma_P^2\left(q_t^e + q_t^d + x_{t-1}^p\right) + q_t^e \sigma_{Pd,Pe} + q_t^e \sigma_{Pe}^2 \\ -\left(q_t^e + q_t^d + x_{t-1}^p\right)\sigma_{P,Pe} - \left(q_t^e \sigma_{P,Pe} + q_t^d \sigma_{P,Pd}\right) \end{bmatrix} = 0 \tag{1.16}$$

$$\overline{P}_t^d - \overline{P}_t - d\left(q_t^e + q_t^d\right) - a_p \begin{bmatrix} \sigma_P^2\left(q_t^e + q_t^d + x_{t-1}^p\right) + q_t^d \sigma_{Pd,Pe} + q_t^d \sigma_{Pd}^2 \\ -\left(q_t^e + q_t^d + x_{t-1}^p\right)\sigma_{P,Pd} - \left(q_t^e \sigma_{P,Pe} + q_t^d \sigma_{P,Pd}\right) \end{bmatrix} = 0 \qquad (1.\ 17)$$

$$P_{t-1}^x - \overline{P}_t - a_p\left[\sigma_P^2\left(q_t^e + q_t^d + x_{t-1}^p\right) - \left(q_t^e \sigma_{P,Pe} + q_t^d \sigma_{P,Pd}\right)\right] = 0 \qquad (1.18)$$

These first order conditions are analogous to those found by Anderson and Danthine (1980; 1981) when they consider multiple cash goods and futures markets. However, since they have not used parametric functions, they expressed the variance and covariance matrix of all prices in their definition of the first order conditions. The use of quadratic cost functions allows us to obtain linear expression for the solutions. Nevertheless, the coefficients of the expressions will be composed of several interactions between all the expected variances and covariances of all prices. It is the introduction of multiple processed goods that presents this complexity in these first order conditions, and that creates the mathematical and analytical complexity for the analysis.

It would be possible to make some reasonable assumptions about the expected covariances and variances in order to reduce the complexity of these expressions. For example, we could assume that the covariances between the input price and the price for the exported and domestic good are equal and reduce the length of the expressions. However, we will leave the analysis as free of such assumptions as possible for the moment, until the analytical analysis demands such simplifying assumptions. Therefore, we allow the model to inform later about the sign and size of those parameters by paying the price of long and complex expressions but receiving the benefits of more understanding of the intuition underlying.

Nevertheless, a theoretical note is convenient to make in this aspect. Any assumption made about the value of the expected variances and covariances must be based on the assumption that the system is stable. If this assumption were not made, given the definition of these expected statistical moments, changes in prices would affect them. In turn, this would generate changes in the values of the parameters and coefficients involved that makes it impossible to establish any assumption about their values. If, on the other hand, we assume the system is stable, changes in prices exert negligible changes in the expectations formed about variances and covariances, implying that the coefficients will not change substantially.

The problem lies in the way expectations are formed. In general, it is unlikely that expectations would not change if later events prove to be against them. Any agent would desire to revise their expectations if newer information suggested that prices, for example, were unstable. However, this expectation adjustment mechanism is not defined and its treatment is particularly complicated. Any alternative specification on the way expectations are formed and/or are adjusted requires judgments about what relevant information may change these expectations and the previous periods relevant in the formation of such expectations. These judgements are generally ad-hoc and can be challenged, and may have important effects on the framework we are presenting here.

Consequently, it is not only convenient in terms of the analysis not to make assumptions about the values of the variances and covariances and their effects on coefficients. It presents theoretical aspects that complicate the analysis. We will eventually assume, to perform the econometric estimations later in this research, that the system is stable or that recent events do not alter the values assigned to these parameters. For the moment, we prefer to leave the treatment free of these assumptions.

1.3.3. Participation in the futures market

From equation (1.18) we can obtain the optimal position in the futures market for the trader:

$$x_{t-1}^{p} = \frac{P_{t-1}^{x} - \overline{P}_t}{a_p \sigma_P^2} + \frac{q_t^e\left(\sigma_{P,Pe} - \sigma_P^2\right)}{\sigma_P^2} + \frac{q_t^d\left(\sigma_{P,Pd} - \sigma_P^2\right)}{\sigma_P^2} \tag{1.19}$$

The first term in equation (1.19) is the pure speculative component in the optimal futures position of the trader. It is the position that the processor will take if his output is zero $\left(q_t^e = q_t^d = 0\right)$ and it can be seen that is similar to the position taken by the pure speculator. The remaining two terms are the hedge component. Each part in the hedge component is proportional to the cash or spot position, q_t^e and q_t^d, and they are more complex than those for the farmer and the storage company are. Note here, that the hedging component is not only affected by the variance of the price of the primary product, but it also affected by how this price varies with respect to the prices of both processed products. This suggests that the evolution of the two processed product prices also affect the hedging decision against the volatility in the price of the input.

We can also see that the hedging components (last two terms) are different from those found, for example, for the farmer, where they tend to replicate the behaviour observed in the output decisions that were made under complete certainty. The hedging components are affected by the expected variance of the primary product, introducing uncertainty in the hedging component. Therefore, the separation result that we have found for the farmer or the storage company does not hold in this context.

We can use equation (1.19) to perform some simple comparative static analysis about the position taken by the processor in the future markets. For example, if both expected covariances between the prices for the processed product and the primary product are negative, the hedging components will be negative. Eventually they could offset for the speculation component, leading to a non-participation in the futures market if the speculative component is positive or contango prevails (a case where the futures price overestimates the expected price). It might eventually surpass the speculative component, making the position of the trader in the futures market long.

However, assume that the bias between the futures price and the expected price is zero such that gains cannot be made by speculation. In this case, the hedging component will be the only determinant of the position in the futures market. This will also hold for farmers and the storage company. In fact, if the bias is zero, we should have that farmers and storage companies will always take short positions since the hedging component will always be positive. Under negative covariances between the processed product and the input price, the hedge component for the processor will be negative, leading the processor to a long position in the futures market. Consequently, in this case, the position taken by processors will be complementary of the position taken by farmers and storage companies. Alternatively, while the farmer and the storage company want to sell futures, the processor wants to buy.

On the other hand, even a positive covariance between the input and the processed product price, lower than the variance of the primary product price, could lead to a long position. Hirshleifer (1988) found that the position taken by the processor is always long since the covariance between the income and the input price is negative if there is input flexibility.

In our framework, with input inflexibility, we have found that a similar result also holds with a positive covariance between the output and the input price. The fact that definitely a negative but also a positive covariance could generate a long position (as long as no bias is present), allows us to give a high probability to this situation.

In general, the probability of a positive covariance between the input price and the output price is small. This is because, in general, as the output is increased the output price will tend to go down and the input price will tend to go up. Of course, if normal backwardation prevails, under negative covariances the position will be definitely long. Given negative covariances are more realistic and likely, the position of the trader will be long if the bias is zero or there is normal backwardation, but also if the size of speculative component under the presence of contango cannot offset the negative hedging components. Moreover, as we have seen, even with positive covariance this result would hold. Consequently, traders will have a tendency to go long in the model that is consistent with previous findings in the literature.

Despite the fact that Anderson and Danthine (1980) analyse multiple cash positions in their approach, they did not explicitly consider whether the processor position in the futures market would be altered by the consideration of multiple cash goods. If the covariances between both processed products' prices and the primary product price have opposite signs and offset the last two terms in the expression above, the last two terms would sum to zero in equation (1.19). This implies that the decision about participating in the futures markets is exclusively governed by speculation, and if the bias is zero the processor will not participate in the futures market. This is because the differences in the covariances between both processed products provide sufficient hedging for the processor and there is no need to engage in future market operations just for hedging, and given that the bias is zero, there are no speculative motives. On the other hand, the possibility that one covariances is positive and outweighs the other negative covariance, could lead to a short position in the futures market rather than a long one; this will only occur if contango prevails. This will be the only possibility in which the processor could be short. However, since contango is unlikely, both theoretically and in reality as we have seen and as the literature suggests, this possibility is unlikely.

The size of the final position will be determined by the quantity produced of the two processed products. If the price of the first product has a higher covariance with the input price than the second product (in absolute value) and more of the first product is produced, the position will be "longer" than if the second product has a higher output. This is because part of the hedging is obtained by increasing the production in the product with the lowest covariance between the input and the output price. In fact, as we will see, the covariance between the output and the input price of the other product will affect the sensitivity of the supply of the first good with respect to their own price.

Consequently, we have seen that only under very restrictive and unlikely conditions, will the trader take a short position or a selling position in the futures market, suggesting that traders, in general, will buy futures contracts. However, the effects that this decision will have on the bias of the future price still need to be analysed. This is done later in this chapter.

1.3.4. The traders supply decision

Solving the system presented by equations (1.16), (1.17) and (1.18) will give the supply for both processed products as well as the quantity to trade in futures by the processor. This will allow us to have a clearer understanding of the output and hedging decisions.

$$q_t^e = \frac{\left\{ \begin{array}{l} \overline{P}_t^e\left(d\sigma_P^2 + a_p k_{PPd}\right) + \overline{P}_t^d\left(-d\sigma_P^2 + a_p\delta\right) + \overline{P}_t\left[-d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) + a_P\lambda_e\right] \\ + P_{t-1}^x\left[d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) - a_p\left(k_{PPd} + \delta + \lambda_e\right)\right] \end{array} \right\}}{a_p\left[d\left(k_{PPd} + 2\delta + k_{PPe}\right) + a_P\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]} \tag{1.20}$$

$$q_t^d = \frac{\left\{ \begin{array}{l} \overline{P}_t^d\left(d\sigma_P^2 + a_p k_{PPe}\right) + \overline{P}_t^e\left(-d\sigma_P^2 + a_p\delta\right) + \overline{P}_t\left[d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) + a_P\lambda_d\right] \\ + P_{t-1}^x\left[-d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) - a_p\left(k_{PPe} + \delta + \lambda_d\right)\right] \end{array} \right\}}{a_p\left[d\left(k_{PPd} + 2\delta + k_{PPe}\right) + a_P\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]} \tag{1.21}$$

$$x_{t-1}^{p}=\frac{\begin{Bmatrix}\overline{P}_t^d\left[-a_p\left(\delta+k_{PPe}+\lambda_d\right)-d\left(\sigma_{P,Pe}-\sigma_{P,Pd}\right)\right]+\overline{P}_t^e\left[-a_p\left(\delta+k_{PPd}+\lambda_e\right)+d\left(\sigma_{P,Pe}-\sigma_{P,Pd}\right)\right]\\+\overline{P}_t\left[-dz-a_P\left(m+\lambda_d+\lambda_e\right)\right]+P_{t-1}^x\left[dz+a_p\left(k_{PPd}+2\delta+k_{PPe}+m+2\left(\lambda_d+\lambda_e\right)\right)\right]\end{Bmatrix}}{a_p\left[d\left(k_{PPd}+2\delta+k_{PPe}\right)+a_P\left(\sigma_P^2 m-\sigma_{Pe}^2\sigma_{P,Pd}^2+2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe}-\sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]}$$

(1.22)

Where,

I) $k_{PPe}=\sigma_P^2\sigma_{Pe}^2-\sigma_{P,Pe}^2\geq 0$ [2]

II) $k_{PPd}=\sigma_P^2\sigma_{Pd}^2-\sigma_{P,Pd}^2\geq 0$

III) $m=\sigma_{Pd}^2\sigma_{Pe}^2-\sigma_{Pe,Pd}^2\geq 0$

IV) $\delta=-\sigma_P^2\sigma_{Pe,Pd}+\sigma_{P,Pd}\sigma_{P,Pe}$

V) $\lambda_e=\sigma_{Pe,Pd}\sigma_{P,Pd}-\sigma_{Pd}^2\sigma_{PPe}$

VI) $\lambda_d=\sigma_{Pe,Pd}\sigma_{P,Pe}-\sigma_{Pe}^2\sigma_{PPd}$

VII) $z=\sigma_{Pd}^2+\sigma_{Pe}^2-2\sigma_{Pe,Pd}\geq 0$

As we can see, these expressions are very complex and the output and hedging decisions are not as clear as in the case of the farmer or the storage company. The coefficients and parameters of the equation appear represented in their basic forms. Since we have not made assumptions about the values or relationships between the variances and covariances, every element present in each parameter is clearly expressed. In reality, as we have discussed, these parameters tend to be constant or stable, since the expected variances and covariances are expected to be subject to little variation. Whilst prices may observe important variation, one can assume that agents will tend to be sluggish to adjust their expectations. Therefore, in a more general or applied context, they can receive similar treatment to any elasticity or other parameter in any equation. However, in this

[2] By the property that establishes that the square of the covariance of two random variables cannot exceed the product of their variances.

stage of the development of the model, it is necessary to leave them with their definitions in order to gain insight in the analysis.

Despite the lack of elegance and length, these expressions are linear; and the parameters involved are repeated in different parts of the equations. For example, the coefficient that multiplies the expected domestic price in the supply of exports is similar to the coefficient that multiplies the expected exported price in the domestic supply function, suggesting a similar cross price effect between both prices. On the other hand, it can be seen that in each supply function the coefficient that multiplies the expected input price is present in the coefficient that multiplies the future price, revealing the hedging nature of futures prices. These elements highlight all the hedging and cross-hedging components of the equations and the high degree of symmetry that exists between them.

We can see that the separation result that we encountered for the farmer and the storage company does not hold. The output decisions are influenced by the attitude to risk, expected prices, and the variances and covariances of the different prices. In contrast to the farmer, the output decision is not made under certainty. The separation result is restored once futures markets are introduced for the processed products or as long as the processor has input flexibility, as suggested by Anderson and Danthine (1983)[3] and Hirshleifer (1988). The separation result fails to hold because the output price of the trader is random or the marginal product is stochastic and the hedging will depend on the effects of input and output prices. In the case of input flexibility, the separation result is restored since the optimal output is chosen once the uncertainty on prices is resolved. Furthermore, the existence of futures markets in the case of input inflexibility has important implications in terms of resource allocation. A quasi-fixed additional input (investment) decided at time *t-1* but keeping the variable input choice after the uncertainty is resolved, will also affect the output decision when futures are considered (Moschini and Lapan, 1992).

As can be seen, it is hard to determine analytically the sign of the coefficients in the equations without making assumptions about the value of the variances and covariances involved in their constitution. They are affected by the expected variances and

[3] The approach we present here can also represent a real processor that produces different products using a common input and in which there are future markets for the processed. Future markets exist for soybean oil and soybean meals.

covariances, as well as other structural parameters. In the supply functions, for example, as long as the denominator is positive, it can be seen that the own price coefficient is definitely positive no matter the value of the covariance. Consequently, an increase in the expected price for the exported product, for example, will generate an increase in the supply of exports. On the other hand, if the covariance between the input and the output prices for both processed products are negative (a case that we have seen is close to reality), the covariance between both processed products must be positive. The reasoning is simple. If the covariance between the exported product price and the input price is negative, for example, it implies that when one price is increasing the other is decreasing. If this holds also for the domestic product, it means that both prices for the processed products move in the same direction, leading to a positive covariance between them.

Through this process, it might be possible to identify the signs of the coefficients in the equations. However, in this case, we will *a priori* be defining the behaviour that these expressions will have, based on *ad hoc* assumptions that we might formulate. It might be possible, of course, to calculate the variance and covariance of these prices and obtain their values. However, even the knowledge of the sign of the covariance is not enough to determine the signs of the different parameters involved in the expressions above, since we also need to know how large these values are with respect to others. As long as additions and/or subtractions of expected variances and covariances are involved, we need to have an idea of their sizes to determine the sign of some expressions.

The fact that the variables involved in the coefficients are repeated in different parts of the equations reveals the hedging and cross hedging that exists in these decisions. The volatility of the profit could be eventually reduced by not only the operation in the futures market for the input, but also by taking different positions in the cash goods.

It is interesting to see that the own price response is more governed by the volatility of the other price as well as the relative volatility with respect to the input price. Depending on the value, a high variance in the export price could lead to an increase in the supply of the domestic good. Note that the coefficient that multiplies the own price in the domestic supply function is affected by the export product variance and the covariance between the price of the input and the export price, as the parameter k_{PPe} suggests in the domestic supply function. Therefore, the own price response (the parameter that multiplies the own price) is affected by the perception of the volatility of the price of the other product rather than the expectation of the volatility of the own price.

It is interesting to highlight the role of the futures price in the supply decision. As can be seen in equations (1.20) and (1.21) the coefficient that multiplies the futures price is composed of several elements present in the rest of the coefficients. It can be seen that all the elements present in the coefficient that multiplies that expected input price, appear in the coefficient of the futures price with the opposite sign. This reflects the hedging effect of the futures price on the supply of both products.

However, it can also be seen that the futures price coefficient is composed of additional elements. Taking equation (1.20), for example, the parameter k_{PPd}, defined in definition II above, is present in the coefficient that multiplies the expected export price and, with the opposite sign, in the coefficient that multiplies the futures price. The same applies to the parameter δ (definition IV) present in the coefficient that multiplies the expected domestic price and, with the opposite sign, in the coefficient that multiplies the futures price. The definitions of both parameters, k_{PPd} and δ, include the covariance between both processed product prices and the input price. The fact that both parameters are present in the definition of the coefficient of the future price reflects that the futures markets is used to hedge not only against the specific volatility of the input price, but also against the effect that this volatility may exert on the supply through the prices of both processed products. Additionally, it can be seen that in the absence of bias in the future price, such that $P_{t-1}^{x} = \bar{P}_t$, the coefficient that multiplies the future price will only contain these two parameters. The operation in the futures market will help to offset part of the volatility in the prices of the export and the domestically supplied product.

So far, it is very hard to get a definite idea of the behaviour of the producer since we need to have information about the values of the covariances and variances involved in the determination of his supply and hedge decisions. Even the assumption about the sign of some of the covariances is not enough to determine the final sign of the coefficients of the equations. However, some comparative statics could help to shed some light on how processors will behave under different conditions or special cases. This will also help us to verify and compare with some results found in the literature. We can see that when the cost of processing tends to infinity, $d \to \infty$, the supply functions and the quantities to trade in futures will be

$$q_t^e = \frac{\sigma_P^2\left(\overline{P}_t^e - \overline{P}_t^d\right) + \left(\sigma_{P,Pe} - \sigma_{P,Pd}\right)\left(P_{t-1}^x - \overline{P}_t\right)}{a_p\left[\left(k_{PPd} + 2\delta + k_{PPe}\right)\right]}$$

$$q_t^d = \frac{\sigma_P^2\left(\overline{P}_t^d - \overline{P}_t^e\right) - \left(\sigma_{P,Pe} - \sigma_{P,Pd}\right)\left(P_{t-1}^x - \overline{P}_t\right)}{a_p\left[\left(k_{PPd} + 2\delta + k_{PPe}\right)\right]}$$

$$x_{t-1}^p = \frac{-\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right)\left(\overline{P}_t^e - \overline{P}_t^d\right) + z\left(P_{t-1}^x - \overline{P}_t\right)}{a_p\left[\left(k_{PPd} + 2\delta + k_{PPe}\right)\right]}$$

In contrast to the case of the farmer, where it can be seen that when the cost of production tends to infinity the farmer ceases production (equation (1.5) tends to zero) and only speculates in the future market. When the trader's costs tend to infinity there must still be a supply decision to be made. In this case, the trader can be seen as a pure intermediary that only buys the primary product and sells it to two different markets without costs. The commercialisation decision in this case depends on the differences between the expected prices of both processed products and on how these products behave with respect to the price of the input.

Additionally, it can be seen that the cost of processing is not a determinant of the separation between the output (pure commercialisation in this case) and hedging decisions. This means that a pure intermediary must jointly decide the amount to trade in futures and the cash positions as well as in the more general case.

It is interesting to see in this context of high processing costs that when the expected prices are equal, the supply of both products is equally determined, since the differential supply of both products is explained by its effects on the reduction of the volatility of profits and not by its effect on the level of them. Since expected prices are equal, both products contribute equally to the level of expected profits; but the different supply is explained by the capability of one of the products of providing less volatility in expected profits with respect to the other. However, we reach a first limit given by the linearity of the model. If both expected prices are equal, in this context, necessarily one of the supplied products should be negative. Therefore, we should be careful when doing these analyses as we can easily reach situations without economic sense. Consequently, rather than saying that both prices are equal, we should say that they tend to be equal to avoid this possibility.

Extending the special conditions even further, we note that, in the case that the prices are equal and the covariances between both processed products and the input prices are equal as well, the commercialisation of both products (now just the same input product) are zero. This implies that there are no profits to be made in dealing with these products. In this case, the processor will behave as a pure speculator and only the last term in the futures equation remains. Finally, if both expected processed product prices are equal and the bias in the futures price is zero, the processor disappears completely since the output will be zero; also there are no speculative gains to exploit.

Whilst the linearity of this model creates some analytical problems, it will prove to be very convenient at the time to econometrically estimating it. If we assume that in the long run the expected variances and covariances tend to be stable and their values are not affected substantially by changes in the prices that explain them, we can express equation (1.20) as

$$q_t^e = a_1^e \bar{P}_t^e + a_2^e \bar{P}_t^d + a_3^e P_{t-1}^x + a_4^e \bar{P}_t \tag{1.23}$$

where

$$a_1^e = \frac{d\sigma_P^2 + a_p k_{PPd}}{a_p\left[d(k_{PPd} + 2d + k_{PPe}) + a_p\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]}$$

$$a_2^e = \frac{-d\sigma_P^2 + a_p\delta}{a_p\left[d(k_{PPd} + 2d + k_{PPe}) + a_p\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]}$$

$$a_3^e = \frac{d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) - a_p(k_{PPd} + \delta + \lambda_e)}{\left[d(k_{PPd} + 2d + k_{PPe}) + a_p\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]}$$

$$a_4^e = \frac{-d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) + a_p\lambda_e}{\left[d(k_{PPd} + 2d + k_{PPe}) + a_p\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]}$$

Making similar replacements, is possible to rewrite equation (1.21) as

$$q_t^d = a_1^d \bar{P}_t^d + a_2^d \bar{P}_t^e + a_3^d P_{t-1}^x + a_4^d \bar{P}_t \tag{1. 24}$$

Equations (1.23) and (1.24) can be estimated econometrically using the appropriate estimation techniques. The econometric estimation allows identifying the signs and the sizes of these coefficients with more precision as well as verifies the empirical validity of the equations. This will be the task to be implemented in later stages of this research.

1.3.5. Restrictive conditions

From equations (1.20) and (1.21) we get some insight about the processor supply decisions. We will see that even under very restrictive conditions, the quantities to be exported and supplied domestically can be decided without having any information about the demand prevailing in time *t*. Therefore, this model will be demand driven only under very restrictive assumptions.

It is convenient to note, as we have seen in the case of a pure intermediary, that if the expected export and domestic prices are equal, the final planned supply is still independently determined. It will depend eventually on the effect that expected variances and the covariances of both products have on the coefficients of the equations, as well as the covariance between the processed products prices and the input prices. Furthermore, the supply decisions are explained not only by differences in expected prices but also by differences in the expected variances and covariances on the prices since the processor is also concerned about the variability of the profits. If we assume that the expected prices for the exported and the domestic product are equal, equation (1.20), for example, becomes

$$q_t^e = \frac{\left\{\begin{aligned}&\overline{P}_t^{e,d} a_p\left(k_{PPd} + \delta\right) + \overline{P}_t\left[-d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) + a_P\lambda_e\right]\\ &+ P_{t-1}^x\left[d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) - a_p\left(k_{PPd} + \delta + \lambda_e\right)\right]\end{aligned}\right\}}{a_p\left[d\left(k_{PPd} + 2\delta + k_{PPe}\right) + a_P\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]} \tag{1.25}$$

where $\bar{P}_t^{e,d} = \bar{P}_t^e = \bar{P}_t^d$ is the common expected price for both products. We will only consider the expression for the exported product. The expression for the domestic product is similar. For example, a negative response to the expected exported product price could appear (if the processor is risk averse) if the covariance between both processed products $(\sigma_{Pe,Pd})$ is greater than the variance of the domestic price (σ_{Pd}^2), and if the covariance between the exported product and the input price ($\sigma_{P,Pe}$) is lower than the covariance between the domestic product and the input price ($\sigma_{P,Pd}$). This can be seen by adding parameters II and IV. In this case, the coefficient that multiplies the common expected price will be

$$a_p\left(k_{PPd}+\delta\right)=a_p\left[\sigma_P^2\left(\sigma_{Pd}^2-\sigma_{Pe,Pd}\right)+\sigma_{P,Pd}\left(\sigma_{P,Pe}-\sigma_{P,Pd}\right)\right]$$

In that case, the processor will find that profits exhibit less variability, which implies that the processor will reduce (increase) the supply of the exported (domestic) product when the common expected price increases. Nevertheless, it is a matter of discussion if this situation can effectively happen. Whilst this situation is possible, its probability of occurrence is an empirical question.

Whilst the expected prices are equal, the variances and covariances could take different values. We will explore how this model can further determine the level of supply of both processed products by assuming different values of the covariances. A candidate is to assume that the covariances between the export and the input price are the same as the covariance between the domestic and the input price $\left(\sigma_{P,Pe}=\sigma_{P,Pd}\right)$, and keeping the equality between the expected prices. In this case, equations (1.20) and (1.21) will become

$$q_t^e=\frac{\left\{\left(\sigma_{Pd}^2-\sigma_{Pd,Pe}\right)\left(-\overline{P}_t\sigma_{P,Pe,Pd}+\sigma_P^2\overline{P}_t^{e,d}+P_{t-1}^x\left(\sigma_{P,Pe,Pd}-\sigma_P^2\right)\right)\right\}}{\sigma_P^2\left(-\sigma_{Pd,Pe}\left(a_p\sigma_{Pd,Pe}+2d\right)+\sigma_{Pd}^2\left(a_p\sigma_{Pe}^2+d\right)+d\sigma_{Pe}^2\right)-a_p\sigma_{P,Pe,Pd}\left(\sigma_{Pd}^2+\sigma_{Pe}^2-2\sigma_{Pd,Pe}\right)} \tag{1.26}$$

$$q_t^d=\frac{\left\{\left(\sigma_{Pe}^2-\sigma_{Pd,Pe}\right)\left(-\overline{P}_t\sigma_{P,Pe,Pd}+\sigma_P^2\overline{P}_t^{e,d}+P_{t-1}^x\left(\sigma_{P,Pe,Pd}-\sigma_P^2\right)\right)\right\}}{\sigma_P^2\left(-\sigma_{Pd,Pe}\left(a_p\sigma_{Pd,Pe}+2d\right)+\sigma_{Pd}^2\left(a_p\sigma_{Pe}^2+d\right)+d\sigma_{Pe}^2\right)-a_p\sigma_{P,Pe,Pd}\left(\sigma_{Pd}^2+\sigma_{Pe}^2-2\sigma_{Pd,Pe}\right)} \tag{1.27}$$

where $\sigma_{P,Pe,Pd}$ reflects the common covariance between the exported product price, the domestic product price, and the input price. We have replaced back the parameters defined by I to VII by their adjusted definitions. Note that the two equations are very similar. The only parameter that differs in these equations is the difference between the respective processed product variances and the covariance between the two products

$(\sigma_{Pe}^2 - \sigma_{Pd,Pe})$ and $(\sigma_{Pd}^2 - \sigma_{Pd,Pe})$, reflecting the role of taking crossed positions to reduce the volatility in profits. This means that the closer the expected variance of the exported product is to the covariance of both processed products, for example, the greater will be the specialisation in the exported product. Nevertheless, the only possibility of this result is violating the non-negativity restriction on some of the prices. Effectively, the only possibility that these combinations of values of expected variances, covariances and expected prices, occur is if some of the prices assume negative values. Given the linearity of the model, this is mathematically possible, but economically meaningless.

However, it is interesting to see the cross effect. If, for example, the variance of the exported product is close to the covariance between the two processed prices the processor will tend to supply only the exported product. To see this we need to analyse what would be the value of the expected variance of the domestic product. In this case, the expected variance of the domestic product will be always greater than the expected variance of the exported product by parameter VII. If the expected variance of the domestic price is greater than the expected variance of the exported product, the processor will find that increasing the supply of the exported product can reduce the variance of the profits, despite both expected prices being equal. However, this result is again only mathematically possible by lifting the restriction of non-negativity on prices.

If, on the other hand, the two variances are equal, the model turns into a demand driven model since the two products have the same expected prices and variances, and only the information given by the demand will help the processor to decide how much must be allocated in each market. Finally, from equations (1.20) and (1.21), we can see that, even in the case that one of the expected prices tends to zero, there will still be supply of that product. In that case, the decision of supplying that product is governed by the possibility of hedging against fluctuations in the processed product prices. Although the expected price could be zero, the expected variance may not. Although unlikely, it is acknowledged as a possibility.

1.4. THE COMPLETE MODEL

The final agents to be described in this framework are the consumers of the two processed goods. These agents are necessary in this framework to analyse the properties and the behaviour of the equilibrium of the model. The domestic and foreign consumers demand their respective goods following non-stochastic linear demand functions that only depend on the spot price of those products at *t*.

$$D_t^d = A_d - \phi_d P_t^d \tag{1.28}$$

$$D_t^e = A_e - \phi_e P_t^e \tag{1.29}$$

Where $\phi_d > 0$ and $\phi_e > 0$. In this framework, both demands are completely independent of each other. Although both goods are similar, in essence it is impossible for the consumer abroad to sell the exported product to a domestic consumer. A richer and more interesting model will be possible if we consider a second set of producers, storage companies and processors located in a different part of the world in such a way that consumers can choose to consume the imported or domestic good. In that case, a certain degree of substitutability between both products will affect the behaviour of prices. However, in that case, the problem will eventually become non-linear, particularly when, for example, the Armington (1969) assumption of imperfect substitution between origins is used.

We will assume that there are a number n_i of homogenous agents of each type. Consequently, the supply equations and the future trading equations can be multiplied by the respective number of agents to get the aggregate supply functions. Finally, we need to establish the market clearing conditions. There are four equilibrium equations: one for each processed product, one for the spot market of the primary product, and one for the future markets. Including these equations, the complete model can be summarised as

$$Q_t = n_f \frac{P_{t-1}^x}{c} + \Phi_t \tag{1.30}$$

$$X_{t-1}^f = n_f \left(\frac{P_{t-1}^x}{c} + \frac{P_{t-1}^x - \overline{P}_t}{a_f \sigma_P^2} \right) \tag{1.31}$$

$$I_{t-1} = I^* + n_c \left(\frac{P_{t-1}^x - \rho P_{t-1}}{h} \right) \tag{1.32}$$

$$X_{t-1}^{c} = I^{*} + n_{c}\left(\frac{P_{t-1}^{x} - \rho P_{t-1}}{h} + \frac{P_{t-1}^{x} - \overline{P}_{t}}{a_{c}\sigma_{P}^{2}}\right) \tag{1.33}$$

$$Q_{t}^{e} = n_{p}\frac{\left\{\begin{array}{l}\overline{P}_{t}^{e}\left(d\sigma_{P}^{2} + a_{p}k_{PPd}\right) + \overline{P}_{t}^{d}\left(-d\sigma_{P}^{2} + a_{p}\delta\right) + \overline{P}_{t}\left[-d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) + a_{P}\lambda_{e}\right]\\ + P_{t-1}^{x}\left[d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) - a_{p}\left(k_{PPd} + \delta + \lambda_{e}\right)\right]\end{array}\right\}}{a_{p}\left[d\left(k_{PPd} + 2\delta + k_{PPe}\right) + a_{P}\left(\sigma_{P}^{2}m - \sigma_{Pe}^{2}\sigma_{P,Pd}^{2} + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^{2}\sigma_{P,Pe}^{2}\right)\right]} \tag{1.34}$$

$$Q_{t}^{d} = n_{p}\frac{\left\{\begin{array}{l}\overline{P}_{t}^{d}\left(d\sigma_{P}^{2} + a_{p}k_{PPe}\right) + \overline{P}_{t}^{e}\left(-d\sigma_{P}^{2} + a_{p}\delta\right) + \overline{P}_{t}\left[d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) + a_{P}\lambda_{d}\right]\\ + P_{t-1}^{x}\left[-d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) - a_{p}\left(k_{PPe} + \delta + \lambda_{d}\right)\right]\end{array}\right\}}{a_{p}\left[d\left(k_{PPd} + 2\delta + k_{PPe}\right) + a_{P}\left(\sigma_{P}^{2}m - \sigma_{Pe}^{2}\sigma_{P,Pd}^{2} + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^{2}\sigma_{P,Pe}^{2}\right)\right]} \tag{1.35}$$

$$X_{t-1}^{p} = n_{p}\frac{\left\{\begin{array}{l}\overline{P}_{t}^{d}\left[-a_{p}\left(\delta + k_{PPe} + \lambda_{d}\right) - d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right)\right] + \overline{P}_{t}^{e}\left[-a_{p}\left(\delta + k_{PPd} + \lambda_{e}\right) + d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right)\right]\\ + \overline{P}_{t}\left[-dz + a_{P}\left(-m - \lambda_{d} - \lambda_{e}\right)\right] + P_{t-1}^{x}\left[dz + a_{p}\left(k_{PPd} + 2\delta + k_{PPe} + m + 2\left(\lambda_{d} + \lambda_{e}\right)\right)\right]\end{array}\right\}}{a_{p}\left[d\left(k_{PPd} + 2\delta + k_{PPe}\right) + a_{P}\left(\sigma_{P}^{2}m - \sigma_{Pe}^{2}\sigma_{P,Pd}^{2} + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^{2}\sigma_{P,Pe}^{2}\right)\right]} \tag{1.36}$$

$$X_{t-1}^{s} = n_{s}\frac{P_{t-1}^{x} - \overline{P}_{t}}{a_{s}\sigma_{P}^{2}} \tag{1.37}$$

$$D_{t}^{d} = A_{d} - \phi_{d}P_{t}^{d} \tag{1.38}$$

$$D_{t}^{e} = A_{e} - \phi_{e}P_{t}^{e} \tag{1.39}$$

$$Q_{t}^{e} = D_{t}^{e} \tag{1.40}$$

$$Q_{t}^{d} = D_{t}^{d} \tag{1.41}$$

$$Q_{t} + I_{t-1} = I_{t} + Q_{t}^{e} + Q_{t}^{d} \tag{1.42}$$

$$X_{t-1}^{f} + X_{t-1}^{c} + X_{t-1}^{p} + X_{t-1}^{s} = 0 \tag{1.43}$$

Where $I^* = n_c i^*$, $\Phi_t = n_f \varphi_t$ and the capital letters represent aggregates over the individuals. In the next sections, we will analyse the equilibrium of the exported and the domestic product. Nevertheless, we will focus first on the analysis of the source of bias in the futures market and the analysis of the equilibrium in this market.

1.4.1. Equilibrium in the futures markets

Since this framework is intended to analyse the influence of futures markets in the supply of exports and domestic products, it is convenient to get some insight about the equilibrium in the futures market. This is necessary in order to verify that futures markets in the model exhibit the standard properties and characteristics found in similar analytical frameworks.

Equation (1.43) establishes that for equilibrium in the futures market the excess of demand for futures contracts traded by farmers, storage companies, processors and speculators should sum zero. Substituting equations (1.31), (1.33), (1.36) and (1.37) into equation (1.43) and solving for the futures price, yields

$$P_{t-1}^{x} = \frac{\left[\begin{array}{l} -I^* + \dfrac{n_c \rho P_{t-1}}{h} - \dfrac{n_p\left(\theta - d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right)\right)\overline{P}_t^{d}}{a_p \Phi} - \dfrac{n_p\left(\chi + d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right)\right)\overline{P}_t^{e}}{a_p \Phi} \\ + \left[\left(\dfrac{n_f}{a_f} + \dfrac{n_c}{a_c} + \dfrac{n_s}{a_s}\right)\left(\dfrac{1}{\sigma_P^2}\right) - \dfrac{n_p \tau}{a_p \Phi}\right]\overline{P}_t \end{array}\right]}{\dfrac{n_f}{c} + \dfrac{n_c}{h} + \left(\dfrac{n_f}{a_f} + \dfrac{n_c}{a_c} + \dfrac{n_s}{a_s}\right)\left(\dfrac{1}{\sigma_P^2}\right) - \dfrac{n_p\left(\theta + \chi + \tau\right)}{a_p \Phi}} \tag{1.44}$$

where

$$\theta = -a_p\left(\delta + k_{PPe} + \lambda_d\right)$$

$$\chi = -a_p\left(\delta + k_{PPd} + \lambda_e\right)$$

$$\tau = -dz - a_p\left(m + \lambda_d + \lambda_e\right)$$

$$\Phi = d\left(2\delta + k_{PPe} + k_{PPd}\right) + a_p\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)$$

Equation (1.44) establishes that the current futures price is affected by the current spot price and the expected future spot price for the primary product as well as the expected domestic and exported product price. This equation is similar to the one found by Turnovsky (1983)[4] and by Kawai (1983). The only relevant difference is the additional terms for the variances and expected prices of the domestic and the exported product, given by the addition of the second stage.

It can be seen, with the assistance of some mathematical software[5], that when the cost of storage tends to zero $(h \to 0)$, the futures price tends to ρP_{t-1}. This implies that the future price at time *t-1* for delivery at time *t*, would be the current spot price adjusted by the cost of taking that product at time *t*, a result previously found and summarised in Carter (1999), and implicit in Kawai (1983) and Turnovsky (1983).

The implications if this condition does not hold can be seen. If the future price for delivery at time *t* is greater than the price at *t-1* adjusted by the carrying costs, an opportunity of intertemporal arbitrage will exist. Therefore, if there were no costs to pay to store the primary product (more than the opportunity cost) we would have an unbiased estimator of the future price. In fact, there is no need to make any expectation about the spot price since this will be just the previous period price adjusted for the carrying cost.

It can be seen, on the other hand, that if the variance of the input price is zero, the future price is unbiased. Alternatively, in the absence of variation in the input price, the future price will perfectly forecast the expected spot price. This result is also consistent with the findings of Turnovsky (1983) and Kawai (1983). On the other hand, when the cost of

---

[4] In fact, if the number of processors, n_p, approaches zero, equation (1.44) tends to be exactly the one found by Turnovsky.

[5] Wolfram Mathematica has been used for these developments.

production of the primary product tends to zero ($c \rightarrow 0$), the futures price will be zero as the output and the quantity of futures sold by the producer will tend to infinity.

However, if the cost of processing the product tends to infinity ($d \rightarrow \infty$), it was seen that the output and the futures traded by the processor would not be zero. The processor will behave in this case more like an intermediary rather than a true processor. However, he must still hedge against fluctuations in the input price. In that case, the bias in the futures is given by

$$P_{t-1}^{x} = \frac{\left[\begin{array}{l} a_p\left(2\delta + k_{PPe} + k_{PPd}\right)\left(\dfrac{n_c \rho P_{t-1}}{h} - I^*\right) + n_p\left(\overline{P}_t^{d} - \overline{P}_t^{e}\right)\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) \\ + \left[a_p\left(2\delta + k_{PPe} + k_{PPd}\right)\left(\dfrac{n_f}{a_f} + \dfrac{n_c}{a_c} + \dfrac{n_s}{a_s}\right)\left(\dfrac{1}{\sigma_P^2}\right) + zn_p\right]\overline{P}_t \end{array}\right]}{a_p\left(2\delta + k_{PPe} + k_{PPd}\right)\left(\dfrac{n_f}{c} + \dfrac{n_c}{h} + \left(\dfrac{n_f}{a_f} + \dfrac{n_c}{a_c} + \dfrac{n_s}{a_s}\right)\left(\dfrac{1}{\sigma_P^2}\right)\right) + zn_p} \tag{1.45}$$

The bias will not disappear since the processor, now a simple intermediary, will still operate in the futures market to hedge against fluctuations in the primary product price, as well as to profit from differences between the futures and the expected price. Moreover, since there is a simplification of coefficients, the effect of the risk aversion will be reduced. Now, it does not affect the coefficients that multiply the expected exported and domestic product prices. However, we cannot say anything yet about the direction of the bias in this case. It is clear that the futures price will still be biased, but the direction will depend on the expected covariances between the input prices and the processed product prices.

The sources of backwardation or contango have been analysed extensively in the literature. In essence, if all agents want to go short (sell futures), necessarily one agent must take a complementary position (buy futures) or take a long position. That agent, the one that is taking the complementary position, will require a premium for the risk assumed and this will lead to a difference between the futures and the expected price. Consequently, a source of bias in the futures price is that the net positions of the different agents involved in the model might not be compatible.

In general contexts without second stages of production, it is seen that the bias should be negative (normal backwardation), since farmers and storage companies have a tendency to go short and in order to make speculators go long, a negative bias will be required. We can see that as long as farmers, storage companies or speculators are risk neutral ($a_f \to 0$, $a_c \to 0$ or $a_s \to 0$), the futures price is an unbiased estimator of the future spot price or $P^x_{t-1} = \overline{P}_t$, as seen in Turnovsky (1983). The definitions of some parameters include the processor's coefficient of risk aversion. In the calculation of the limits, these interactions of the coefficient of risk aversion in the parameters defined have been considered. The zero bias when producers and storage companies are risk averse is also present in Sarris (1984). When these agents tend to neutrality against risk their speculative component will tend to infinity, and they will share the speculation between them. The rest of the agents will limit their activities in the futures markets to hedge their output in a bias free market as suggested by Newbery and Stiglitz (1981). This implies that the two components of the futures decision (hedging and speculation) are separated. The speculation is resolved by these neutral to risk agents, leaving the futures markets to hedge the output of the risk averse agents. Furthermore, the action of neutral to risk agents will make the bias in the futures price disappear.

1.4.1.1. Neutral to risk traders

The idea that a two-stage production process also leads to backwardation has been suggested by Anderson and Danthine (1983) and Hirshleifer (1988) in a framework with only one cash good, even when processors have a tendency to go long in the futures market. Since the activities of processors affect the risk and hedging activities of the rest of the agents, they are forced to go short by an amount larger than the one necessary to offset the long position taken by processors. This means that although processors constitute a natural counterpart on the hedging strategies of the rest of the agents, a speculator will still be necessary to close the gap. Consequently, risk averse agents with complementary positions cannot guarantee an unbiased futures price.

So far, we have found that in this model, processors or traders have a tendency to buy futures, and we have seen that when the rest of the agents are neutral to risk, the future price is unbiased. A natural question is what happens when processors are neutral to risk. If the processors are risk neutral $(a_P \to 0)$ the current futures price is

$$P_{t-1}^{x} = \overline{P} + \frac{\left(\overline{P}_t^{e} - \overline{P}_t^{d}\right)\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right)}{z} \tag{1.46}$$

This implies that if processors were risk neutral the futures price would still be a biased estimator of future spot price. The direction of the bias will be given by the difference between the expected covariances between the primary product price and the exported and the domestic products price, respectively, and the difference between the expected prices for both processed products.

As in the case of farmers, one should expect that if processors are risk neutral, they will share the speculation between them that should generate and unbiased price. However, in contrast to the case analysed above, when farmers or store companies were risk neutral, the bias would still be present since now the agents that are short (farmers and storage companies) need to hedge their output and they will need to find an agent (speculator) willing to take the complementary position that traders are not taking. This means that neutral to risk traders imply the disappearance of the natural counterpart in the futures markets for the rest of the agents.

This chapter (as well as previous findings) has found that processors tend to be long, providing the rest of the agents with a 'natural' counterpart in the futures trade, but still with a bias in the futures price. If processors are risk neutral, their speculative component could eventually outweigh the tendency to adopt a long position driven by the hedge component in the futures position. Furthermore, a bias is necessary to attract other agents (a speculator, for example) to take a long position that would close the gap. Therefore, the short hedgers must pay a premium to hedge effectively their output. This analysis can be seen as the missing case in Hirshleifer (1988). In his analysis, farmers do not participate in the futures market, leaving the future with a positive bias since *"...while short-hedging growers are driven out, the long-hedging processors for the most part remain in the market*" (Hirshleifer, 1988, p. 1218). The case presented here would be similar to one in which the number of processors tends to zero. On the other hand, a situation in which the speculative component makes the processors position 'too long' will call for an upward bias to make the rest of the agents positions' shorter and close the gap effectively. This implies that neutral to risk processors in this model could generate either backwardation or contango. Therefore:

PROPOSITION I. *When output is stochastic, non-stochastic linear demand schedules, two processed products use a common primary product, if processors are risk neutral, the futures price for the primary product is a biased predictor of the expected spot price for the primary product*.

We have assigned a low probability to a positive covariance between both processed product prices and the input price. Therefore, the sign of the differences in the covariances will be positive (negative) as long as the expected covariance between the exported product and the input price is larger (smaller) in absolute value than the covariance between the domestic product and the input price. Eventually, if the covariances and the prices are equal, the bias will disappear.

The final sign of the bias will be determined by the differences between the expected prices for the exported and the domestic product. *A priori*, we cannot say anything about the difference between these two prices, which will lead us to the possibility of contango if the difference in expected prices has the same sign as the difference in covariances. On the other hand, the unbiasedness will be restored when common expectations are formed about both prices.

To summarise, risk neutral processors do not guarantee an unbiased price. However, as long as they have common expectations about the processed product prices or common expectations about the covariances between the processed product and the input product price, the bias will disappear.

1.5. EQUILIBRIUM IN THE EXPORT AND DOMESTIC MARKETS

Up to this point, we have focused our attention on the behaviour of traders in their supply in the exports and domestic products and in their participation in the futures market. We have found that there is a tendency in this framework, and shared by the literature, to buy futures contracts. Moreover, we have also analysed the equilibrium in futures market and have found that, although the position taken by traders tends to be complementary to the position of the rest of the agents, the future price is still biased. We need now to focus our attention on how the spot market behaves in equilibrium.

The first thing we need to find is the equilibrium in both processed products markets that are met at time *t*. When expectations are realised or the model is in equilibrium, we will have that

$$\overline{P}_t^e = P_t^e = \overline{\overline{P}}^e$$

$$\overline{P}_t^d = P_t^d = \overline{\overline{P}}^d$$

$$\overline{P}_t = P_t = \overline{\overline{P}}$$

where $\overline{\overline{P}}^e$, $\overline{\overline{P}}^d$and $\overline{\overline{P}}$ are the long-run equilibrium price for the exported, domestic and primary product respectively. On the other hand, in the absence of basic risk[6], we should expect that the futures prices at delivery should be equal to the spot price at that time or $P_t = P_{t-1}^x = \overline{\overline{P}}$. Considering this and using equilibrium conditions (1.40) and (1.41), we get.

$$\overline{\overline{P}}^e = \frac{a_p\left[n_p\overline{\overline{P}}(k_{PPd} + \delta) + \Delta A_e\right] + n_p\overline{\overline{P}}^d\left(d\sigma_P^2 - a_p\delta\right)}{a_p\left(\Delta\phi_e + n_p k_{PPd}\right) + n_p d\sigma_P^2} \tag{1.47}$$

$$\overline{\overline{P}}^d = \frac{a_p\left[n_p\overline{\overline{P}}(k_{PPe} + \delta) + \Delta A_d\right] + n_p\overline{\overline{P}}^e\left(d\sigma_P^2 - a_p\delta\right)}{a_p\left(\Delta\phi_d + n_p k_{PPe}\right) + n_p d\sigma_P^2} \tag{1.48}$$

Where

$$\Delta = d\left(k_{PPd} + 2\delta + k_{PPe}\right) + a_P\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)$$

If we further solve the system presented by equations (1.47) and (1.48), we will get the long run equilibrium price for the exported and the domestic product:

---

[6] The basis is defined as the difference between the future price for delivery at time t at maturity and the spot time at t. In reality, there is always a difference explained by the cost of delivery of the product. However, we assume that this is negligible.

$$\bar{\bar{P}}^e = \frac{\begin{array}{c} n_p\left[a_p\left(\Delta\phi_d(k_{PPd}+\delta)+n_p(k_{PPe}k_{PPd}-\delta^2)\right)+n_p d\sigma_p^2(k_{PPe}+k_{PPd}+2\delta)\right]\bar{\bar{P}}+ \\ \Delta\left(a_pA_e(\Delta\phi_d+k_{PPe}n_p)-\delta a_pA_dn_p+dn_p\sigma_p^2(A_e+A_d)\right) \end{array}}{\begin{array}{c} a_p\left[\Delta\phi_e(\Delta\phi_d+k_{PPe}n_p)+n_p\left(\Delta\phi_dk_{PPd}+n_p(k_{PPe}k_{PPd}-\delta^2)\right)\right] \\ +n_p\left[d\sigma_p^2\left(\Delta(\phi_d+\phi_e)+(k_{PPe}+k_{PPd}+2\delta)\right)\right] \end{array}} \tag{1.49}$$

$$\bar{\bar{P}}^e = \frac{\begin{array}{c} n_p\left[a_p\left(\Delta\phi_e(k_{PPe}+\delta)+n_p(k_{PPe}k_{PPd}-\delta^2)\right)+n_p d\sigma_p^2(k_{PPe}+k_{PPd}+2\delta)\right]\bar{\bar{P}}+ \\ \Delta\left(a_pA_d(\Delta\phi_d+k_{PPd}n_p)-\delta a_pA_dn_p+dn_p\sigma_p^2(A_e+A_d)\right) \end{array}}{\begin{array}{c} a_p\left[\Delta\phi_e(\Delta\phi_d+k_{PPe}n_p)+n_p\left(\Delta\phi_dk_{PPd}+n_p(k_{PPe}k_{PPd}-\delta^2)\right)\right] \\ +n_p\left[d\sigma_p^2\left(\Delta(\phi_d+\phi_e)+(k_{PPe}+k_{PPd}+2\delta)\right)\right] \end{array}} \tag{1.50}$$

Both expressions are very complex. However, the high degree of symmetry between them can be seen. The only price that intervenes in the determination of the long-run processed product equilibrium price is the long-run equilibrium price of the input. The futures price disappears in the definition of these prices. Nevertheless, as in any analysis of the solutions in the long run, all the structural parameters determine the equilibrium. Of course, these parameters are affected by the expected variances and covariances that will be in equilibrium.

Some comparative statics can be performed to verify some properties of the equilibrium. It can be seen that if the number of producers tends to zero $(n_p \to 0)$, the long run equilibrium export and domestic prices converge to ${A_e}/{\phi_e}$ and ${A_d}/{\phi_d}$, respectively. This is consistent with the rational expectations hypothesis and establishes that, in that case, the price depends only on the structural parameters of the demand functions as Turnovsky (1983) highlights.

If on the other hand, the processors are risk neutral, $a_p \to 0$, and considering the definition of the parameter Δ, the long run equilibrium price for the processed product will be

$$\overset{=}{P}^{e} = \overset{=}{P}^{d} = \frac{n_p \overset{=}{P} + d\left(A_e + A_d\right)}{n_p + d\left(\phi_d + \phi_e\right)} \tag{1.51}$$

Alternatively, in the long run, neutral to risk processors will generate the arbitrage conditions for the price of both products to be identical. This means that, in equilibrium and with risk neutral processors, trade will make domestic and international prices converge to a common price. Effectively, in the long run equilibrium, there should not be any speculative gain to make, and neutral to risk processors will assure that equality between prices will prevail. Note that both equilibrium prices will still be affected by the long-run equilibrium in the input price that is a consequence of the lack of futures prices for both processed products.

Another interesting possibility to analyse will be if the cost of processing tends to infinity, $d \to \infty$. In this case, the processor behaves as a pure intermediary and the long run equilibrium prices for both products will be

$$\overset{=}{P}^{e} = \frac{a_p A_e \phi_d \left(k_{PPe} + k_{PPd} + 2\delta\right) + n_p \sigma_P^2 \left(A_e + A_d\right)}{a_p \phi_e \phi_d \left(k_{PPe} + k_{PPd} + 2\delta\right) + n_p \sigma_P^2 \left(\phi_d + \phi_e\right)} \tag{1.52}$$

$$\overset{=}{P}^{d} = \frac{a_p A_d \phi_e \left(k_{PPe} + k_{PPd} + 2\delta\right) + n_p \sigma_P^2 \left(A_e + A_d\right)}{a_p \phi_e \phi_d \left(k_{PPe} + k_{PPd} + 2\delta\right) + n_p \sigma_P^2 \left(\phi_d + \phi_e\right)} \tag{1.53}$$

We will have two different long run equilibrium prices but the difference between them will depend only on the structural parameters of the demand functions. Finally, it can be seen that if any or both of the demand parameters ϕ_d or ϕ_e tend to infinity, the equilibrium price for that product will tend to zero, that is also a result that can be verified in Turnovsky (1983) and Kawai (1983).

Consequently, we have found that the long run equilibrium properties are consistent to the standard theory and literature on futures markets. The differences are explained by the interactions of the expected variances and covariances that generate very complex expressions in the analysis. However, as in any rational expectations equilibrium, the long run prices of the processed products depend exclusively on the structural parameters of the system.

1.6. LONG RUN SOLUTION IN THE FUTURES MARKETS

The long-run equilibrium in the futures prices needs to be analysed. Unfortunately, the resultant expression is very complex and tedious to be presented here. It is recognised that this analysis will be incomplete without a proper analysis of the equilibrium in this market. However, whilst it is extremely important for the study of the equilibrium properties of the futures price, its study is outside the scope of this chapter where the focus was in the supply decision rather than the futures market *per se*.

However, we would not finish this analysis without, at least, presenting some results in the equilibrium of the futures price that are obtained when some particular cases are observed. The long run equilibrium in the futures price can be obtained by replacing equations (1.49) and (1.50), the solutions for the long run equilibrium in the processed product market, into equation (1.44). The resulting expression, as we mentioned, is beyond reproduction.

However, we can make some simple analysis about it without the need for further development. If we assume that in equilibrium, $\bar{P}_t = P_t = P_{t-1} = \bar{\bar{P}}$, which is the condition for a rational expectations equilibrium, the resulting expression, the long run future price, will depend only on the long run equilibrium price for the input product $\bar{\bar{P}}$. The rest of the prices in the model do not affect it and only the structural parameters of the system will only have an effect on its level.

However, the bias in the futures price would still be present in the long run. If in the long run there were no bias, there would be no speculation and, since traders tend to be long, an unbiased price should emerge. However, we have seen that, although unlikely, traders might also sell futures, in which case a bias will be still necessary to attract someone to take a complementary position to the rest of the agents.

Some comparative static exercises could be performed in this case. If the number of processors tends to zero, $n_p \rightarrow 0$, the expression converges to equation (36) in Turnovsky (1983, p. 1374), where he shows that a bias will still be present in the long run. This suggests that the presence of a second stage of production is irrelevant in the determination of an unbiased futures price in the long run.

What is interesting to see is that if the coefficient of risk aversion of the processor tends to zero $a_P \rightarrow 0$, we will obtain

$$P_{t-1}^{x} = P_{t}^{x} = \bar{\bar{P}}$$

Effectively, in equilibrium, with risk neutral traders and under the assumption of no basic risk, the futures price with a two stages of production would generate neither backwardation nor contango. In other words, under the assumption of no basis risk, such as the futures price at maturity is equal to the current spot price and given that we are looking for a rational expectations equilibrium (such as the spot prices are equal to the long run average price) the futures price will eventually be equal to this price. Consequently, Proposition I, presented previously, does not hold in rational expectations equilibrium. Furthermore, although in the short run risk neutral processors could generate a bias in the futures price. This bias must be zero in the long run or in equilibrium.

The long run equilibrium in the spot price input could be obtained as well. This can be obtained by replacing equations (1.30), (1.32), (1.34), (1.35) into equilibrium condition (1.42), and considering the price equilibrium (1.49) and (1.50) together with the equilibrium price for the futures market (not presented in this chapter). However, whilst its study may help to gain additional knowledge about the model we are presenting here, we prefer go no further since we have obtained enough insights on the determination of the exports and the domestic supply under the context we have presented, which was the main objective of this chapter.

1.7. CONCLUSIONS

The existence of intermediate agents (traders) between producers and demand is a distinctive feature of agricultural commodity markets that should not be overlooked. Their trading decisions affect the spot market as well as the futures markets. This means that a more accurate representation of these markets will definitely need to incorporate them. On the other hand, commodity markets are internationally integrated and any analysis on the trade of commodities should include how futures markets and the existence of traders affect the supply in the domestic and the export markets.

This chapter has analysed how the decision of trade (in terms of supplying the world or the domestic market) is governed when risk in output of a storable agricultural commodity is present, and when futures markets are available to hedge against fluctuations in the price. In order to achieve this goal, a model that includes farmers, storage companies and processors or traders has been developed in which agents have the possibility of operating in futures markets. This model is intended to reflect the situation of agricultural commodities.

In order to develop the model, we have tried to reconcile the development of cross-hedging and second stage of production when future markets are available. We have used the role of the processor to consider the existence of traders. However, this model is flexible enough not to consider only the domestic/international trade decisions, but also to accommodate a model for more general second stage models.

The model has proved to be extremely complex in its mathematical formulation. The inclusion of a second stage of production with two markets to supply implies the interaction of complexly defined parameters and coefficients and the variables of the model. Particularly, the different interaction of prices and the expected variances and covariances of all the prices involved in the model generated very complex and unwieldy expressions. This implies that parametric approaches, such as the one applied here, may not be the most appropriate treatment for this problem. Instead, a more general mathematical formulation may have been convenient given its tractability. However, the parametric approach is convenient to perform simulations and for future econometric estimations.

Nevertheless, it was possible to show that this model, under special circumstances, observes similar behaviour and compatible results to similar models found in the literature. This implies that, despite its mathematical complications, the model does not depart substantially from the literature more than because of the elements introduced. It has been shown that some of the theoretical results already found in the literature remain valid within this framework.

It was confirmed in this framework that traders tend to buy futures, providing a natural counterpart to producers that tend to sell them. In our case, we could also observe this result under even less restrictive conditions than previous findings. Traders will take a short position only under very restrictive and unlikely conditions related to the values of the expected covariances. Therefore, traders in this model have a tendency or high probability to buy futures or go long.

However, this still cannot guarantee an unbiased futures price. Only under very restrictive conditions, similar to those found previously in the literature, may the bias disappear. Moreover, even with processors neutral to risk, and in contrast to the rest of the agents under similar risk attitudes, the futures price will still be biased since the short position taken by the rest of the agents may require a bias to attract an agent to take a long position that is not taken by the trader.

It was shown that, in the context presented here, and in contrast to the cases without second stages of production, the equilibrium price of both processed products depend not only on the structural parameters of the model but also on the price of the input. This is explained by the lack of futures markets for the processed products. On the other hand, the bias in the future price remains in the long run and, in contrast to the short run solution, the bias disappears when traders are neutral to risk.

The behaviour of the supply of export and domestic products has been analysed in depth. It can be shown that the decision to supply any of the markets is not only governed by all the intervening prices, but also that the parameters of the supply equations depend on the expected variances and the covariances between all prices. The decision on which markets to supply (the export or the domestic markets) is explained not only by differences in prices but also by differences in the variances and covariances on the prices. This is because the processor is concerned about the level but also the variability of the profits. Therefore, the trader may allocate output in one particular product if the other product exhibits a high variance in the price.

An additional interesting result is the effect of the futures price on the supply of exports and the domestic supply. The future market cannot only offset the volatility in the price of the input, but can also reduce part of the volatility in the effect that the price of both processed products has on the supply. In the absence of a bias in the future price, the effect of the future price on the supply offsets for part of the effect explained by the prices of both processed products.

Nevertheless, the precise sign and size of the coefficients of the export and domestic supply equations cannot be determined. Whilst we have identified precisely the definition of each parameter of the equations, the sign can only be determined by making assumptions about some values of their components. The parametric approach followed in this chapter, however, may allow the posterior econometric validation of the equations. Particularly of interest are the export and domestic supply equations. This is performed in the following stages of this research.

The model developed in this chapter might be enhanced in several ways. A more general approach in their formulation might address some of the analytical difficulties found. An approach like this might help to shed additional light on the behaviour of the export and domestic supply of agricultural commodities. However, a change like this might require a completely different validation strategy, where simulations exercises could be the most appropriate approach.

This theoretical model can accommodate the reality of the export and domestic supply of annual agricultural crops. However, non-storable agricultural or continuously produced commodities, and oil and mining commodities, also operate in similar contexts, where future markets are used for analogous purposes. Of course, the nature of the commodities involved would require further analysis on the type of technological risk they face (if they are affected by any). A study on the structure of these markets should be undertaken first. This is because, particularly in the case of oil and mining markets, the characterisation made in this research of traders intervening in these markets might not be appropriate given the observed vertical integration that exists in production and commercialisation in these markets.

CHAPTER TWO

SOME NOTES ON ECONOMIC HISTORY, AGRICULTURAL PRODUCTION AND FUTURES MARKETS IN ARGENTINA

Summary

Argentine economic history cannot be disassociated from the evolution of the agricultural sector. The country's current economic performance is still heavily influenced by the performance of this sector, which is crucial for Argentina. However, it is also an important international supplier of agricultural commodities – particularly wheat, maize and soybeans –and a vital player given its extensive export-oriented food industry. Therefore, the simple characterisation of agricultural production through farmers without the action of traders or processors does not represent the supply of these commodities in Argentina. Moreover, the existence of future markets for these products there cannot be ignored, given the importance they have in the production, hedging and commercialisation decisions.

Additionally, the performance of the Argentine economy over the last century has been characterised by high instability from a macroeconomic and institutional perspective. In the last 25 years, in particular, the economy has been subject to serious economic collapses. These latter explain the many unique economic decisions taken by agents that require special analytical treatment.

This chapter provides a context for the selection of both Argentina and the set of products analysed. Moreover, it serves as a reference point for some analytical approaches, treatments and results that will appear later in this research.

2.1. INTRODUCTION

The economic history of Argentina is associated with the performance of the agricultural sector. During colonial times, the country known today as Argentina was linked to the production and trade mainly of livestock products such as live animals, jerky, leather and animal fats. Any other economic activity was dedicated mainly to the provision of the domestic market or of goods or services for these export sectors.

This agricultural profile deepened after independence in the early decades of the 19th century and with the establishment, under Britain's leadership, of the new-world economic paradigm based on free trade. The international division of labour and the development of new transport and logistic technologies led to specialisation in the provision of agricultural products –particularly beef and wool – in Argentina by the mid-19th century.

Investment in transport notably extended the country's agricultural frontiers,[7] reaching areas that were previously considered unproductive given their location far from ports or populated areas. At the same time, these investments implied important changes in the location of the different agricultural activities, as some activities that were occupying "traditional" areas were displaced to new areas. The effective government occupation of the Patagonia region in the 1870s displaced wool production to the south, as Giberti (1986) suggests, liberating additional areas for grain production.

At the same time, the important European immigration that started in the 1870s implied a change in the pattern of specialisation. Whilst the majority of the immigration tended to be located in urban areas, an important number of European farmers introduced new agricultural techniques and, more importantly, much needed labour to agriculture and, in particular, grain production. Beef continued to be present, but cereals, in particular wheat and maize, were added to the export supply. By the end of the 19th and the beginning of the 20th centuries, Argentina was one of the largest exporters of food in the world.

[7] Development of the railway began in Argentina in 1857 and, by the beginning of 1914, the network extended almost 30,000 km. By the 1950s, it reached 47,000 km and was the fifth largest railway network in the world.

After the crisis of the 1930s, a policy-led industrialisation period dramatically changed the shape of the country. Important internal migration and the continued influx of foreign migrants notably increased the urban population at the expense of growth in rural areas. Whilst the agricultural frontier slowed down its advance, the relative fall in the rural population and the supply of agricultural labour exacerbated the extensive agricultural pattern of specialisation.

These changes, from the early stages, meant that agricultural production in Argentina took on a different format to the rest of the developing world. Important investments in mechanisation and storage facilities were necessary to offset scarce rural labour in the production of grain. This was associated with the particularly large average size of the farms, in comparison with other countries. Between 1988 and 2002, the average farm size increased from 421 to 587 hectares (INDEC 1998, 2002). In the EU and the US, the average farm size is 12 hectares and 180 hectares respectively (European Commission, 2013). Argentina's farm size is substantially higher than the average farm size in South America and only smaller than the farm size in Australia, as Eastwood, Lipton and Newell (2010) indicate.

Although the large farm size is a characteristic of agricultural production in Argentina, this is not translated into an unequal land distribution. Whilst the distribution of land in Argentina may seem unequal when measured by the Gini coefficient, it is below the average coefficient for South America, and only Colombia (with an average small farm size) has more equally distributed land, as shown by Eastwood, Lipton and Newell (2010). This suggests that the large average farm is not only the effect of just few extremely large farms; all the farm size distribution seems to be scaled up.

This large farm size (together with the low availability of agricultural labour) has led to a particularly capitalist organisation of agricultural production in Argentina, where farmers hire labour, buy inputs and contract services in the market. This contrasts to the production organisations in other developing countries, where other types of economic relationship, such as peasantry, tend to have a more important role. At the same time, as farmers require the participation of external investors (sowing pools, for example) as well as financial providers, the need for transparent markets is manifest, since these agents would prefer not to participate or to finance activities in contexts of limited and private information.

These large agricultural enterprises and investors were in a position to finance the existence of organised trading markets and to demand the existence of instruments to hedge against fluctuations in the prices. Since the early stages in the development of grain production in Argentina, futures markets were available and provided this coverage, suggesting that they offered a necessary service for farmers and other agents. Moreover, they were also in a position to finance these markets, given their capitalist organisation and volume of production.

With a low rural population, important focalised demand centres[8] and large distances to cover, farmers are not in a position to trade their own output. They need to rely on traders and intermediaries to market their output. At the same time, the development of the vegetable oil industry generated the appearance of focalised and limited but key demanders. These agents were also heavily engaged in the operations in futures.

However, the operation of futures markets in Argentina was not problem-free. Several developments in economic policy during the 20th century substantially complicated their operation and, for many years, they were virtually at zero. Moreover, the introduction of Import Substitution Industrialisation (ISI) strategies affected the production and trade of agricultural products. At the same time, it reinforced the self-sufficient character and anti-Government intervention of the sector. The recent tension (that started in 2008) between farmers and the Argentine Government cannot be seen in isolation from the historical distrust between farmers and the policy-makers who have ruled in Argentina during many decades of the last 70 years.

Three crops are extremely important in the production and trade of Argentina: wheat, maize and soybeans. Each presents different characteristics in terms of their use and they constitute relevant and interesting cases to study. The purpose of this chapter is to motivate discussion of the whole research project and to provide a background to the understanding and justification of some of the approaches used and treatments applied during the study. In the first part of this chapter, I discuss and justify the importance of agriculture for Argentina and the relevance of the three crops analysed through studying the importance of these crops for Argentina and the importance of the country as a

[8] Just three cities of Argentina account for almost 50% of the total Argentine population.

supplier of these crops in the world market. I also highlight the importance of traders/processors as agents in the commercialisation of the commodities identified.

In the second part, I present in more detail the production and marketing characteristics of each of these crops, in particular with respect to the deterministic patterns of their seasonality. This helps to identify the main elements that explain and characterise each of the seasons in the commodities commercialisation.

In the third part, I focus on characterising the economic and agricultural policy of the last 25 years, highlighting the important volatility in economic policies and, consequently, their effects on economic performance. This provides a background for some of the modelling strategies to be discussed later in this research.

In the fourth part, I provide a short history of the operation of futures markets in Argentina. The idea behind this is to present their relevance and to associate performance in these markets with the economic and institutional background in operation in Argentina over the last 100 years.

2.2. AGRICULTURE IN ARGENTINA

The importance of agriculture in Argentina is difficult to see in the standard economic figures, as they tend to understate its importance. The value added generated in this sector represented between 5% and 8% of GDP between 1993 and 2006[9]. This is lower than the average of South America and substantially lower than the average of the Developing Countries where this figure tends to be more than 20%.

In terms of employment, the agriculture sector is not the main employer in Argentina. According to the Food and Agricultural Organisation (FAO)[10], in 2010, less than 7.3% of its population lives in rural areas and 7.6% of the economically active population is employed in agriculture. This suggests that the importance of this sector in the Argentine economy cannot be seen through the participation in total employment. This contrasts with the 14% share in the economically active population in the rest of South America and with the 40.3% in the World according to FAO as well. Although agriculture is the

[9] World Bank - World Development Indicators.
[10] FAO – FaoStats.

main and practically only source of income and employment in rural areas, the importance in terms of generating employment for the economy as a whole is particularly small. This suggests that it is not through its contribution to total employment that we can identify its importance.

However, the agriculture sector is extremely important in terms its contribution to the finance of imports. Without considering the food industry, agricultural exports in Argentina represented 17.5% of total exports in 2007[11]. The agri-food sector (agriculture + food industry) represented 47% of Argentine exports in the same year. The availability of important extensions of arable land combined with a relatively low population generate an important agricultural export surplus that is behind these figures.

The importance of the agricultural manufactures in Argentine exports implies that not only Argentina exports agricultural commodities but also products that have been subject to industrial transformation. In fact, the most important exported products in Argentina are soybeans cakes and oils. Moreover, the agricultural and the food exports exhibit important product diversification. Among the top agri-food exports, it is possible to find cereals, oilseeds, fruits, vegetables, meats, dairy, processed food and the production of oils and oilseed cakes already mentioned.

Nevertheless, the production and export of cereals and oilseeds are notable. Cereals and oilseeds represent, respectively, nearly 8% and 6% of total Argentine exports in 2007. Among cereals, wheat and maize are clearly the most important products. They represent almost 8% of total Argentine exports as we can see from Table 2.1.

In the case of soybeans, whilst exports of sunflowers have been historically important, soybeans are definitely the most important exported oilseed. They represented 5.6% of total Argentine exports during the period 2007-12 and accounted for around 88% of total exports of oilseeds during the period.

[11] INDEC – Intercambio Comercial Argentino.

Table 2.1. Wheat, Maize and Soybean Exports (average 2007-12)

	Value of exports (in thousands of USD)	Share in total exports
Wheat	1,987,945	2.9%
Maize	3,317,263	4.8%
Soybeans	3,888,144	5.6%
Total exports	9,193,352	13.30%

Source: INDEC – Intercambio Comercial Argentino

It is important to highlight that these commodities are not only important for Argentina in terms of its share in its trade; Argentina is an important producer and supplier of these products. This suggests that these products are important for Argentina but at the same time, Argentina is an important supplier of these products. Table 2.2 presents the different measures of Argentina as supplier of these three commodities between 2008 and 2011.

Table 2.2. Share of Argentina in world's area harvested, production and exports of wheat, maize and soybeans.

	Total area harvested (2008-2011)	Production (2008-2011)	Exports (2007-2010)
Wheat	1.9%	1.8%	5.0%
Maize	1.9%	2.4%	13.4%
Soybeans	17.5%	18.2%	12.6%

Source: FAO – FAOSTAT – Production and Commodity Balances.

Whilst in the case of wheat and maize its share in the area harvested and in production might not be seen as important as many countries produce these cereals to supply their own domestic markets, the importance in world trade is manifested. Argentina accounted for 5% and more than 13% of the world trade in wheat and maize, respectively. In the case of maize, Argentina is the second most important exporter behind the United States.

On the other hand, in the case of soybeans, its share in world supply is even higher. More than 18% of world production and more than 12% of world trade in soybeans are originated in Argentina. Argentina is the third largest producer and exporter of soybeans behind the United States and Brazil.

Nevertheless, the magnitude of the share of these commodities in Argentinean trade, as well as the importance of Argentina as supplier of these products, suggest the existence of minor domestic markets. Whilst the export market tends to dominate the domestic supply of wheat and maize, the situation is the opposite in the case of soybeans where the domestic market clearly dominates. Nevertheless, even in the case of wheat, the domestic supply tends to account for around 30% of the production of this product.

There are, however, differences in the way these products are used domestically. These differences are associated not only with the characteristics of the products but also by their main demand or use of them. In this way, soybeans and maize is used as feed, whilst wheat has a predominant use as food. Therefore, the export availability is affected, among other factors, by the size of the local population (in the case of those products that are used a food) and by the existence of other activities that may demand these products, such as livestock farming.

For each of the products considered, Table 2.3 indicates the volumes of exports and domestic supply during 2008. Focusing on wheat, the majority of the output tends to be exported[12], although nearly one third of output is domestically used mainly by mills in the production of flour for domestic consumption or used in the food industry. Given its special conservation requirements, trade in wheat flour tends to be particularly limited to cross-border trade. The relatively high ratio output to population of Argentina allows having a large export surplus of wheat.

Table 2.3. Distribution of production 2008 (in thousands of metric tonnes)

	Export supply	Domestic supply
Wheat	10,238	4,175
Maize	15,454	6,176
Soybeans	11,734	36,992

Source: FAOStat Food balance sheets

Note: Domestic Supply includes feed, seed, food manufacture, other uses and food.

In the case of maize, the majority is exported and the domestic supply is mainly used for feeding purposes. This contrasts to other countries where maize has also a food use. For example, in Mexico (with a comparable production although marginal exports) the use of maize as food surpasses the use as feed; the opposite of Argentina where the use of food is almost marginal and the main domestic use is feed. This implies that consumer

[12] Although production can be stored and not consumed, we assume that this is part of the domestic supply.

preferences may be behind these figures, as well as the determination of a particular use for these products. On the other hand, increases in intensive forms of livestock farming as well as the increase in the production of poultry and pork has pushed up domestic demand in Argentina. Although domestic uses seem to be on the rise, in terms of share, exports tend to dominate the distribution of production of maize.

Whilst wheat and maize can be seen as "traditional" exports of Argentina, soybeans are a relatively new product. In fact, there are no records of soybean production before the 1960s but its production has been steadily growing since that time to become, by the end of the 1990s, the most important grain produced in Argentina.

In the case of soybeans, the domestic supply is substantially more important, with domestic supply larger than export supply by a factor of 3. The main domestic, and almost unique, use in Argentina is the production of soybean oil and soybean meals, Argentina being the most important exporter of these products. In fact, the "soybean complex" (soybeans, soybean oil and soybean meals) represented, on average, nearly 23% of total Argentine exports of goods in 2008[13].

A final element to consider, given the importance this research has put on them, is the role of traders and processors in the trade and domestic supply of these products. As we mentioned, only extremely large farmers or producers are in a position to export or trade in the domestic market without the intervention of traders. Moreover, as these grains tend to be intermediate products in the production process of other food items, large processors such as mills or vegetable oil producing companies play a decisive role in the determination of spot and future prices.

Table 2.4 tries to assert the importance of these types of company by identifying their share in the total Argentine exports. It can be seen that among the top 20 Argentine exporter firms, 12 of them are either traders or processors of grains. Moreover, the exports of these companies represented more than 30% of the Argentine exports. Although it is difficult to assess how much they represent of the exports of wheat, soybeans and maize, given the high share of these products in the Argentine exports, it is clear that their share in the exports of these products could not be lower than their share in total exports.

[13] INDEC – Intercambio Comercial Argentino.

Table 2.4. Top 20 Argentine exporter firms 2011

Firm	Sector	USD millions	Share
Minera Alumbrera	Mining	4,132.0	5.7
Cargill	Trader/Processor	3,737.0	5.1
Pan American Energy	Oil/Gas	3,608.0	5.0
Bunge	Trader/Processor	3,517.0	4.8
LDC	Trader/Processor	3,060.0	4.2
Aceitera General Deheza	Trader/Processor	1,931.0	2.7
Volkswagen	Automobiles	1,616.0	2.2
ADM	Trader/Processor	1,602.0	2.2
Vicentin	Trader/Processor	1,554.0	2.1
Noble Argentina	Trader/Processor	1,412.0	1.9
Alfred Toepfer	Trader/Processor	1,376.0	1.9
Molinos Rio de la Plata	Trader/Processor	1,356.0	1.9
Nidera	Trader/Processor	1,162.0	1.6
Transportadora de Gas del Sur	Gas	942.0	1.3
Ford	Automobiles	873.7	1.2
Oleaginosa Moreno	Trader/Processor	873.0	1.2
Asoc. De Coop. Arg	Trader/Processor	764.0	1.1
YPF	Oil/Gas	655.0	0.9
Medanito	Oil/Gas	632.0	0.9
Siderca	Steel	621.0	0.9
TOTAL Top 20		35,423.7	48.7
TOTAL Trader/Processors		22,344.0	30.7

Source: Own elaboration based on data from the Asociación de Importadores y Exportadores de la Republica Argentina (AIERA)

Their shares in domestic supply are harder to assess as the data are substantially scarcer. However, these companies tend to operate in the domestic and export market. Many of these companies could be buyers in the domestic market and sellers in the export market. For example, vegetable oil producers tend to be the main buyers of soybeans but also the exporters of soybean oil. As a reference, the top 20 facilities used in the production of vegetable oil and fats (with many facilities owned by a single company) represented

94.5% of the gross value of output in 2003[14]. This indicates the high degree of concentration in the production of these products and, consequently, in the use of oilseeds.

The high share in the total and in the agricultural exports that a reduced number of traders and processors present implies additional evidence to the strategy followed of including traders and/or processors in the development of the model developed in Chapter 1. Moreover, given the evidence presented, it remains difficult to conceive their exclusion in the analysis.

The description made so far allows the assertion of the relevance of the problem under analysis. The focus on the study of the marketing and trading aspects of agricultural commodities in Argentina is of extreme importance given the weight that the agricultural sector has in the Argentine economy. On the other hand, the three commodities identified are important in terms of their contribution to Argentine exports and constitute relevant and interesting cases of analysis.

On the other hand, it can be seen that the characterisation made of the commodities markets in Argentina, emphasizing the role of traders and processors, given the importance of them in the commercialisation of commodities, is adequate and in line with the reality. Farmers and producers in Argentina are not the direct exporters or suppliers of the domestic markets and the role of traders and processors cannot be downplayed.

2.3. PRODUCTION TIMING AND LOCATION

Wheat, maize and soybeans are repeated or annual crops implying that the production of the grain is associated with the life cycle of the plant. This contrasts to permanent crops such as fruits where a single plant produces several times during its lifetime that may extend over several years. In the case of annual crops, there are clearly identifiable moments such as the sowing or implantation and the harvest times. Depending on the crop and the location, among other factors, these moments may be unique or multiple, presenting the possibility of unique or multiple cycles in the same year. For example, given the latitude and climate conditions, in the United States or China is possible to have

[14] INDEC – Censo Nacional Economico 2004/05 – *Cuadro 4. Concentración en la ocupación, la producción y en la producción según rama de actividad.*

two planting and harvest seasons of wheat in a single year, creating two distinct production cycles.

The two production moments (sowing and harvesting), in terms of its location in the calendar, are affected by multiple factors and each product may have particular requirements. Some crops require higher sun radiation at the end of the production cycle, for example, wheat and other so-called winter grains such as barley, oats and rye. The water input requirements may be different in specific times of the life of the plant, implying that the location of the production moments is affected as well by the rainfall regime. At the same time, climate conditions may affect the cultures associated with each production moments. For example, snow or excessive rain in certain key moments may make complicate the work of heavy machinery. These climate factors will affect the "deterministic" production times. However, weather, among other factors, may introduce instability to these patterns.

This suggests that these deterministic conditions will tend to vary enormously, in particular, with the size of the production area under consideration. The larger the area the more diverse the climate will be, especially because the area will be diverse in terms of altitude and the different intensity and time availability of the sun radiation. This is particularly the case of "long" areas[15]. This implies that the optimum sowing or harvest times may differ not only between crops but also for the same crops if the area or country under consideration is particularly big.

For the reason explained, it remains particularly complicated to identify sowing and harvest seasons in a country of the size of Argentina. There are approximately 16 degrees of latitude between the north and south extremes of the productive areas[16], implying that the number of hours of daylight and, consequently, sun radiation may differ within the area, with the southern parts exhibiting more variability during the year. Moreover, given the extension, it is frequent that sowing and harvesting periods may overlap as in the north harvest may begin when sowing is finishing in the south.

[15] We define areas as "long" to those that are spread over several latitude parallels.
[16] It remains particularly difficult to define the extremes of the arable land, but it can be identified between parallels 23 and 39 below the Equator, roughly equivalent to the distance between London and Gibraltar.

Moreover, considering also the width of the productive area (around 800km), the climate is also diverse with the areas closer to the sea presenting different rain patterns than the areas located in the interior. This should be added to the variability that might be observed between north and south, with lower temperatures in the south and different rain patterns. All this implies that seasons tend to be particularly wide in terms of length of time, with sowing and harvest seasons occupying more than one month.

Particularly in the case of Argentina, production of wheat, maize and soybeans is mainly located in the top part of the country with the Patagonian region supplying marginal volumes of the commodities. In general, the Region Pampeana (roughly between parallels 30 and 38) is historically the most important area given its meteorological, edaphic and pedological conditions. However, in the last two decades, given the advances in terms of direct sowing, the introduction of genetically modified crops and the improvements of transport infrastructure, among other factors, marginal or traditionally livestock farms areas in the Northeast and Northwest of the country have increased their share in the area sown and harvested in the three products.

Consequently, with this increase in the productive areas, the amplitude of the implantation and harvest seasons has widened, making the identification of seasons particularly complicated. Nevertheless, it is possible to identify broadly and anticipating the changing nature of these borders, the extremes of the different periods where the different productive moments take place.

An additional element that exacerbates the seasonal pattern, especially in the case of exports, is the availability of vessels. Although the storage capability has increased, foreign trade operations outside the traditional harvest time imply higher transport costs as other areas in the world are also harvesting and, therefore, there is increased demand on these services. In contrast, during harvest time, the exporters in Argentina would be the only ones that would be demanding these services. Therefore, as we will see in the following chapter, the seasonal pattern of the exported products is influenced by this aspect.

Wheat is a cereal sown at the beginning of winter and harvested around the beginning of summer. In this sense, implantation might begin as early as the last week of May in the north of the country and by the last week of July, around 90% of the intended sown area would be covered[17]. Wheat harvest season tends to start by mid-November in the north of the country and might extend until mid-January in the south. The peak in the activity, measured by the volume of wheat harvested, is observed during December.

Maize is generally sown in spring (although it is possible that by the end of December, nearly 10% of the area remains to be implanted), and it is harvested between the end of summer and the beginning of winter depending on the region. Harvest may begin by the end of February and may extend until the end of June, although the peak of activity is generally observed during April.

In contrast to the other two cereals, soybean presents two sowing seasons. The first season begins by mid-October and the second season is implanted immediately after wheat has been harvested. The adoption of direct sowing has allowed this practice. This has led to the identification of two types of soybean associated with the time it has been implanted: First and Second soybean. Second soybean is implanted generally between mid-December and mid-January presenting a particularly short span of time. Although soybean is planted in two different periods, there is no distinction in the time of the harvest. Harvest occurs mainly in autumn, starting in March and extending until the end of May. This implies that, although soybeans may present two different sowing times, it is still a one season crop as their harvest (and eventual supply) is the same regardless of when was planted.

The analysis made has allowed us to identify the main moments in the production and supply of the three commodities identified. In the case of wheat, exports and domestic supply tend to be higher in December and January; whilst for maize and soybeans autumn is the time with the highest activity. However, given the important latitude in which this production is located, this period may present some variability.

2.3.1. Technological innovations

So far, we have discussed the elements that characterise the deterministic seasonality present in the commercialisation of these three commodities. We have also discussed about some stochastic elements that might affect seasons and that they would be

[17] Ministerio de Agricultura – Estimaciones Agricolas semanales – weekly reports.

exacerbated in the case of Argentina, given its size. We would also like to present some technological innovations that have occurred in the last 20 years that might have introduced, if they had a cumulative effect, some stochastic seasonality in the series.

The technological innovations and institutional changes constitute two important factors that might have a cumulative effect that may introduce stochastic seasonality. Although they might be, in their conception, one-off events that will not be repeated in the future, their effect may spread over many periods. This might be the case if the adoption or implementation of new technology or institutional changes takes time.

The authorisation for the implantation of glyphosate-resistant soybeans in 1997 in Argentina had important effects on the quantity of hectares implanted with this crop. This, consequently, affected the area implanted with the other crops; but also it has affected the total agricultural land. However, as the implementation of this variety has not been immediate, the effects on the total area implanted and on the total production have been cumulated over many periods. Producers required time to become used to the new variety and to include the rest of the elements necessary for its implantation. Therefore, this innovation, through its cumulative effect, could have affected the seasonal pattern of the series; moreover, it could affect its trend as well.

Additionally, and in conjunction with the authorisation of genetically modified soybeans, the spread in the use of direct sowing also had important effects. Whilst in the campaign 1993/94 the area under direct sowing in soybeans, wheat, maize and sunflower where slightly under 2 million hectares, by the campaign 2004/05, this area had grown tenfold, as can be seen in Table 2.5.

The double effect, genetically modified crops and direct sowing, had important effects in the agricultural sector. There have been changes in the composition of crops (with soybeans reaching 50% of the area implanted) and there has been an increase in the total area implanted that went from 20.5 million hectares in 1993/94 to 33.3 million in 2007/08. These changes cannot be attributed exclusively to the rise in international prices of agricultural commodities observed in the last 10 years, given that they started to operate before. By 2004/05, the area had already increased by more than 8.3 million hectares, and soybeans already represented half the area implanted. Additionally, as we have seen above, direct sowing facilitated the introduction of a short-cycled soybean immediately after the harvest of wheat.

Table 2.5. Annual crops implanted area (in thousands of hectares)

	1993/94		2004/05		2007/08	
	Area	Share	Area	Share	Area	Share
Wheat	4,910.0	23.9	6,039.9	20.9	5,947.8	17.8
Maize	2,781.0	13.5	2,988.4	10.4	4,239.7	12.7
Soybeans	5,817.5	28.3	14,526.6	50.4	16,603.5	49.8
Sunflower	2,205.8	10.7	1,848.0	6.4	2,612.6	7.8
Oats	1,971.4	9.6	1,344.0	4.7	1,112.9	3.3
Cotton	503.6	2.4	266.4	0.9	310.4	0.9
Other crops	2,386.6	11.6	1,827.4	6.3	2,537.9	7.6
Total	*20,575.9*	*100.0*	*28,840.6*	*100.0*	*33,364.9*	*100.0*
Direct sowing area	*1,900*	*12.1*	*19,800*	*77.9*	*25,500*	*86.7*

Source: Ministerio de Agricultura, Ganadería y Pesca – Sistema Integrado de Información Agropecuaria and Asociación Argentina de Productores en Siembra Directa.

The important point to highlight in these innovations is that, although it might be possible to identify the precise moment when they started to operate, their effects have been cumulative rather than instantaneous. This implies that these changes, whilst one-off, might have introduced stochastic elements in the seasonal pattern as producers adopted these innovations. Seasons might have been affected as new areas are implanted with these crops and/or, in the case of wheat, as the harvest time is speeded up to implant short-cycled soybeans. This must be distinguished from the standard structural break observed in a single period.

The possibility that these innovations might have introduced stochastic elements in the seasons is what would motivate the seasonal unit root analysis and posterior econometric estimation of these series.

2.4. RECENT ECONOMIC POLICY

The performance of the Argentine economy during the 1980s was clearly unsatisfactory. The economy did not grow and inflation accelerated dramatically during the period. Average consumer price index annual variation between 1980 and 1990 was 1,175%;

with a maximum observed in March of 1990 of 20,263%[18]. Between 1980 and 1990, two new currencies were introduced with the currency denomination losing seven zeroes in the process[19]. The average annual real GDP growth rate in the same period was -0.88%. Just between January 1989 and February 1991, the price of the US dollar in local currency terms had multiplied by a factor of 582.

Although the 1980s were part of a longer period of economic stagnation, the accumulation of bad results in terms of economic policy led to the exacerbation of the problems observed and to the eventual economic collapse at the end of the decade. Different stabilisation policies applied during the previous 30 years could not solve the structural problems that the Argentine economy had observed. Each policy applied created new disequilibria that feedback an on-going unstable process.

On the contrary, by the end of the 1980s it was perceived that the roots of the economic problems were more related to the structure of the Argentine economy and, in particular, with the development strategy followed in the previous 40 years. This view was coincidental with the approach sketched in the so-called Washington Consensus (Williamson, 1990) at the end of the 1980s suggested that stable growth and inflation can only be dominated by abandoning the old development strategy and embracing deregulation, more free market-oriented economies, less government intervention and smaller public sectors.

In terms of policies implemented in Argentina, and in terms of the relevance of the problems under study, the reforms must be classified in two groups: those general policies applied across the whole economy and those policies with a primary objective of the agricultural sector.

In March 1991, a currency board was established. This fixed the price of the US dollar but more importantly, implied the virtual elimination of the possibility in the use of monetary policy. The *Ley de Convertibilidad* established that the Central Bank made the strong commitment of keeping a monetary base of no less than the value of the US dollar denominated international reserves. In fact, the original law established that for the

[18] Same month previous year variation of the Índice de Precios al Consumidor (IPC) between December 1980 and December 1990.

[19] 10,000 Pesos Ley 18.188 were equivalent to one new Peso Argentino in 1983. 1,000 Pesos Argentinos were equivalent to one Austral in 1985.

purposes of the currency board, a limited share of US dollar denominated Government bonds would be considered as international reserves. The rest of the international reserves should be constituted in deposits of US dollars in the Federal Reserve of the United States. This meant that any expansion in the monetary base was constrained to the availability of additional physical international reserves.

In terms of addressing inflation, the results were more than satisfactory. CPI variation between March 1991 and March 1992 was 30% and by 1994, the annual variation of the CPI was just around 3%. CPI variation during the rest of the 1990s was even lower and below the inflation in the United States. Fix nominal exchange rate and higher international inflation generated an important appreciation in the real exchange rate.

An important part of the stabilisation package was the authorisation to operate, make contracts, make bank deposits and contract loans denominated in US dollars. This was the legal recognition of a long-term practice in virtue of the lack of stability in the domestic currency and its failure in meeting its functions: store of value, measure of value and standard of deferred payment. Of course, given these problems, the currency was no longer a medium of exchange. Therefore, even after the success of the Ley de Convertibilidad, the US dollar continued to circulate alongside the peso. This means that, although inflation had fallen dramatically, the peso did not assume all the standard functions of money. The US dollar continued to be the denominator of contracts and, generally, any business transaction. For example, not only transactions with foreign residents were denominated in US dollars; bank deposits, debts, mortgages and other domestic transactions did not make use of the peso to denominate their value. This, as we will see, will have implications in the modelling stages as it complicates the use of peso denominated variables.

The stabilisation package included a vast programme of privatisations of public enterprises and the deregulation of markets. This was accompanied by the acceleration in the trade liberalisation schedule. Tariff protection was decreased substantially and the Mercosur customs union was established. Protection in “inefficient” sectors was reduced with the objective of achieving efficiency by specialising trade in those sectors with clear comparative advantages. The agricultural sector was among those sectors[20].

At the same time, different forms of direct subsidies to sectors, particularly specific sector funds, and other ways of promotion were cut or eliminated, affecting particularly non-traditional sectors. In addition, the privatisation and de-regulation efforts implied the end of the subsidised provision of different goods and services via the existence of different cross-subsidies from the taxpayer to sectors.

As mentioned previously, there were also specific policies addressed to the agricultural sector or that at least had agriculture as its main affected sector. Government intervention in the trade of grains was eliminated in 1991 when the Junta Nacional de Granos was dissolved. Privatisation and deregulation of grain elevators had the objective of reducing the cost of trade particularly in the products that use these services.

Moreover, export taxes were eliminated in 1991 immediately affecting the distribution of output between export and domestic markets, and increasing the domestic price by the elimination of the wedge with the export price. However, an export tax of 3.5% was kept in the exports of oilseeds with the objective of reducing the price paid by the oil crushing industry, and has proved to be central in the expansion and consolidation of Argentina as an important exporter of oils and meals.

In general, the 1990s saw a more amicable approach and attitude to the agricultural sector and proved to be key, albeit with exogenous factors, in the performance of the crops under study and in the development and use of futures markets. Not only specific policies but also general policies worked in favour of these aspects.

Different external but also internal factors led to a progressive deterioration of the economic situation. It is beyond the scope of this study to analyse its causes. Calvo, Izquierdo and Talvi (2003) suggest that Argentina was vulnerable to the sudden stop in capital flows of the Russian crisis in 1998. Although specific elements explain the differences in the effects between Argentina and other emerging economies, the origin of the crisis is exogenous. De la Torre et al (2003) provide an endogenous explanation related to the intrinsic nature of the currency board. However, generally it is accepted that

[20] The role of Mercosur in creating trade is debatable, as there is controversy about the role of FTAs in the creation of trade and in particular the case of Mercosur.

whilst the currency board established in 1991 has paid all its benefits in terms of inflation control, it was generating serious costs in terms of unemployment. During the decade, a progressive process of appreciation of the real exchange rate was taking its toll in the economic activity in the last years of the decade, and there was no possibility of nominal exchange management under the Ley de Convertibilidad.

At the same time, commodity prices by the end of the 1990s and beginning of the 2000s were particularly low. The average IMF food price index in 1999-2001 was 25% lower than in 1995-1997[21]. By the end of the 1990s, the Argentine economy had entered into a deep recession. Real GDP growth was -3.4%, -0.8% and -4.4% in 1999, 2000 and 2001 respectively.

The impossibility of performing monetary policy was combined with a severely constrained fiscal policy due to high levels of debt and the burden of its services. This made the government unable to make policy under the corset of the currency board. In the last 10 days of 2001, Argentina had four presidents and, in terms of economic policy, abandoned the currency board, devaluated its currency and suspended the payment of its external debt obligations.

The collapse of the Argentine economy in 2002 had immediate negative effects. First quarter growth in 2002 compared to the same quarter in the previous year was -16.3% and although, measured in deseasonalised terms, economic activity started to recover by the third quarter of 2002; real GDP growth in 2002 was -10.9%.

Facing the need of creating some subsidy for the thousands of families without income[22] and without access to new debt given the unilateral suspension in the payments of the external debt, the Government looked for alternative sources of income. Eyes were put immediately on the extraordinary windfall gains that exporters were obtaining because of the devaluation of the peso. At the same time, there was an objective that the price effects of the devaluation were not transmitted, at least immediately, to the domestic prices. Consequently, export taxes were re-introduced in March 2002. Tax rates of 20% for cereals and 30% for oilseeds were established[23]. This implies the reintroduction of the

[21] IMF – Food Price Index. Base 1995=100
[22] Unemployment and underemployment were 21.5% and 18.6%, respectively, in May 2002.
[23] These rates were put on top of the existing rates. In this way, soybeans exports faced a rate of 33.5%.

wedge between the export and the domestic price with the size of the wedge closely similar to the rate.

The small exchange rate pass-through on prices of the devaluation generated some real effects[24]. Nominal exports grew by 81% between 2002 and 2005, and the economy bounced with real GDP growing no less than 8% in each of the years between 2003 and 2006.

In 2006, two related phenomena began to be clear. On one side, international prices for commodities began a sustained growth. Measured by the index of export prices, average export prices were 112% higher in 2008 than in 2002. In terms of the products under study, prices for wheat, maize and soybeans were 119%, 124% and 139% respectively between the same two years.

On the other hand, with the economy operating at higher capacity and with the effect of international prices in commodities, domestic prices began a new ascension path. Unfortunately, official consumer price index information is not considered reliable since the beginning of 2007 as the Government started to manipulate price statistics as a way of hiding the real magnitude of the inflation problem. Whilst the official consumer price index reveals an average growth in the index of 10% between 2007 and 2012, private estimations revealed substantially higher inflation. This lack of reliability in the price statistics is recognised by the IMF as The Economist (2012) publishes; the press has shown in The Guardian (2012) and has been study in Cavallo (2013)

At the same time, additional pressure on public finances, and the impossibility of accessing international markets made the Government look again into the even higher rents of the agricultural sector because of higher international prices. Therefore, the Government intended to apply a variable export tax on soybeans whose rate evolved positively with the international price. This implied that immediate increase in the export tax rate from 30% to 44%.

The so-called "125", given the number of the Ministerial Resolution of 11th March 2008, was heavily opposed by farmers. This started a conflict, "Conflicto del Campo", which included demonstrations, pickets, riots and several days of complete paralysation of grain

[24] The small pass-through is the result of an accumulated increase in the nominal exchange rate of 190%, and an increase in domestic prices of just 53% between 2002 and 2005.

commercialisation. The conflict was resolved by the decision of the National Congress of rejecting the Ministerial Resolution in July. However, in these four months the commercialisation of grains was altered completely.

The relationship between farmers and the Government was never the same after the conflict. The general distrust between some policy-makers (particularly those coming from the Peronist extraction) and farmers reappeared after the conflict. Increased intervention on the commercialisation of grains such as the pre-authorisation of exports or export quotas and the general intervention in the economy have tensioned the relationship. Nevertheless, the latest events of the economic and agricultural policy are outside the scope of this research and are not reflected in the data used. This analysis of the recent economic history indicates an important volatility in the main economic variables. Moreover, it suggests that the economic and general policy had also observed important instability and a conflictive relationship between the Government and the agricultural sector.

Moreover, although macroeconomic stability had improved in comparison with the decades before 1990, it had not been enough to eliminate the double currency standard of the Argentine economy. Given the instability experienced, agents tend to use the US dollar as a measure of value; and only the variation in the prices denominated in this currency are expected to have real effects. Recent attempts to make agents "to think in pesos" Moffet (2012) have generated serious falls in the activity of many sectors such as real estate (The Economist, 2013), and have been ineffective in their objective as Cachanosky (2013) suggests.

This suggests that any estimation that tries to capture the effect of nominal variables in real variables in Argentina need to consider this fact and look for alternative specifications. It is expected that only extreme changes in the nominal exchange rate, as the one that occurred in 2001, for example, would have important real effects in Argentina. Moreover, any evaluation of projects of investments that require months to mature, like agriculture production, cannot be formulated in an extremely volatile currency.

2.5. EVOLUTION OF THE FUTURES MARKETS IN ARGENTINA

Organised markets for the commercialisation of grains have existed in Argentina since 1854. However, 1907 is the year that is normally regarded as the beginning of the operation in futures markets with the creation of the Mercado a Termino de Buenos Aires (MATBA). It is important to highlight that, in Argentina, there are two main organised futures markets: the MATBA and the Mercado a Termino de Rosario (now known as Rofex), created in 1909.

However, the existence and performance of futures markets in Argentina could not escape the changes in economic and agriculture policy and their effects in the Argentine economy. Moreover, futures markets operations have been permeable to the changes in economic and political conception of the different Governments that introduced the different economic policies. This means that their performance and development have been closely associated to how the market, as a coordinator of economic resources, was seen by the different policy makers over the last hundred years.

Measured by volume operated, and following Olivo (2010), it is possible to distinguish four stages in the history of futures markets. However, as we will see, there is a close connection between the stages used to separate the evolution of future markets and the stages in what could be separated the evolution of the Argentine economy and the economic policy in the last century.

2.5.1. First stage: 1907-1930

This period observed important dynamism in the operation of these markets. Operations were growing and in every single year between 1919 and 1926, operations were larger than the output of grains. Whilst the volume operated was around 18 million tonnes, grains output in Argentina was between 10 and 16 million tonnes. This is important to highlight since this phenomenon has not occurred again and revealed the importance of these markets during that period.

At the same time, the institutional framework and general market friendly policies followed during that period were particularly beneficial to the development of these markets. This was accompanied by the important growth in grain production that nearly doubled during these two decades. It is important to highlight that, given soybean was non-existent in Argentina, futures operations did not existed for this oilseed in this period.

It is important to highlight that during this period, Argentina observed important economic growth. Between 1900 and 1929, real GDP in Argentina grew by 275%. The GDP per capita in Argentina in 1912 was higher than the GDP per capita of France and Germany (Bolt and van Zanden, 2013). This coincides with a period of massive immigration from Europe. As we will see, the performance of the futures markets in Argentina will follow very closely the general economic performance.

2.5.2. Second stage: 1931-1940

The second stage is characterised by the effects of the world economic crisis that started in 1929 and the policies implemented to address them. Increasing controls and intervention on foreign trade, the establishment of capital controls and different interventions in the exchange markets, and the eventual creation of the Central Bank in 1935 after abandoning the Gold Standard in 1929, were the most important general monetary and economic measures that affected the operation of futures markets.

Also given the importance of grain exports in the total Argentine trade at that time[25], policies addressed primarily to the agricultural sector were used also to achieve other general economic objectives. The creation of the Junta Reguladora de Granos (grains board) in 1933 (that established minimum prices for wheat and maize) and the intervention in providers of services to the agricultural sectors, were among the most prominent sector specific policies but with a far more general scope. In 1933 the operation of grains elevators were put under the control of the National Direction of Grains Elevators.

These increasing Government interventions had their effects on the operations of the futures markets. In a context of stagnant grains production, operations in futures decreased. Volumes operated were substantially below output and under the volumes operated in the previous decade. Although the fall in operations tended to reflect the path of introduction of newer policies and interventions, the operations tended to be around

[25] Two-thirds of total Argentine exports were explained by wheat, maize and linseed during the decade of the 1920s as seen in (Gerchunoff and Llach, 1998)

half those seen in the previous period. This suggests that even though government intervention was increasing and affecting operations, it still allowed the existence of significant futures markets.

It is also important to highlight that this period coincides with the start of the long period of economic stagnation that Argentina faced after the crisis of the 1930s. This economic stagnation has been studied in depth by Diaz Alejandro (1970), Street (1974) and Gerchunoff and Llach (1998) among others. This reinforces the association between the performance of the futures markets in Argentina and its economic performance.

2.5.3. Third stage: 1941-1991

This stage begins with an even more important Government intervention combined with an import substitution industrialisation (ISI) development strategy. The creation in 1946 of the *Instituto Argentino para el Intercambio* (IAPI) implied the complete intervention in foreign trade, the Government being the only buyer and seller of grains to the domestic and foreign markets.

The ISI policy followed from 1946 and the important expansion of the Government and public consumption had its macroeconomic effect with the appearance of inflation by the end of the 1940s. It is within this framework that the so-called *stop-and-go* (Braun and Joy, 1968) cyclical macroeconomic process of expansions-devaluations-recessions was verified. Furthermore, a process of almost systematic devaluations that took its toll on the price level began in this period and extended until the 1990s.

The result of this major government intervention, that was an expression of a general distrust in the capability of the market to allocate resources, was the almost complete disappearance of the operation of futures in Argentina. Moreover, the additional macroeconomic instability and general political anti-agricultural bias had its effects on grain production and on the operation in futures. Between 1946 and 1976, average grain production was similar to the average for the 1930s and slightly higher than that observed during the first three decades of the century.

In 1976, the military *coup-d'état* brought new policies implying financial and trade openness and the reduction of Government intervention. Although this helped to increase grains production, it did not imply the reappearance of futures markets. Grains operations continued to be regulated, controlled and monopolised under the Junta Nacional de Granos. The financial liberalisation pursued during the second half of the 1970s did not reach the futures markets. Future markets operations continued to be minimum even though the economic policy seemed to move in the direction of more deregulated markets. The presence of a grains board meant that, during a period of particularly high instability in commodity prices, futures markets in Argentina remained severely underused.

The ever-increasing macroeconomic instability during the second part of the 1970s brought higher inflation and increased devaluation expectations. Producers' primary concern was not the fluctuations on the international prices of products, but fluctuations in the purchasing power of the peso. Operations in foreign currency provided a more convenient hedging strategy. Speculators, on the other hand, had other more profitable opportunities in the speculation in the currency markets than in the futures markets.

The Crisis of the Debt of the beginning of the 1980s did not help in securing macroeconomic instability. The hyperinflation observed in 1989-1990 and the change in the political economy (Washington Consensus) implied the abrupt and revolutionary end of a stage that was characterised by stagnant grains production and almost non-existent futures markets operations.

2.5.4. Fourth stage: 1992-2013

As we have seen in the previous section, the changes in the economic policies at the beginning of the 1990s had important implications in the economy. Not only did they influence the macroeconomic variables, they also implied a change in the perception of the market as a coordinator of economic resources. This meant that, following the Washington Consensus, a radical structural reform of the economy was carried out. Introduction of more market friendly institutions, privatization of public enterprises and the general deregulation of markets had important effects on the economy.

This implied the end of a period in which the agricultural sector was seen by policy makers as merely a provider of foreign currency to assist the industrial sector. It also implied the end of the intervention of the Government in the commercialisation of grains (the Junta Nacional de Granos was dissolved in 1991), and in the provision of inefficient and expensive services (Grains elevators and port facilities were privatised). Although international prices were better than during the 1980s and some newer technologies were behind, the economic reforms implemented during the 1990s constituted a key element in the explanation of the doubling of production of grains between 1991 and 2001.

The control of inflation, the deregulation of the grains market and the generally more favourable view of the market economy, implied newer impetus in the operation in futures. Although it has never surpassed the values of output, operations in futures have reached the same values observed during the first stage identified here. The increase in the agricultural output experienced during this decade also boosted the growth in the future markets.

Although by far, the Chicago Board of Trade is the most important futures market in the world with volumes surpassing output by a factor of 80, both Argentine markets are important international future markets. Considering soybeans, the MATBA and the Rofex are the fourth and fifth most important markets in the operation of futures as Facciano (2011) shows. However, the difference in operated volumes between both Argentine markets and the Chicago Board of Trade is not associated with the differences in output but to differences in the regulatory system and in the enforcement of contracts that facilitate the operation in the American market.

Moreover, the use of international futures markets has been facilitated by the fall in the cost of telecommunications and the deregulation of financial markets. This has provided agents with a hedging tool against fluctuations in prices and a profit opportunity in the speculation on the differences between futures and expected prices. This means that, although the Argentine futures markets might be small, the utilisation of international futures is a widespread practice. Moreover, even if agents do not participate in them, as we have seen in the previous chapter, they provide agents with a reference price on which to base economic decisions.

2.6. CONCLUSIONS

This chapter has justified the choice of commodities, period and country, and the precise topic of futures markets. It is intended to serve as a contextual and referential guide for the empirical chapters and to provide a background to the whole research.

With respect to the relevance of wheat, maize and soybeans in the context of the Argentine economy, figures have shown that these three are extremely important for the country, representing together nearly 13% of total Argentine exports. Moreover, they have indicated that Argentina is a key supplier of the three commodities in the world market.

On the other hand, characterisation of the commercialisation and identification of the key moments in their production processes were carried out for the three commodities. This was explored with the objective of identifying the seasonality and its sources in the supply of these commodities. Given the dimensions of Argentina, the identification of seasonality was highlighted as not being as straightforward as in other cases.

At the same time, a history of the operation of futures markets in Argentina was presented and their relevance in Argentina described. It was seen that farms tend to differ from other developing countries given their size and the intensive use of technified production. Other specific characteristics of the agriculture production and geography of Argentina justify the existence of organised markets and, given the volatility, the existence of futures markets. At the same time, producers can finance their existence and the volumes operated justify their existence.

A review of the history of economic policies in Argentina, with an emphasis on the last 25 years, has been undertaken, enabling the highlighting of the different general and sector-specific policies followed by Argentina and their effects in the agricultural sector. The idea behind this is to provide evidence of and justification for some unique or unorthodox approaches that will be followed in the estimation stages – for example, the few real effects that a nominal variable may have in a context of historical and expected extremely high inflation and a lack of confidence in the country's own currency.

CHAPTER THREE

SEASONAL UNIT ROOTS AND STRUCTURAL BREAKS IN AGRICULTURAL TIME SERIES: MONTHLY EXPORTS AND DOMESTIC SUPPLY IN ARGENTINA

Summary

Monthly time-series data based on agricultural commodities tend to present strong and particular patterns of seasonality. The presence of zero values in some of the seasons is not explained by the absence of reporting but is the result of actual features of agricultural processes. Seasonal unit root tests have never been applied to data that exhibit these characteristics, with a consequent lack of critical values to be used in the inference. Monte Carlo simulations are performed to obtain critical values that can be used for this type of data. In addition, seasonal unit roots under the presence of unknown structural breaks have never been applied to any kind of monthly time series, with the associated absence of critical values to be used in the testing procedure. Monte Carlo simulations are also performed to tabulate these critical values. It is observed that the presence of zero values does not invalidate the critical values available, with or without unknown structural breaks; the values obtained here for the monthly seasonal unit root tests under unknown structural breaks can be used in any other kinds of exercise. A seasonal unit root test with more power is also considered and critical values are obtained to perform the inference. The capability of the seasonal unit root tests to select the right break date is analysed, with some divergent results with respect to previous findings. An application of these techniques on the monthly quantities of exports and domestic supply of three agricultural commodities in Argentina between 1994 and 2008, which observe the patterns of seasonality described, is presented. Although, some evidence of stochastic seasonality has been found in some of these series, in general a deterministic approach can adequately describe their seasonality.

3.1. INTRODUCTION

Seasonality is a distinctive feature of many economic time series. In some cases, seasonal patterns may be responsible for the explanation of an important part of the variation observed in the series. The factors that generate seasonality are diverse. Weather, climate, institutional arrangements, and even the culture could affect the moment at which certain activities are performed. Specifically in the case of time series associated to agricultural processes, given its exposure to the climate cycle and in addition to other factors, they tend to present important seasonal patterns associated with specific moments such as the implantation and the harvest.

The econometric treatment applied to series that exhibit seasonality depends on the nature of that seasonality. If the seasonality is deterministic, in the sense that its pattern can be predicted and is stable over time, there are direct approaches that could be used to adjust, control or model the effect of seasonality in the variable of interest. The introduction of dummy variables to capture the incidence of each of the seasons is the general approach used either to model seasonality or to remove their effect in the series.

However, the analysis becomes more complex when stochastic elements are considered and they affect permanently the seasonal pattern. Although, stochastic elements whose properties do not violate the assumptions behind the estimation methods would not present particular problems, if these stochastic elements, given their nature, have a cumulative effect on the series, they could complicate the estimation and the inference. These permanent effects may introduce additional unit roots to the one that might be seen at the zero frequency, indicating the presence of a stochastic trend. The application of a difference operator (to remove the stochastic trend) and estimate the model in their differences, would not solve the issue as it is expected that there would be more than one unit root (one for each season).

Moreover, even in the case where this procedure may fix the issues related to the estimation, the procedure implies the waste of relevant information about the long-run relationship between the variables and the change in the nature of the problem under study. If the model is trying to capture the presumably stable relationship between two variables, the estimation of the model in differences reflects how changes in one variable affects changes in the other, which is a short-run response rather than the long-run relationship hypothesized.

These aspects were addressed, in the context of stochastic trends, with the developments of integration and cointegration. The Engle and Granger (1987) cointegration approach considers that if two (or more) series are integrated of the same order, there may exist a linear combination of the series integrated from a lower order. In that case, even though the series may contain stochastic trends (i.e., be non-stationary), they will move closer together over time such that a linear combination of them will be stationary. Therefore, it is possible to estimate a single equation model even if the two series are not stationary (i.e., both containing a unit root) and the residuals of that estimation are integrated of order zero or free of unit roots. This implies that, as long as the cointegration assumption can be sustained, the estimation of that model will provide consistent estimates and the inferences offered will be statistically valid.

In order to meet the necessary conditions for cointegration, a test for the presence of unit roots in the series under study must be undertaken. The Dickey and Fuller (1979) test, or its parameterized version including lagged values of the autoregressive process of the variable, the Augmented Dickey-Fuller (ADF) test, may do the work. The distribution of the statistic of this test is not standard and special critical values are used for inferential purposes.

If these seasonal patterns are stable or repetitive over time, dummy-style deterministic approaches can be attempted to describe its behaviour. In this case, the standard ADF test for the presence of stochastic trends remains valid. However, if the innovations have persistent effects that reshape or introduce a new seasonal pattern, the use of deterministic approaches that do not consider the presence of these seasonal unit roots will result in inappropriate adjustments. Given that changes in technology or in the institutional frameworks, although one-off events, may take time to be widely implemented, the possibility that agricultural products present stochastic elements in the seasons cannot be excluded.

This suggests that unit roots may be present in the long run (or at the zero frequency) and/or in each of the seasons. Therefore, any potential cointegration relationship might occur at seasonal cycles as well as in the zero frequency. If seasonal roots are present and the cointegration relationship is thought to be a long run one (only at the zero frequency), the relationship between the two series might give inconsistent estimates. Therefore, it is important to test for the presence of seasonal unit roots before applying the appropriate cointegration technique.

Hylleberg et al (1990) (HEGY) developed a technique for testing unit roots at different frequencies. Their technique clearly distinguishes between long run roots (or standard unit roots) and seasonal roots at different cycles (semi-annual, bi-monthly, etc.). Initially applied to quarterly data, their technique can be extended easily to monthly data.

The applications of this technique have dealt primarily with indexes (production indexes) or some types of aggregated data (GDP, investment, etc.). In general, this type of data, by construction, is always positive or contains only non-zero values. The possibility that in one of the seasons the series could contain observations with a value of zero is not considered by construction. However, several agricultural series exhibit a strong seasonal pattern where in some of the seasons the series adopt a value of zero. This is typical in agricultural annual crops where the harvest seasons are determined clearly by climate, although weather may introduce some noise, and where a sequence of positive observations is followed cyclically by a sequence with zero values. Whilst this fact does not invalidate *per se* the HEGY test, the presence of zero values affects the data generation processes (DGP) underlying the test and the critical values used for making the inference. One of the objectives of this chapter is to apply the HEGY technique to this type of data and analyse its implications for the testing procedure.

On the other hand, as in the popular Dickey-Fuller test for unit roots, where the statistic of interest follows no conventional probability distribution, in the HEGY approach the set of statistics for testing seasonal unit roots requires the tabulation of special critical values. While the literature has found sets of critical values for quarterly and monthly data, it is necessary to find out whether those critical values are affected by the fact that the series could contain zero values. Therefore, in this chapter, we have also tabulated critical values for testing monthly unit roots when zeros are present in the series using Monte Carlo techniques.

In addition, Perron (1989) has found that the standard unit root tests could lack power if the series contain a structural break. In this case, the structural break could be disguising an otherwise stationary process as a non-stationary one. This also applies when testing for seasonal unit roots. In the presence of breaks, the HEGY approach will indicate too many unit roots when the process could be stationary. Developing countries, such as Argentina, are generally subject to severe and sudden institutional and legal changes that can introduce modifications in the way and time commodities are traded. Therefore, it is also relevant to consider the possibility of breaks in the series when testing for unit roots.

When the date of the break is known beforehand and can be identified, the correction of the unit root tests is relatively straightforward. It is possible and likely, however, not to know where the break is located. In that case, the date of the break must also be estimated. As in the non-seasonal case, critical values are non-standard and differ from those used when no structural breaks are assumed. In this chapter, we have also tabulated critical values for monthly data that are not available even for more general data generation processes, when the data generation process contains breaks and the series is allowed to take zero values.

Testing for unit roots using the HEGY approach when unknown breaks are present entails fitting a model with several parameters. When series are too short, this may generate some power problems given the number of parameters to be estimated with a limited length of data. Therefore, in this chapter, we have also considered the possibility of testing for seasonal unit roots using a single dichotomous variable to capture the break. This increases the power of the tests by decreasing the number of parameters to be estimated, but it does not allow for breaks being present in seasons. If the nature of the breaks is assumed to be affecting the level of the series rather than the seasonal pattern, the procedure can be applied without loss of generality. However, in this case, rather than considering the possibility of a structural break affecting the seasonal pattern, we are considering a break affecting the trend or the zero frequency. Appropriate critical values, given the presence of zero values in the variables, have also been tabulated for this procedure.

Some authors, such as Harvey, Leybourne and Newbold (2002) have found that in quarterly data when the size of the break is very large, the procedure to identify the time of the break in the HEGY approach tends to misplace the time of break with respect to the true time. Because of this, the HEGY approach could suffer from test size problems if the time of the break is not corrected. This chapter also tries to verify this problem using monthly data when zeroes are allowed for in the series, either in the full model for identification of seasonal unit roots under the presence of breaks, or in the simplified model using a single dummy variable.

As an illustration of these techniques, we have applied an intensive analysis on the seasonality of six series of monthly exports and monthly domestic supply of three annual crops (soybeans, maize and wheat) in Argentina between 1994 and 2008. First, an inspection on the series through the analysis of the autocorrelation function (ACF) and partial autocorrelation function (PAF) has been performed with the objective of evaluating how these series compare to the theoretical functions for monthly seasonal process. Additionally, a deterministic approach has been applied with the objective of assessing how much of the variation in the series can be explained.

In addition, a HEGY test has been applied on the series with the objective of evaluating the presence of stochastic seasonality. In one of these series, in non-harvest time the series take zeroes in those specific seasons. Additionally, in those series where it was not possible to reject unit roots at the seasonal frequencies, the HEGY approach has been applied allowing for the possibility of structural breaks using the full model and the simplified one, in order to verify if the unit root were spurious or not. Moreover, the misplacement on the identification of the unknown structural break has been evaluated in the context of these series.

This chapter is structured as follows. In the first and second parts, we give a description of the nature of seasonality in agricultural time series and different treatments generally used to address it. In addition, we discuss their problems and then introduce the effects on seasonality of stochastic elements. In the third part, we make a brief presentation of the HEGY approach using monthly data. In the fourth part, we present the critical values for this test when zero values are allowed in the series and we compare them to the critical values already found in the literature. Then, an application on real data is presented to illustrate how to implement the procedure. In the sixth part, a discussion on the procedure for testing seasonal unit roots when the location of structural breaks is unknown is considered, and critical values for both the extended and simplified model are tabulated. The misplacement of the date break is also evaluated. Finally, an application of these techniques is considered to ascertain if the unit roots found before are present when structural breaks are also relevant for the series.

3.2. SOME FEATURES OF AGRICULTURAL TIME SERIES AND THEIR SEASONALITY

Many economic processes present some form of seasonality. Series associated with tourism, retail, and, in particular, agriculture present clear seasonal patterns, generally associated to weather and climate. Sometimes, the seasonality is of such importance that it can explain by itself alone the complete variance of the series. This implies that forecasts that ignore the seasonal pattern are expected to present higher variance (Enders, 1995), and not provide an accurate representation of the process. The objective of this section is to motivate the discussion of seasonality in agriculture production by highlighting the elements that generate these type of phenomena. The next section will give a theoretical discussion on the deterministic seasonality.

Production of agricultural products, particularly annual crops, presents particularly identifiable moments that are generally located at a known period. It is known that wheat is harvested at the end of spring or the beginning of the summer, and soybeans, for example, are harvested in autumn. This gives seasonal characteristics to agricultural series. However, in agriculture, seasons cannot be clearly defined and several factors may alter their length and time.

The length of seasons depends heavily on latitude, where areas located far from the Equator exhibit more variability in solar radiation, thus amplifying the effect of seasons. Moreover, the existence of *a priori* stable cyclical climate phenomena that vary according to the area, and vary in duration and occurrence, also introduce distinctive effects on seasons. This is in addition to the other effects on the level and trend of the series. This means that the length of the seasons cannot be universally defined, affecting the time when harvests (and other production activities related to agricultural production) take place.

This characteristically seasonal feature of agricultural production notably affects all related economic activities. Consequently, production, commercialization and the income of producers and their consumption exhibit similar seasonal patterns. Under the assumption that agents prefer smoother patterns of consumption, the presence of seasonality in their incomes presents a problem. If financial instruments are available, it is possible to transfer income between seasons and reduce the seasonal effect in consumption. Nevertheless, if these elements or other economic phenomena (such as inflation) are present, smoothing the consumption pattern may not be possible.

Nevertheless, some instruments exist to reduce or smooth the seasonal pattern in agricultural production and its effects, for example, in the export and domestic supply of a commodity. Storage is notably one of them. The possibility of postponing the supply of a production and avoiding selling at a time when the supply is very high can help to reduce seasonality. This means storing the harvest and selling it when the total supply is lower and prices are higher. Nevertheless, several factors can reduce this capability and the availability of storage.

Storing a commodity postpones an immediate and certain income for an expected higher income in the future. Nevertheless, an agent would often like to bring that future income to the present in order not to alter substantially his current consumption. Moreover, debts originated in the past (to finance the sowing of the current harvest, for example) may require the availability of current proceeds. If financial instruments are available, this can be done relatively easily. The agent can take a loan that will allow him to wait until a more appropriate time for selling its product. However, if standard financial instruments are underdeveloped or absent, a typical feature in developing countries, storage becomes difficult and producers are forced to sell as soon their product is ready.

On the other hand, storage requires a physical place to store the production. Whilst today it is possible to store in large bags that makes building infrastructure (silos) unnecessary, it is still a costly activity. Nevertheless, only a part of the production can be stored and the rest must be sold definitely at the time of harvest. Therefore, storage physical capacity places a limit on the capability of the seasonality reduction of storage.

Moreover, changes in the regulatory, legal and tax frameworks may substantially affect the convenience and availability of storage. Governments needing income or facing problems in the balance of payments may encourage (if not force) the rapid liquidation of stocks. This is particularly present in countries with weak institutional frameworks and/or when taxes on particular commodities generate a large share of Government income or represent an important share of exports. Because of all this, the capability of storage to reduce or dampen the effect of seasonality on exports and the domestic supply may be reduced.

On the other hand, seasonality in production, exports and domestic supply of some commodities is particularly acute. Whilst aggregated series such as consumption or production of industrial goods present seasonality, generally there is always activity in every season. This may not be the case in some agricultural time series where a zero value reflecting, for example, no exports is definitely common. This is probably the most distinctive characteristic of the seasonal pattern in some agricultural time series. This does not imply that observations are missing or they have not been collected. It means that zero is part of the domain of the series. A feature that is not very common in other economic series or that is hidden by temporal aggregation. Therefore, this feature must be taken into account at the time of working with time series in agriculture whatever the purpose of the analysis.

Since seasonality is a distinctive feature of agricultural time series (and other economic series) some approach or procedure must be applied at the time of the analysis. Ignoring seasonality may generate different estimation problems and inaccurate forecasts. Enders (2010) highlights that forecasts whose models have ignored seasonality, may have higher variance. In addition, ignoring seasonality implies losing important and relevant information about the economic process under analysis.

The application of filters or other seasonal smoothing procedures, before estimation, have been a common way of treating seasonality. It entails applying some filter that removes the seasonal component and leaves the series only with their trend and/or cyclical and irregular components. The US Census Bureau has developed the X-11 and X-12 methods that have been applied extensively for these purposes.

The first problem that appears with this type of filtering is the assumption that a series can be decomposed in such components as highlighted by Franses (1996). If seasons are not regular, in the sense that weather or other phenomena, including changes in the behaviour of agents, affect the specific time when a recurrent event occurs, the possibility of separating the series into these components is reduced. In this case, seasonal adjustment may remove or hide relevant information on time series that may describe the behaviour of agents.

Additionally, Ghysels (1990) suggests that seasonal adjustment implies important changes in the data generation process. Since seasonal adjustment implies the smoothing of the series, higher persistence and higher first-order autocorrelations are introduced. This implies that, for a given sample size, it may be harder to reject a null of unit-root hypothesis or the power of the unit root tests are reduced.

This implies that whilst seasonal adjustment may be useful for presentational purposes or simple analysis about the tendency of a particular series, its use in the estimation of econometric models, testing and inference is avoided. Therefore, some procedure to consider and account for seasonality is required.

3.2.1. Theory of seasonality in time series

As we have seen, implicit in any method of seasonal adjustment is a two-step procedure where first, seasonality is removed, and second, the autoregressive and moving average components are estimated using the Box-Jenkins method. However, as identified by Bell and Hilmer (1984), frequently the seasonal and ARMA coefficients are best identified and estimated at the same time. Consequently, it is convenient to avoid using seasonal adjusted data and use a technique that allows the joint estimation of the seasonal and the non-seasonal components.

In general, it is possible to use the Box-Jenkins method for modelling seasonal data as this does not differ substantially from that of non-seasonal data. The difference introduced by seasonal data of period *s* is that the seasonal coefficients of the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) appear at lags *s, 2s, 3s,…,* rather than at lags 1, 2, 3,… in the standard non-seasonal time series data. In this sense, it can be shown in Enders (1995) that the pure seasonal autoregressive model

$$y_t = a_{12}y_{t-12} + \epsilon_t \qquad |a_{12}| < 1 \qquad (3.1)$$

with y_t being a monthly time series and ϵ_t, a process with mean zero, constant variance and serially uncorrelated, present the correlogram given by the expressions

$\rho_i = (a_{12})^{\frac{i}{12}}$ if ${}^{i}/_{12}$ is an integer and

$\rho_i = 0$, otherwise. Therefore, the ACF of this pure monthly seasonal model will present decreasing spikes in periods 12, 24, 36, …. In reality, the identification will be more complicated because of the interaction of seasonal and non-seasonal components. This means that the ACF (and the PACF) for these processes will combine elements from both types of processes. Therefore, it is possible to represent the more general process by

$$y_t = a_1 y_{t-1} + a_{12} y_{t-12} + \epsilon_t + \beta_1 \epsilon_{t-1} \qquad (3.2)$$

This process presents the seasonal component presented before, plus the addition of an autoregressive (AR) component and a moving average (MA) component. The three components entered **additively** to the expression, implying that there is no interaction between the seasonal component and the ARMA components. It is possible, on the other hand, to consider the interaction of the different components, including the seasonal, through **multiplicative seasonality.** Consequently, it is possible to consider the process

$$(1 - a_1 L)(1 - a_{12} L^{12}) y_t = (1 + \beta_1 L) \epsilon_t \qquad (3.3)$$

Which can be rewritten as

$$y_t = a_1 y_{t-1} + a_{12} y_{t-12} - a_1 a_{12} y_{t-13} + \epsilon_t + \beta_1 \epsilon_{t-1} \qquad (3.4)$$

The process in (3.3) above differs only in the fact that the autoregressive term in lag 1 is allowed to interact with the seasonal autoregressive effect at lag 12. The advantage is that this form allows for a rich interaction with a small number of coefficients. Note that the process described by (3.4) differs from the one described in (3.2) only by the addition of the addition of the interaction of the autoregressive and seasonal components at lag 13. Therefore, by estimating three coefficients $(a_1, a_{12}, and, \beta_1)$ it is possible to obtain the moving average term at lag 1 and the autoregressive terms at lags 1, 12 and 13.

Nevertheless, the estimation of three autoregressive coefficients in (3.4) is interrelated. If the estimation of the model $y_t = a_1 y_{t-1} + a_{12} y_{t-12} - a_{13} y_{t-13} + \epsilon_t + \beta_1 \epsilon_{t-1}$, is attempted, a smaller sum of squared residuals is expected since a_{13} is not requested or constrained to be equal to $a_1 a_{12}$. However, the model given by (3.4) would be preferable since it is more parsimonious. This means that if the unconstrained estimate of a_{13} approaches the product $a_1 a_{12}$, the multiplicative model would be preferable. However, there are not, in principle, any theoretical reasons or foundations for the election of one type of seasonal modelling over the other.

The objective of the identification of the type of seasonality is an adequate methodology to treat it in the analysis or inferences. One approach, as we have seen, is to eliminate the seasonality through the application of filters or smoothing procedures. This not only has implications and costs in terms of information that is thrown away, but also presents conceptual complications with respect to the feasibility of such an approach as Franses (1996) identifies.

Another possibility is the use of regression procedures in which the seasonal and non-seasonal components are explained through a linear relationship. Following Pierce (1979), the additive seasonal model can be described as

$$y_t = p_t + s_t + e_t \tag{3.5}$$

where $p_t, s_{t,} and\ e_t$ are the trend cycle, seasonal, and irregular factors of y_t respectively. If, consequently,

$$p_t = \sum_{i=1}^{I} \alpha_i c_{it} \tag{3.6}$$

$$s_t = \sum_{j=1}^{J} \delta_j d_{jt} \tag{3.7}$$

And e_t has an expected value of zero, constant variance, and is not serially correlated (or the process is white noise); the components $p_t, s_{t,}$ are estimated for a sample $y = (y_1, \dots, y_n)'$ using the model

$$y = C\alpha + D\beta + e$$

That resulted from replacing (3.7) and (3.6) in (3.5). The elements d_{jt} are periodic variables, generally seasonal dummies. It could also include interactions with the time variable in order to capture a changing seasonal pattern. However, whilst the pattern might change, it is done through a deterministic pattern and not because of the effects of non-stationary innovations. This means that the seasonal component in (3.7), assuming knowledge of the error, can be predicted without error. The case where the seasonal pattern is affected by stochastic elements will be treated later in this chapter.

A simple example of a deterministic seasonal is the fixed periodic function

$$s_t = \sum_{j=1}^{12} \beta_j d_{jt} = \beta_t \qquad (\beta_t = \beta_{t\pm 12k}, k = 1,2, \dots) \tag{3.8}$$

Where $d_{1i}, \dots, d_{12t}$ are seasonal dummies variables and $\sum_{j=1}^{12} \beta_j = 0$, implying that the cumulative effect of the seasons should be zero. Therefore, the seasonal component for January is β_1, for February is β_{12} and so on.

This form of adjustment or control of seasonality tends to possess more desirable properties than the filters or seasonal removal procedures, as we have seen in the previous section. Therefore, the use of deterministic seasonality tends to be preferred in the applied work, in the inference as well as in the estimation of models.

For example, as seasonality might complicate the identification, inference and estimation using non-stationary series, it is possible to use dummy variables to remove the deterministic seasonal components and perform standard unit root tests on the residuals. In this sense, the estimation of the regression equation

$$y_t = \alpha_0 + \sum_{i=1}^{12} \alpha_i D_i + \hat{y}_t$$

removes from series y_t the seasonal components leaving in the residuals $\hat{y}_t$, the "de-seasonalised" value of y_t. Therefore, it is possible to evaluate the presence of unit root in these adjusted series by the Augmented Dickey-Fuller test as considered by Enders (1995, 2010).

In a more general framework, it is possible to consider a series given by a representation, where a trend term given by αt has been added.

$$y_t = \theta_0 + \alpha t + \textstyle\sum_{i=1}^{12} \alpha_i D_i + e_t \qquad (3.9)$$

Deterministic seasonality exists, following Pierce (1979) if all the α_i in (3.9) are not zero. Therefore, it is possible to test for the presence of deterministic seasonality with the hypothesis

$$H_0: \alpha_1 = \cdots = \alpha_{12} = 0 \qquad (3.10)$$

Rejection of this hypothesis may lead to the conclusion that some adjustment for deterministic seasonality may be necessary.

Consequently, if the seasonal pattern presents deterministic features, the methods presented in this section can be used straightforwardly to control or adjust for deterministic seasonality. However, the unit root testing procedure presented above, and other econometric techniques, may be invalid if a deterministic seasonal treatment is given to series that present seasonal unit roots. Before discussing the stochastic seasonality, we will devote some time to motivate the origin of this type of seasonality in the specific case of agriculture. This will help to analyse and understand the problem under study clearer.

3.2.2. Stochastic elements in the agricultural seasonality

So far, the discussion has been centred on the nature and effects of the deterministic seasonality in agricultural activity, and about different approaches to adjust, control or model it. If these seasonal patterns were stable or predictive over time, a simple deterministic approach, as we have seen, may effectively deal with them. As long as the effect of seasons is identified clearly, and it is stable in magnitude and in the season, the addition of dummy or dichotomous variables in the estimation to capture or control for the seasonal effect may be sufficient to obtain estimates of the parameters of interest that satisfy the model assumptions.

Nevertheless, changes in weather, calendar, and the behaviour of agents or other phenomena during particular periods may alter the seasonal pattern of a time series. Therefore, stochastic elements can affect the nature of seasonality and, consequently, this may require special treatment.

However, as time series data tend to be affected by multiple types of stochastic elements, it is important to emphasize the effects of those that are of relevance to the problem of stochastic seasonality. Among the stochastic components, it is important to highlight the difference between those that may have temporary effects, where these are not spread into the future (no serial correlation); from those shocks or innovations whose effects might influence the future values of the series. For example, whilst the effects of weather are expected to be concentrated in a single period; technological innovations such as the introduction of a new practice may spread over several periods.

Nevertheless, it is also important to make the distinction between what would constitute an innovation whose effects may cumulate over time, from an innovation that might change a deterministic seasonal pattern. Frequently, these structural changes might be confused by stochastic seasonality. We will discuss this type of phenomena and their treatment within the context of stochastic seasonality later in this chapter. Therefore, the focus is put on those innovations or changes whose cumulative effects do not die out fast enough or introduce stochastic elements that affect seasonal patterns.

Before presenting the theoretical discussion about the treatment of stochastic seasonality and how is detected, we will devote some time to introduce the elements that may generate this type of phenomenon in agriculture, putting emphasis on the specific case of Argentina. In Chapter 2, we made an exhaustive description of the features of agricultural production in Argentina, especially with respect to the location of the production and the factors that explains its seasonality. The discussion, in this section, will pay particular attention to those features that might exacerbate the case for stochastic seasonality or not.

When agricultural production is distributed widely within large countries, the sowing and harvest times tend to vary enough to impede the identification of a single time of harvest or season. As we have seen, agricultural areas in Argentina extend for many degrees of latitude. Additionally, as we have seen in the previous chapter, harvesting soybeans can begin as early as the end of February in some areas, and can finish as late as June in other parts of the country. Therefore, it is difficult to identify and determine a single period for each activity. If this pattern were stable and predictable, its behaviour could be determined and considered at the time of the estimation of an econometric model. Nevertheless, there are factors or stochastic elements that may affect locally and globally the harvest decision with its effect on the series of exports and domestic supply.

Weather is the first channel through which stochastic elements may be introduced in seasonality. For example, heavy rains may alter the location of the season. Heavy rains impede the work of agricultural machinery, delaying the effective time of harvest. In large agricultural areas, the heterogeneity in weather conditions imposes an additional instability component. Harvests, generally occurring at a particular time of year, may be delayed and their production value added to the following seasons, generally the harvest time of a different region. Therefore, not only does weather affect seasonality *per se*, but also the heterogeneity in weather conditions in large countries adds an extra instability to seasonal effects. However, as long as the weather has effects that are limited to the period when they occur, they would constitute standard innovations from a white noise process that would not require special treatment, and would not constitute a case of stochastic seasonality.

On the other hand, economic conditions may also affect seasonality. In large countries, if the level of the price used to decide about sowing has meant certain regions (whose harvest is placed generally during a certain time of the year) reduce the area effectively sowed and consequently its output, the value of the series in that particular season will be affected, altering the seasonal pattern. This means that the economic information may generate changes, not only on the value of the series, but also in the seasonal pattern. A similar effect may occur if policies applied by state or regional governments, alter the quantity or time of the sowing and harvest of a particular region associated with a particular season, particularly when these legislations or regulations are changed with high frequency.

Additionally, there is the issue of temporal aggregation. It is straightforward to define seasons or periods according to the convenience of the analysis. Monthly data can be aggregated easily into quarterly data, sometimes based solely on subjective elements. In that case, the presence of stochastic seasonality may be reduced since it is more likely that the instability of the season will be contained within a given quarter. Only at the borders of the quarters may the possibility of instability in the season appear. In this case, a careful examination of the data and the adjustment of the beginning of the quarters may reduce this effect. The definition of the quarters does not need to follow the year definition of the calendar. Whilst in general the first quarter always contains the first three months of the year, no econometric assumption or property will be affected if a quarter is defined by November, December and January. In fact, in agriculture (particularly in the South Hemisphere) will be perfectly justified to begin the agricultural year in the months of July or August, the months of the beginning of the sowing of summer crops. Nevertheless, this aggregation may be neither convenient nor desirable.

Monthly data aggregated into a lower frequency aggregation (such as quarterly) implies either the sum or the average over time of the data. We focus here on the aggregation of flow variables where aggregation is made by sampling every period of high frequency (Silvestrini and Veredas, 2008). In the case of averaging, this will smooth data (as the application of a seasonal filter) with similar implications for the estimation and the inference, as explained before. More importantly, Rossana and Seater (1995) suggest that the aggregation will eliminate long-term variation, since cycles that last more than a year, obvious in monthly data, tend to disappear when data are aggregated. This means that temporal aggregation implies the loss of information of the data generation process, and the real possibility of extracting false conclusions from estimation.

On the other hand, the aggregation may make the process lose economic sense. Whilst quantities could be summed eventually over three months to obtain a quarterly observation, some economic variables, such as prices, cannot be added. Given their importance in many economic functions, such as demand or supply of goods as well as factors of production, a treatment that can address this issue on prices is key. Whilst prices could be averaged, with the implications highlighted, the relationship that may need to be estimated may have very little economic sense, particularly when the volatility in prices is too high. Moreover, it has been suggested that predictors based on high frequency rather than those based on lower frequency have a superior performance (Amemiya and Wu, 1972). Additionally, they suggest that the least square estimator in the aggregate model will be inconsistent and will require additional lags and instrumental variables to make it consistent.

However, Wilcox (1992) highlights that, given the way that some series are constructed; monthly data can suffer from more measurement error than quarterly data. The reason is that, generally, monthly economic data are constructed from estimates based on samples rather than a complete survey of economic units. According to the author, the aggregation (using averages) into quarterly data may substantially reduce the measurement errors observed.

In the case of agricultural variables, such as the quantity produced of a good or the quantity exported, the incidence of sampling error is less important. Given that exports or the domestic supply of commodities are made by a relatively few traders or companies, measures of these quantities are collected taking declarations of the entire population of traders. This means that the aggregation into quarters is of very little added value, while creating serious estimation and inference problems.

An additional feature in seasonality is the presence of innovations that can introduce non-random or non-stochastic and permanent changes in the seasonal pattern. The introduction of new technologies and practices, such as direct sowing, that allows minimum previous tillage tasks, they also allow for the implantation of soybeans in January (immediately after wheat has been harvested), in addition to the regular sowing season in August/September (in the southern hemisphere). This may not only affect the seasonality in soybeans (given the new extra sowing season), but also the seasonality in the commercialization of wheat; this is because the costs of these short-cycled soybeans are generally paid with the proceeds of the sell or liquidation of the wheat harvest, reducing the stored wheat to be sold in other seasons. This structural break, whilst it may have a gradual effect in their introduction, redefines the seasonal pattern by affecting it once and for all.

Government regulations or legislation may also permanently affect seasonal patterns in the commercialization of products. The introduction of commodity boards, for example, where the body buys the harvest from producers and then sells it on the export or domestic markets, can also affect seasonality. Given its less binding financial constraints, the commodity board can hold the product or release it when it considers necessary or convenient for their operations in a different way from a single producer. In fact, the introduction of a commodity board eventually, if it has sufficient storage capability, may eliminate seasonality in the commercialisation of commodities.

These technological innovations and institutional changes constitute two important factors that might have a cumulative effect as they are adopted that might introduce stochastic seasonality. Although they might be, in their conception, one-off events, their effect may spread over many periods. This might be the case if the adoption or implementation of the new technology or institutional changes takes time.

For example, the authorization for the implantation of glyphosate-resistant soybeans in 1997 in Argentina had important effects on the quantity of hectares implanted with this crop. This, consequently, not only affected the area implanted with other crops, but also affected all the agricultural land. Areas previously idle were brought into production or changed from livestock farming. However, as the implementation of this variety in production has not been immediate, the effect on the total area implanted and, consequently, on the total production has cumulated over many periods. Producers required time to get used to the new variety and to include the rest of the elements necessary for the implantation of it. Therefore, this innovation, through its cumulative effect, could have affected the seasonal pattern of the series; moreover, it could affect its trend as well.

As they could be confused in the empirical analysis, it is important to highlight the distinction between what would be a structural break and an innovation that might have permanent or cumulative effects. A structural break would generate a punctual and identifiable effect in the any of the deterministic components of the series such as the trend or the deterministic seasonality. However, its effect is immediate or, as we will see, it is possible to identify or model its implementation process. This is in contrast to the case of an innovation that had a cumulative effect as it is adopted.

If the effect of the innovation is punctual and immediate, a deterministic approach can be followed and the treatment of a structural break in the context of seasonality is relatively straightforward. A dichotomous variable that separates the series into periods before and after the break may be effective. Of course, if the break is such that eliminates the seasonal pattern, it may be convenient to consider the estimation of two separate models.

However, a different problem appears when testing for stochastic seasonality. A structural break changes the nature of the seasonality pattern since the effect of this break may be conflicted and confused with the stochastic element, leading to the possibility of extracting erroneous conclusions from the seasonal unit root tests. Moreover, the problem may be more complicated if the location or the existence of it is unknown. The special treatment of the seasonal unit root tests under structural breaks will be discussed later in this chapter.

The stochastic elements in the seasonal pattern in agricultural time series only matter if they have a permanent effect on the seasonal patterns that are reflected into the future values of the series, or if they make the series permanently divert from a stable or repetitive pattern. If the stochastic elements only have a temporal or ephemeral effect, such that they can be considered part of the general stochastic elements present in the series, a deterministic treatment may be appropriate. However, only a proper testing procedure would shed some light on the nature of these stochastic elements. The HEGY test is the suggested tool for this purpose. However, the use of this technique must consider the distinctive features of agricultural time series in that series may contain zero values that are part of their domain, and that temporal aggregation into higher frequencies may be neither advisable nor convenient. This is the topic of the following section.

3.3. SEASONAL UNIT ROOTS IN THE CONTEXT OF MONTHLY AGRICULTURAL TIME SERIES

The possibility that agricultural time series, such as exports or domestic supply of commodities, contain stochastic seasonality cannot be ruled out. It was also highlighted that ignoring seasonality, or treating it inadequately, may have important implications in the estimation of econometric models and the inference that can be drawn from them. Consequently, this section will be devoted to the presentation of the methodology applied to test for the presence of seasonal unit roots and their implications when using agricultural time series available for monthly data.

The Hylleberg *et al* (1990) or HEGY test on the presence of seasonal unit roots is the basis of this analysis. The test has been developed to deal with seasonal unit roots in quarterly data and has been applied extensively on this frequency of data. Whilst the analysis of seasonal unit root tests (or standard unit root tests) may have important economic implications *per se*[26], its application is frequently associated as a prior step for cointegration analysis. Given that exercises on seasonal cointegration on monthly data are not very frequent in contrast to quarterly data, it is natural that the HEGY test has not received frequent attention on monthly data.

Since then, this procedure has been used intensively and its properties have been studied. Consequently, the HEGY procedure has become part of the econometric toolbox. Ghysels, Lee and Noh (1994) using Monte-Carlo techniques, have concluded that HEGY is the most appropriate testing technique for seasonal unit roots as well as for testing standard long-run unit roots when seasonal unit roots are present.

In addition to the original application made by Hylleberg *et al* (1990) on income and consumption in the UK, notable applications can be highlighted. Engle, *et al*. (1993)

---

[26] The validity of the purchasing power hypothesis, for example, has been frequently addressed by testing unit root tests, since the presence of a stochastic trend suggests that the hypothesis does not hold.

applied this technique at the time of analysing seasonal cointegration between consumption and disposable income in Japan. The HEGY technique has been extensively applied in the context of monetary economics in the search for seasonal cointegration relationships. Bohl and Sell (1998), Bohl (2000) and Herwartz and Reimers (2003) are some of the contributions that can be identified, among others here. These applications, as well as the gross of the applied literature, have been made on quarterly data.

Nevertheless, some applications can be found in data of other frequencies. The extension to monthly data has been done and applied on industrial data on The Netherlands by Franses (1991). Beaulieu and Miron (1993) analysed different monthly US aggregates, whilst Taylor (1998) applied HEGY on US unemployment and Canadian industrial production. More recently, Mugambe and Reilly (2007) analysed stochastic seasonality on different industrial aggregates in Uganda. However, all these applications have been made primarily on indexes or different economic aggregates.

The application on agricultural products has been scarce. De Pablo Valenciano, Perez Mesa and Levy Mangin (2008) find some evidence of monthly seasonal unit roots in the exports of tomatoes from a particular region in Spain. However, the perishable nature of this product, as well as its more continuous production, reduces its value as a reference.

It is convenient to make a brief review of the HEGY testing procedure on monthly data.[27] Let $y_t$be the monthly series in question, generated by an autoregressive process of the form

$$\varphi(B)y_t = \varepsilon_t$$

where $\varphi(B)$ is a polynomial in the backshift operator and ε_t is a standard white noise disturbance. Let λ_k be the roots of the characteristic polynomial of $\varphi(B)$. Some or all of the λ_k may be complex. In the polar representation of the characteristic root,e^{ai}, the value of α is the frequency associated with a particular root. A root is seasonal if $\alpha = 2\pi j/S, j = 1 \dots, S-1$, being S the number of observations per year. Therefore, for monthly data, the seasonal unit roots are

[27] A concise explanation of this procedure on quarterly data can be found in Ghysels and Osborn (2001), Harris and Sollis (2003) and Enders (2010).

$$-1;\ \pm i:\ \left(-\frac{1}{2}\pm\frac{\sqrt{3}}{2}i\right);\left(\frac{1}{2}\pm\frac{\sqrt{3}}{2}i\right);\ \left(-\frac{\sqrt{3}}{2}\pm\frac{i}{2}\right);\left(\frac{\sqrt{3}}{2}\pm\frac{i}{2}\right)$$

These roots correspond to the frequencies π; $\pm\pi/2$; $\pm 2\pi/3$; $\pm\pi/3$; $\pm 5\pi/6$; and $\pm\pi/6$. These frequencies correspond to the bi-monthly case[28], four-month case, quarterly case, six-month and 12-month cases. The idea is to know whether the polynomial has roots equal to one at the zero or seasonal frequencies. The testing procedure consists of linearizing the polynomial around the zero frequency plus the S-1 roots given above.

The HEGY technique allows the determination of the presence of unit roots in the long run as well as in each of the seasons. In the monthly case, the procedure requires the estimation of the following expression

$$\Delta_{12}y_t = \mu + \beta t + \sum_{k=2}^{12}\delta_k D_{kt} + \sum_{k=1}^{12}\pi_k Z_{k,t-1} + \sum_{i=1}^{\rho-1}\psi_i \Delta_{12}y_{t-i} + v_t \qquad (3.11)$$

Where μ and t are the drift and deterministic time trend terms, respectively, D_{kt} are deterministic seasonal dummies, and Z_k are transformations of the y_t that provide the basis for testing unit roots at zero and the rest of the frequencies. The definition of these variables can be found in Appendix I, and a more theoretical and technical description can be found for the monthly case in Beaulieu and Miron (1993).

In addition, as in the Augmented Dickey-Fuller tests, some lagged values of the dependent variable are included to assure the appropriate behaviour of the residuals. The literature suggests, among other criteria, to include all the lagged values until $\rho-1$; where ρ is suggested to be determined using the general to specific approach Ng and Perron (1995). The procedure starts from a very long lagged model and reduce the number of lags until the last lag included is statistically different from zero at some pre-specified level of significance (in general, a 10% level of significance is used). The introduction of these lags (as in the Augmented Dickey-Fuller tests) is to assure that the residuals have the standard properties (no serial autocorrelation). Finally, v_t is an error term with the standard white noise properties.

[28] If the unit circle is 2π, a month has π of a cycle; therefore, every two months there is one cycle and there are six cycles in the year. On the other hand, if a month has $\pi/2$ of a cycle, every four months there is a cycle and there are three cycles in the year.

The procedure entails testing for unit roots at different frequencies of the series. In order to test for the presence of a unit root at zero frequency, a left-sided t-test is used on the hypothesis $\pi_1 = 0$, against the alternative that $\pi_1 < 0$. In the same way, it is possible to test for the presence of a unit root at the π frequency using a similar one-sided t-test on π_2. These test procedures are similar to those applied in the Dickey-Fuller test. For the remaining frequencies, an F-test is used on the joint hypothesis $\pi_k = \pi_{k-1} = 0$, k being an even integer.

Alternatively, instead of using F-tests for the remaining frequencies, it is possible to use t-tests for the joint hypothesis. For testing unit roots at the 0 and π frequencies, the procedure is the same one explained above. For the remaining frequencies, a test on $\pi_k = 0$, where k is even, is performed first with a two-tailed test. Under this test, the alternative hypothesis states that the even coefficients may be positive or negative. If the test fails to reject the null, a one-sided test is conducted on the null hypothesis $\pi_{k-1} = 0$ against the alternative $\pi_{k-1} < 0$. This is because we restrict the attention to alternatives having assumed that $\pi_k = 0$, although Hylleberg, *et al*. (1990) suggests that this procedure may lack power. Moreover, the procedure could be a little cumbersome in the case of data with high frequency. For these reasons, the first approach is generally preferred (using F-tests for higher order frequencies).

It is interesting to note that the usual Augmented Dickey-Fuller test of the null hypothesis of a unit root at the zero frequency is still valid, even when other unit roots at different frequencies are present. As long as sufficient lagged terms of the dependent variable are included, Ghysels, Lee and Noh (1994) show that the usual unit root test at the zero frequency is compatible with tests on seasonal roots. In this sense, one may view this procedure as a generalization of the standard unit root test.

As mentioned earlier, it is suggested that the lag selection criterion is based on the general to specific criterion. The reason for adding lags of the dependent variable in the equation rests on the necessity of having white noise residuals. Too many lags could lead to low powered tests, while too few lags could increase the empirical size of the tests. It is important to highlight that, while the general to specific criterion is the most popular criterion and seems to select a correct lag specification that generates well-behaved residuals, it is not the only criterion employed. For example, it is possible to base the lag length on some information criterion such as the Akaike Information Criterion (AIC).

Despite some claims that this criterion could lead to a very parsimonious model with implications for the size of the tests (Perron, 1997), some authors, such as Enders (2010), recommend its use. In addition, it is possible to use a specific to general determination procedure to select the lag length. Nevertheless, Hall (1994) shows that this procedure is inferior to the general to specific criterion since it is not asymptotically valid.

On the other hand, if the objective is to achieve properly behaved residuals, a more pragmatic approach could be employed. Mugambe and Reilly (2007) have followed this approach where the definition of the lag length is not based on any rule in particular. Instead, they selected the lag length that yielded better white noise properties based on different tests on normality and serial correlation on the residuals. This suggests that they have used a trial and error approach until the desired properties were achieved. However, Rodrigues and Osborn (1999) have warned about basing the order of augmentation based on serial correlation tests. They have found that passing those tests does not guarantee that seasonal unit root tests have the size close to the nominal one, particularly in monthly data.

As in the standard Dickey-Fuller test, the critical values used to validate the test are not the standard ones. In the HEGY approach, neither the standard *t*-distributed nor the *F*-distributed critical values are valid since the null hypothesis is formulated in terms of a non-stationary process. Using Monte-Carlo simulation techniques, different authors have obtained simulated critical values for different series lengths. In the case of monthly data, Beaulieu and Miron (1993) have tabulated the appropriate critical values for the tests for different specifications of the estimation equation (in terms of the presence of drifts and trends). These critical values have been augmented by Franses and Hobijn (1997) by providing an almost full set of critical values for almost any model specification in quarterly and monthly data. For a discussion on the distributional properties see Smith and Taylor (1998).

3.3.1. Zero values and their effects on critical values of seasonal unit root tests

Very few applications of the HEGY test have been made on standard raw quantities of data, and almost none on series with acute seasonality where in some of the months; zero is the value of the series. The problem of zero values in the series goes beyond the complexity in the estimation of the equation. The fact that it is impossible the application of logarithms to the series is the least important problem. A more important problem is the validity of the unit roots tests performed on series that present this characteristic.

Whilst the original development of the test made by Hylleberg, *et al*. (1990), and its posterior analysis and extensions by Ghysels, Lee and Noh (1994) and Beaulieu and Miron (1993), does not indicate that the tests are restricted to a particular set of values, the zero value effect resembles the case of truncation of dependent variables frequently observed in other types of applications, particularly in cross-sectional data. In these cases, the truncation problem is generated for working with a sample of elements, from a more general population, whose particular attributes exceed or do not surpass a particular value (i.e., samples of workers with wages above a certain value). Nevertheless, in contrast to the standard truncation case in econometrics, in the case we present here, the population distribution of the variable is the one that presents this phenomenon and not the distribution of a sample of it. The data we are using, that is a sample of a general time series process, has not been subjected to a truncation of any nature and the lower bound of the variable is a "natural" characteristic of the process.

A deep analysis of the implications of the inferences made on this type of data as well as its asymptotic properties may be required. The interest of this chapter is mainly on determining the presence of stochastic seasonality in agricultural time series of exports and domestic supply required for the eventual seasonal cointegration analysis. Therefore, the theoretical analysis on the asymptotic properties of the estimators and inference under this type of data is left for future work. Thus, we will focus on some practical implications that the use of this type of data may have at the time of testing for the presence of seasonal unit roots.

Probably, the most important practical question is if the critical values already tabulated are still valid for data generation processes (DGP) where seasons could repeat values of zeroes. If this affects the critical values already tabulated, the application of the HEGY approach to this type of data will not be valid since it could lead to incorrect conclusions. Additionally, some model specifications (in terms of the inclusion of trend, deterministic seasonal dummies, and constant) have not been tabulated previously in the literature. If the presence of zero values does not affect the distribution of statistics, the additional model specifications can complement the tables already available.

Therefore, this chapter conducts a set of Monte Carlo experiments to obtain the appropriate critical values to verify if they differ from those already found in the literature and determine whether they can be used in the application presented below. In order to obtain the critical values for the zero and π frequencies, 24,000 replications were carried out using a simulated process in order to obtain the critical values to test the zero and π frequencies using t-tests. For the F critical values for the rest of the frequencies, 120,000 replications were used following Beaulieu and Miron's, (1993) suggestion. Simulations were done using STATA 10. The disturbance in the series follows a standard normal distribution. The simulated DGP takes this form

$$\Delta_{12}\tilde{y}_t = e_t$$

Where e_t is an standard white noise process and $\tilde{y}_t$is a modified process where

$\tilde{y}_t = y_t$ if $y_t > 0$

$\tilde{y}_t = 0$ if $y_t \leq 0$

Table II.1 (in Appendix II) presents the critical values obtained. It can be seen that the values obtained do not differ substantially from the ones found by Beaulieu and Miron, (1993, pp. 325-326) and Franses and Hobijn (1997, pp. 29-32). For example, when intercept, seasonal dummies and trends are considered and for a series of 240 observations and 5% level of significance, Beaulieu and Miron, 1993) tabulate -3.28, -2.75 and 6.23 for the `t':$\pi_1$ , `t':π_2 and F:$_{\pi odd,\ \pi even}$ critical values, respectively; while this analysis has tabulated critical values of -3.29, -2.78 and 6.06 for the same specification and similar series length. The small differences may be attributed to sampling error and the fact that our simulations did not consider the exact series lengths that these authors tabulated.

On the other hand, the general features of the different values according to the model specification are shared between the critical values presented here and those previously found. In general, specifications with no deterministic seasonal dummies generate critical values that favour the rejection of the null hypothesis of a unit root, for example. This suggests that specifications that include these elements will require high values on the statistic in order to reject the null.

Since the critical values tabulated in this analysis do not differ from those already found, and additional model specifications (all combinations between trend, intercept and seasonal dummies were considered) are tabulated, this table augments the critical values found by those authors. The fact that the critical values are similar to those already tabulated suggests that the presence of zero values in series that exhibit strong patterns of seasonality, do not seem to invalidate the HEGY procedure and its critical values and they can be used in more general time series. The additional model specifications will thus be useful for further research in this field or in others. Consequently, we will proceed in the next section to the application of the HEGY test on a particular set of data that present the mentioned characteristics.

3.4. EVIDENCE OF SEASONAL UNIT ROOTS IN AGRICULTURAL COMMODITIES

We will apply the HEGY procedure to six series of exports and domestic supply of three agricultural commodities in Argentina; their definitions can be seen in Table 3.1. The series used are quantities and, since zero values are present in some of them, we have made the analysis using the data in levels rather than transforming the variables using logarithms. This treatment has been followed for all series in order not to lose generality in the treatment.

Before entering into the analysis of seasonal unit roots and its extensions, we will devote some time and space to inspect and analyse the series as much as possible. This will be done by describing the graphical depiction of the series, as well as presenting the autocorrelations and partial autocorrelations. This should help to extract as much as information as possible from the series before entering into more formal analysis.

Table 3. 1 Series used in the analysis of seasonal unit roots

Series	Description	Time span
Qesoy	Quantity of exports of soybeans (in 000' tons)	Jan/94 – Jan/09
Qdsoy	Quantity of domestic supply of soybeans (in 000' tons)	Jan/94 – Jan/09
Qemaz	Quantity of exports of maize (in 000' tons)	Jan/94 – Sep/08
Qdmaz	Quantity of domestic supply of maize (in 000' tons)	Jan/94 – Sep/08
Qewht	Quantity of exports of wheat (in 000' tons)	Jan/94 – Sep/08
Qdwht	Quantity of domestic supply of wheat (in 000' tons)	Jan/94 – Sep/08

Source: Ministerio de Agricultura, Ganadería, Pesca y Alimentos de la Republica Argentina.

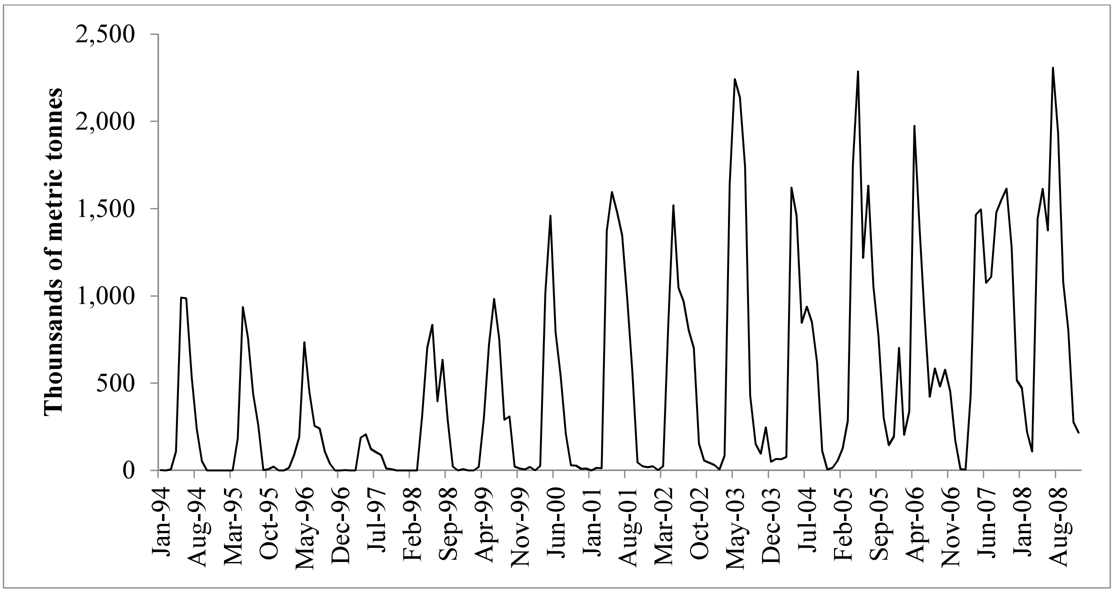
3.4.1. Graphical inspection and analysis of the series

Figure 3.1 presents the monthly exports of soybeans in physical quantities from January 1994 until August 2008. A clear seasonal deterministic pattern emerges, with export activity after or during the harvest time (between February and August), with no activity in the remaining months of the year. This pattern seems to be stable in the sense that tends to repeat over the sample.

Additionally, in Table 3.2 the corresponding autocorrelation and partial autocorrelations is presented. Whilst in standard non-seasonal process the coefficients of the autocorrelations and partial autocorrelations would present spikes at lags 1, 2, 3, … (presenting a rapid decay if the process tends to be white noise), in a purely deterministic seasonal process, the spike would appear at each season. This means that in a purely deterministic monthly seasonal process, the theoretical autocorrelation function will present spikes at 12, 24, 36, … and zero for the remaining periods.

As can be seen, the shape of the actual autocorrelation function tends to mimic the theoretical seasonal autocorrelation function. Spikes tend to appear in months 12, 24 and 36. However, the remaining periods are not zero as in the theoretical autocorrelation function. Moreover, the Q-statistic, based on Ljung and Box (1978), also suggests that these are not statistically different from zero.

Figure 3.1 Argentina's exports of soybeans 1994-2008 (in thousands of metric tonnes)



Source: Ministerio de Agricultura, Ganadería, Pesca y Alimentos de la Republica Argentina

Table 3.2 Autocorrelations and Partial autocorrelations – QESOY

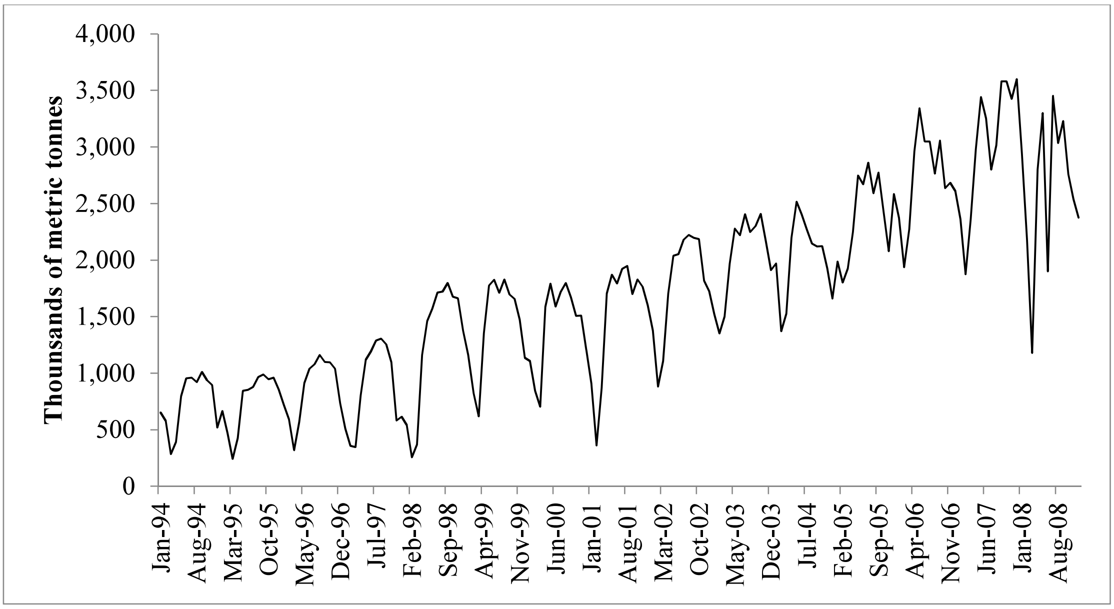
LAG	AC	PAC	Q	LAG	AC	PAC	Q
1	0.7423	0.7423	101.39	21	0.0662	0.1125	453.93
2	0.4092	-0.3159	132.37	22	0.2636	0.0119	468.4
3	0.1587	-0.016	137.06	23	0.4978	0.2112	520.36
4	-0.073	-0.229	138.06	24	0.6381	0.1036	606.26
5	-0.1939	0.0542	145.14	25	0.4891	-0.1122	657.07
6	-0.2266	-0.0661	154.85	26	0.2995	0.1036	676.23
7	-0.1833	0.0762	161.25	27	0.163	0.1101	681.95
8	-0.0529	0.1088	161.79	28	-0.0145	-0.0805	681.99
9	0.1446	0.2167	165.81	29	-0.1537	0.0569	687.14
10	0.3321	0.1371	187.17	30	-0.2181	0.0151	697.57
11	0.5721	0.473	250.94	31	-0.2218	-0.0741	708.44
12	0.6997	0.1142	346.89	32	-0.1434	-0.1125	713.01
13	0.5196	-0.2694	400.11	33	0.0202	0.0656	713.1
14	0.281	0.0346	415.77	34	0.201	-0.0185	722.21
15	0.1118	0.1586	418.27	35	0.4341	0.0932	764.96
16	-0.0608	-0.0604	419.01	36	0.5614	-0.0649	836.96
17	-0.1872	-0.0963	426.09	37	0.4167	-0.089	876.91
18	-0.2565	-0.1336	439.45	38	0.2484	-0.0086	891.21
19	-0.2287	0.0711	450.15	39	0.1019	-0.0798	893.63
20	-0.1182	-0.1098	453.02	40	-0.0619	0.0562	894.53

Source: Own calculations

Figure 3.2 presents the monthly supply (in thousands of metric tonnes) of soybeans to the domestic market between January 1994 and August 2008. The graph suggests the presence of seasonality with systematic higher activity during the harvest periods; however, in contrast to the case of the exports presented before, the activity does not disappear in the remaining months.

The autocorrelations and partial autocorrelations in Table 3.3 indicate that, although there are spikes in the periods associated with the seasonality component, the non-zero values for the remaining periods suggest that a deterministic seasonality might be insufficient in providing an adequate explanation for the seasonal component. On the other hand, the slow general decay suggests the existence of unit roots, which could be at the zero or at the seasons.

Figure 3.2 Argentina's domestic supply of soybeans 1994-2008 (in thousands of metric tonnes)



Source: Ministerio de Agricultura, Ganadería, Pesca y Alimentos de la Republica Argentina

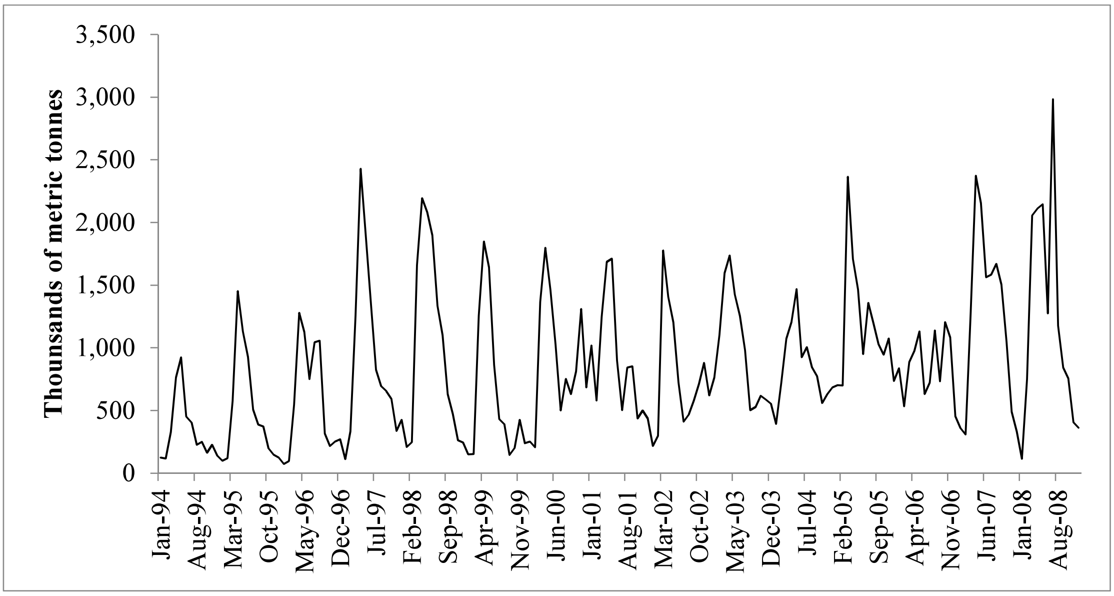
Monthly exports of maize between January 1994 and August 2008 are presented in Figure 3.3. Although some seasonality can be observed, with more active exports immediately after the harvest, this pattern is less pronounced than in the previous cases and cannot be seen as repetitive or systematic. However, the autocorrelation and partial autocorrelation functions in Table 3.4 present spikes at 12, 24 and 36 months, typical of deterministic seasonality. In addition, negative spikes appear in the intermediate months (6, 18, 30). The resemblance between this autocorrelation function and the theoretical one suggests that seasonality is present in this series.

Table 3.3 Autocorrelation and Partial autocorrelations – QDSOY

LAG	AC	PAC	Q	LAG	AC	PAC	Q
1	0.9012	0.9012	149.44	21	0.4813	0.0388	1698.9
2	0.8027	-0.0502	268.66	22	0.5237	0.0339	1756
3	0.7336	0.1022	368.8	23	0.5802	0.0155	1826.5
4	0.6675	-0.0255	452.17	24	0.6078	0.0859	1904.5
5	0.6165	0.0628	523.7	25	0.5522	-0.1563	1969.2
6	0.5916	0.1112	589.95	26	0.4779	-0.0044	2018.1
7	0.6036	0.2085	659.3	27	0.4155	-0.0152	2055.2
8	0.6303	0.1377	735.36	28	0.3681	0.0998	2084.5
9	0.6564	0.1049	818.32	29	0.3195	-0.0417	2106.8
10	0.7027	0.224	913.96	30	0.2939	0.0338	2125.7
11	0.7784	0.3466	1032	31	0.2978	-0.0055	2145.3
12	0.8041	0.0028	1158.7	32	0.3082	-0.0117	2166.4
13	0.7485	-0.275	1269.2	33	0.3319	0.0492	2191.1
14	0.6613	-0.2647	1355.9	34	0.3726	0.0139	2222.3
15	0.5895	-0.0592	1425.3	35	0.4235	-0.0182	2263
16	0.5305	0.0198	1481.7	36	0.4364	-0.0418	2306.5
17	0.4754	-0.054	1527.4	37	0.3777	-0.1645	2339.4
18	0.445	-0.0937	1567.6	38	0.3079	-0.0117	2361.3
19	0.4409	-0.1007	1607.4	39	0.2469	-0.056	2375.5
20	0.4599	0.0728	1650.9	40	0.1955	-0.0053	2384.5

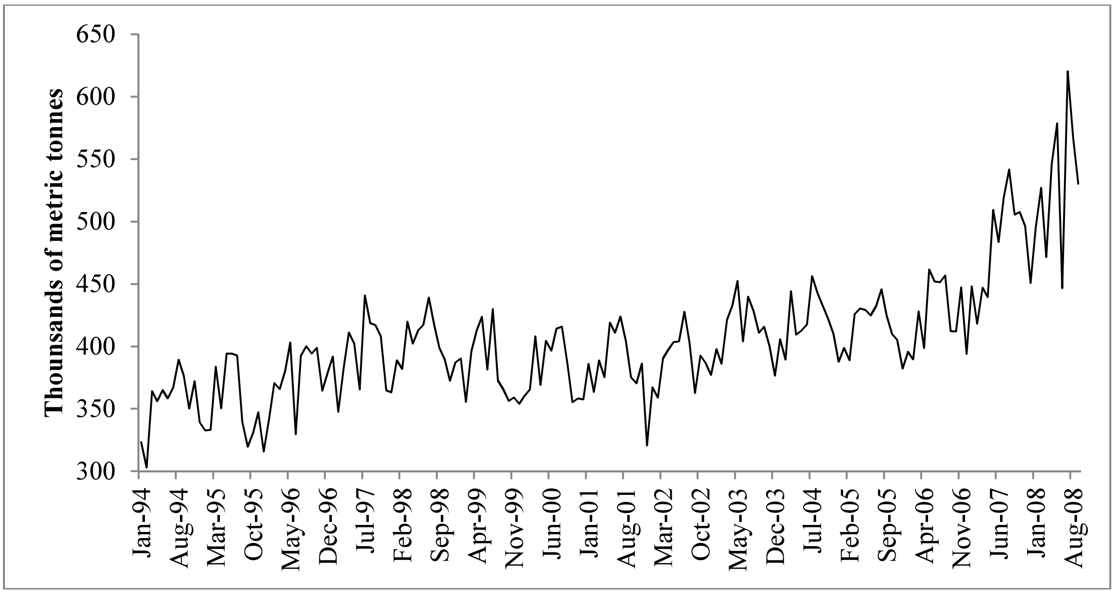
Source: Own calculations

Figure 3.3 Argentina's exports of maize 1994-2008 (in thousands of metric tonnes)



Source: Ministerio de Agricultura, Ganadería, Pesca y Alimentos de la Republica Argentina

Figure 3.4. Argentina's monthly domestic supply of maize 1994-2008 (in thousands of metric tonnes)



Source: Ministerio de Agricultura, Ganadería, Pesca y Alimentos de la Republica Argentina

Table 3.4. Autocorrelation and Partial autocorrelations - QEMAZ

LAG	AC	PAC	Q	LAG	AC	PAC	Q
1	0.6695	0.6695	80.688	21	-0.0385	-0.0806	324.51
2	0.3225	-0.2279	99.516	22	0.135	-0.0365	328.24
3	0.0389	-0.1408	99.791	23	0.3783	0.1945	357.69
4	-0.0922	0.0145	101.35	24	0.4915	0.1612	407.72
5	-0.1822	-0.1278	107.47	25	0.3391	-0.0482	431.69
6	-0.2212	-0.067	116.53	26	0.1051	-0.0434	434.01
7	-0.1836	0.0323	122.81	27	-0.0824	-0.0603	435.44
8	-0.1199	-0.0268	125.51	28	-0.1347	0.0867	439.3
9	-0.01	0.0801	125.53	29	-0.1637	0.0375	445.04
10	0.2422	0.3609	136.65	30	-0.1811	-0.0624	452.11
11	0.4963	0.2725	183.66	31	-0.1791	-0.0644	459.07
12	0.5997	0.1382	252.73	32	-0.1364	0.054	463.14
13	0.3957	-0.2135	282.98	33	-0.0703	-0.0377	464.23
14	0.122	-0.1195	285.87	34	0.1125	0.0376	467.03
15	-0.0848	-0.0323	287.28	35	0.3471	0.0767	493.91
16	-0.1689	0.0472	292.89	36	0.4559	0.0564	540.61
17	-0.2217	-0.0353	302.62	37	0.266	-0.1698	556.62
18	-0.2331	-0.0243	313.45	38	0.0444	0.0105	557.07
19	-0.1925	0.0233	320.88	39	-0.1173	-0.0486	560.23
20	-0.1286	-0.0229	324.21	40	-0.1274	0.0996	563.98

Source: Own calculations

In Figure 3.4 we can see the monthly domestic supply of maize from January 1994 to August 2008. In contrast to the other series seen so far, it is very hard to distinguish a clear deterministic seasonal pattern. It is not possible to identify clearly periods of intense or low activity that systematically repeat. At the same time, the autocorrelation and partial autocorrelation functions, presented in Table 3.5, do not exhibit the typical repetitive spikes every twelve months, rather a continuous but slow decay, typical of non-stationary processes. This implies that deterministic seasonality cannot characterise adequately or contribute to the explanation of the variation in this series.

Table 3.5. Autocorrelation and Partial autocorrelations – QDMAZ

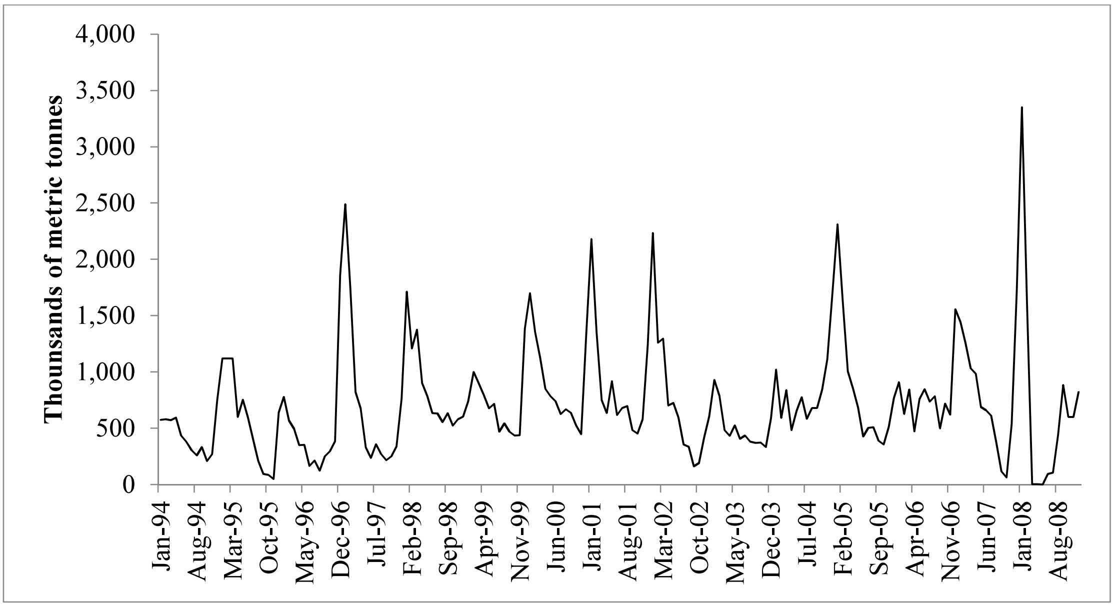
LAG	AC	PAC	Q	LAG	AC	PAC	Q
1	0.8703	0.8703	136.35	21	0.2212	0.0122	1101.8
2	0.8096	0.2151	255.01	22	0.213	0.0666	1111
3	0.7562	0.0691	359.14	23	0.2352	0.1139	1122.4
4	0.7268	0.1104	455.88	24	0.229	-0.059	1133.3
5	0.6848	-0.0027	542.25	25	0.1985	-0.0988	1141.5
6	0.6244	-0.0965	614.49	26	0.1558	-0.0857	1146.6
7	0.6017	0.0966	681.97	27	0.1195	-0.0872	1149.6
8	0.5723	0.0146	743.37	28	0.1024	-0.0427	1151.8
9	0.5336	-0.051	797.07	29	0.0527	-0.1072	1152.4
10	0.498	0.0044	844.13	30	0.0168	-0.0354	1152.5
11	0.489	0.0987	889.76	31	-0.0285	-0.0632	1152.6
12	0.4826	0.0432	934.47	32	-0.0431	0.0423	1153
13	0.422	-0.1918	968.88	33	-0.0759	-0.0315	1154.3
14	0.3697	-0.0809	995.44	34	-0.0883	0.0322	1156
15	0.3244	-0.0473	1016	35	-0.1033	-0.048	1158.4
16	0.3259	0.1338	1036.9	36	-0.107	0.058	1161
17	0.2938	-0.0254	1054	37	-0.1424	-0.0632	1165.6
18	0.2695	0.0268	1068.5	38	-0.1932	-0.0831	1174.1
19	0.2463	-0.0235	1080.7	39	-0.2253	-0.0621	1185.7
20	0.2352	0.0173	1091.8	40	-0.2339	0.0466	1198.4

Source: Own calculations

The monthly exports of wheat between January 1994 and August 2008 can be seen displayed in Figure 3.5. Although the seasonal pattern is not as clear as in the other export series, it is still possible to see some repetitive patterns with important supply moments during harvest (December-January). However, this pattern is not as distinctive as in the other cases where zeroes were observed in the rest of the seasons.

This seasonality is also present in the autocorrelation and partial autocorrelation functions where is possible to see single spikes at months 12, 24 and 36 (Table 3.6). Although in the rest of the months the autocorrelations are not zero, it is possible to observe a stressed seasonal pattern. This implies that deterministic seasonality might help to explain the variability in this series.

Figure 3.5. Argentina's monthly exports of wheat 1994-2008 (in thousands of metric tonnes)



Source: Ministerio de Agricultura, Ganadería, Pesca y Alimentos de la Republica Argentina

Table 3.6. Autocorrelation and Partial autocorrelations – QEWHT

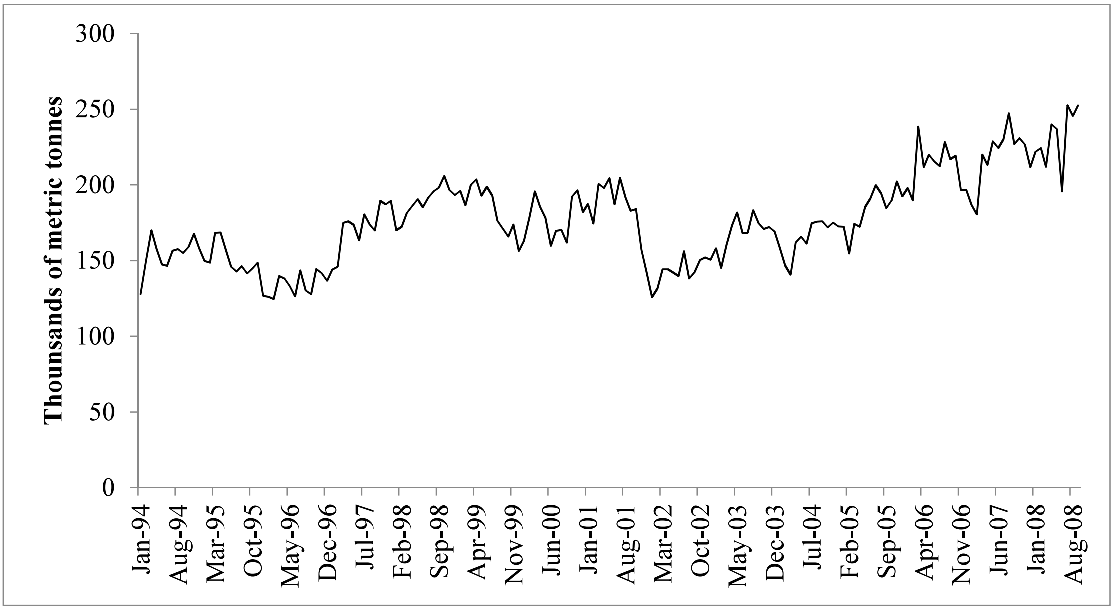
LAG	AC	PAC	Q	LAG	AC	PAC	Q
1	0.6668	0.6668	80.036	21	-0.0387	0.0808	219.19
2	0.2264	-0.3929	89.317	22	0.1028	0.0847	221.35
3	-0.0686	-0.036	90.174	23	0.2582	0.0525	235.07
4	-0.1917	-0.0502	96.903	24	0.3349	0.0439	258.3
5	-0.2354	-0.1177	107.11	25	0.2124	-0.1093	267.7
6	-0.2395	-0.076	117.74	26	0.028	0.0625	267.86
7	-0.1929	-0.0191	124.67	27	-0.1176	-0.0472	270.78
8	-0.1452	-0.0929	128.62	28	-0.1931	-0.0645	278.71
9	-0.0718	0.0261	129.6	29	-0.2117	-0.0099	288.31
10	0.0715	0.1387	130.56	30	-0.2319	-0.1289	299.9
11	0.2678	0.1902	144.26	31	-0.2134	-0.0301	309.78
12	0.3936	0.1048	174.01	32	-0.1556	0.0069	315.07
13	0.2813	-0.1626	189.29	33	-0.0502	0.0194	315.63
14	0.0274	-0.1097	189.44	34	0.1228	0.0711	318.97
15	-0.1275	0.0888	192.61	35	0.3315	0.1768	343.49
16	-0.1843	-0.0607	199.3	36	0.4429	0.0735	387.57
17	-0.1887	-0.0213	206.35	37	0.2877	-0.1613	406.3
18	-0.179	-0.0349	212.73	38	0.0468	0.0523	406.8
19	-0.1373	-0.008	216.51	39	-0.1135	0.0002	409.76
20	-0.1085	-0.0724	218.89	40	-0.1798	-0.0062	417.24

Source: Own calculations

As can be seen in Figure 3.6, the monthly domestic supply of wheat between January-1994 and August 2008 does not seem to present a clear seasonal pattern. Its behaviour seems to be represented better by a non-stationary process with a likely structural break around January 2002. Although the autocorrelation function depicted in Table 3.7

suggests some spikes every twelve months that resembles the theoretical autocorrelation function, it is very hard that deterministic seasonality might be an important explanation for the variation of this series.

Figure 3.6. Argentina's monthly domestic supply of wheat 1994-2008 (in thousands of metric tonnes)



Source: Ministerio de Agricultura, Ganadería, Pesca y Alimentos de la Republica Argentina

Table 3.7. Autocorrelation and Partial autocorrelations - QDWHT

LAG	AC	PAC	Q	LAG	AC	PAC	Q
1	0.7531	0.7531	102.1	21	0.1869	-0.022	774.35
2	0.7052	0.3188	192.13	22	0.2306	0.061	785.22
3	0.6247	0.0587	263.19	23	0.2422	0.0429	797.28
4	0.5376	-0.0491	316.12	24	0.2768	0.1169	813.15
5	0.4587	-0.05	354.88	25	0.2131	0.0023	822.62
6	0.454	0.1438	393.07	26	0.2197	0.0226	832.75
7	0.399	0.0089	422.74	27	0.1203	-0.0879	835.8
8	0.4504	0.2117	460.76	28	0.1102	0.0386	838.39
9	0.4761	0.147	503.5	29	0.0318	-0.1512	838.6
10	0.5107	0.1102	552.98	30	0.0103	-0.0954	838.63
11	0.5097	-0.0013	602.57	31	-0.0021	0.0302	838.63
12	0.5352	0.0551	657.58	32	0.0616	0.0588	839.46
13	0.4035	-0.2911	689.03	33	0.0795	0.0921	840.85
14	0.401	0.0197	720.28	34	0.1664	0.1089	846.98
15	0.2929	-0.1345	737.05	35	0.1659	-0.0325	853.12
16	0.2172	-0.0878	746.34	36	0.2185	0.0624	863.85
17	0.1562	-0.0567	751.17	37	0.1737	-0.0515	870.67
18	0.1697	0.0861	756.91	38	0.1519	-0.0892	875.93
19	0.1198	-0.0498	759.79	39	0.0836	0.0227	877.54
20	0.1923	0.0925	767.25	40	0.0675	0.022	878.59

Source: Own calculations

The visual inspection of the series and the analysis of the autocorrelation and partial autocorrelation function, allows us to identify some patterns of deterministic seasonality, particularly in the export series. This might be associated with the lack of storage facilities and the availability of vessels that make it imperative to export at the moment of harvest in the southern hemisphere. This is not the case in the domestic supply series where the supply can be spread more easily over different periods, suggesting a less stressed pattern of seasonality.

In addition to the graphical evidence and the analysis of the autocorrelation and partial autocorrelations, it is possible to characterise the seasonality as a deterministic pattern and evaluate its relevance in terms of how well they represent the data. Table 3.8 presents a deterministic approach to capture the seasonality in the series presented. The approach is simple and implies the definition of monthly dichotomic variables to capture a deterministic seasonal pattern. This is the approach followed by Mugambe and Reilly (2007). In addition, we present the Breusch-Godfrey indicator of serial autocorrelation on the residues, an F-type test on the joint significance of the explanatory variables and the R-squared.

Table 3.8. Quantities supplied – Deterministic approach

	qesoy	qdsoy	qemaz	qdmaz	qewht	qdwht
Jan	42.6	-139.8	-46.2	-1.4	528.0	22.6
	0.29	*0.02*	*0.27*	*0.57*	*0.00*	*0.00*
Feb	-48.0	-341.4	56.1	-2.2	-454.9	-9.7
	0.19	*0.00*	*0.40*	*0.44*	*0.00*	*0.06*
Mar	56.3	-3.1	867.3	17.9	-267.8	30.6
	0.07	*0.98*	*0.00*	*0.00*	*0.03*	*0.00*
Apr	860.0	614.9	380.2	0.6	-222.2	-1.8
	0.00	*0.00*	*0.01*	*0.86*	*0.00*	*0.82*
May	331.1	359.5	-134.1	0.5	-16.8	24.1
	0.00	*0.00*	*0.02*	*0.84*	*0.73*	*0.00*
Jun	-292.4	-125.6	-498.3	-8.7	-79.1	-25.8
	0.00	*0.20*	*0.00*	*0.00*	*0.02*	*0.01*
Jul	-99.2	128.7	-59.9	11.2	-58.7	37.4
	0.31	*0.23*	*0.69*	*0.00*	*0.06*	*0.00*
Aug	-228.4	-41.7	-142.9	-1.4	-12.5	-13.6
	0.03	*0.36*	*0.29*	*0.63*	*0.71*	*0.04*
Sep	-224.0	54.4	-173.8	-4.9	-40.8	-22.0
	0.00	*0.31*	*0.00*	*0.01*	*0.38*	*0.00*
Oct	-192.1	-108.0	-58.3	6.4	-0.5	-4.5
	0.00	*0.02*	*0.35*	*0.02*	*0.99*	*0.28*
Nov	-86.6	-156.0	1.7	-1.8	92.4	-3.5
	0.05	*0.00*	*0.98*	*0.56*	*0.02*	*0.56*
Dec	-85.6	-134.0	-161.6	-8.3	637.3	-20.5
	0.12	*0.03*	*0.01*	*0.00*	*0.00*	*0.00*
B-G	*52.7*	*49.2*	*47.8*	*31.8*	*29.7*	*88.8*
F-statistic	*14.3*	*12.4*	*12.4*	*6.5*	*14.1*	*9.2*
R-squared	*0.47*	*0.43*	*0.44*	*0.27*	*0.47*	*0.36*

Note: Newey-West estimators with 12 lags.

Source: Own calculations

The intention of this exercise is to try to measure how good the deterministic approach is in explaining the variation observed in the dependent variable. Of course, there are additional elements that can explain their behaviour such as trends, cycles and breaks, however, the effort is to analyse how much of the variation can be characterised with a deterministic approach.

In all the export variables, the R-squared tends to be above 0.40. This is quite high considering that only these dichotomic variables have been included. This suggests that the deterministic seasonality could explain nearly 40% of the variation observed in the dependent variable. The explanatory power of this approach is lower for the domestic supplies (except soybeans) but still can characterise more than 27% of the variation in the case of the domestic supply of maize.

At the same time, the high BG statistic rejects the null of no autocorrelation in the residuals. This suggests that the residuals are not behaving according to the theoretical model and, in addition, explains the low explanatory power of the model. This implies that, as expected, there is still explanatory power in the residuals that could explain the variation of the dependent variable.

High significance in the individual variables indicates that a particular month is of particular importance in the explanation of the variation. In this sense, in April exports of soybeans tend to be around 860 thousand tonnes higher than the average. Note that the months that tend to be significant are those that lead to the peak of the supply and when the peak has been passed. Therefore, there is a combination of positive and negative coefficients before and after. For example, in the case of the export of soybeans, it can be seen that March, April and May presents high and significant positive coefficients, whilst the following months also present high and significant but negative ones. This pattern, however, is less pronounced in the domestic supplied variables where the seasonality is not so stressed.

In addition, the joint significance test rejects the null that all coefficients are simultaneously equal to zero, suggesting that the deterministic model, with the probable exclusion of some monthly dummy variables, can provide a good initial base for the explanation of the variability of the dependent variables. However, as we have mentioned, other elements need to be added to complement it.

3.4.2. Testing seasonal unit roots

A deterministic approach might provide a sufficient explanation for the seasonality observed in the series. However, the possibility that unit roots might be present in the series cannot be discarded easily from the inspection of the plots and the analysis of the ACFs and PACFs. Therefore, in this section, we perform the HEGY unit root test on the six series presented.

The series may be seen particularly short for this type of analysis (around 180 months) given the number of parameters that must be estimated from them. The power of the tests may be affected and that may lead to incorrect conclusions on the presence of seasonal unit roots. If longer series are available, they should be used. At the time of the writing this chapter, the latest information available has been used. Unfortunately, we could not find consistent older series to extend the sample. Where some older series existed, these were collected by different bodies and using different methodologies. This means that additional noise, related to measurement error or with different ways of measurement, may be introduced that can lead potentially to incorrect conclusions. This noise will be added to the structural breaks on the series (changes on regimes, legislation, technological change, etc.) as we have discussed previously.

In fact, using longer series adds the additional problem of potentially many structural breaks. This is particularly true in cases of developing countries where legislation and institutions tend to be more volatile. In the case of Argentina, extending the series further to the past will include important changes in regime such as the stabilization programme of 1991 and, more importantly, the liberalisation of the trade and commercialisation of grains after the dissolution of the previous grains board (Junta Nacional de Granos). As we have seen, since the effect of grains boards may notably reduce the seasonality in the commercialisation of grains, including data under this regime may change the nature of the analysis. The data segments created by this particular regime may be seen as two different data processes that should not be taken altogether; at least for the seasonal analysis.

On the other hand, a close inspection of other applications reveals that the length used in this analysis is not particularly short. Whilst Beaulieu and Miron (1993) use a sample of almost 240 months of series of real wages in the US, Franses (1991) uses just 120 months of new car registrations in the Netherlands, and only nine years of monthly data on the number of airline passengers.

Studies on quarterly series use samples with fewer observations but the model estimates fewer parameters as well. Franses and Volgelsang (1998) use only 40 observations of quarterly GDP in the Netherlands. In the context of monthly data, this would be equivalent to 160 observations. Moreover, the same paper also uses this series to test for seasonal unit roots under the presence of unknown structural breaks that will require even more parameters to be estimated, as we will see below.

However, it is fair to say that studies have highlighted that the power of tests may be more related to the span rather than the number of observations (Shiller and Perron, 1985). In this sense, tests based on samples of, for example, 20 years of quarterly data may have more power than based on 10 years of monthly data. However, Haug (2002) shows that the increase in the power associated with the use of a longer sample when series are temporally aggregated may be offset by the reduction in the number of observations, particularly in finite samples. Therefore, we recognize the limitations on the inference made on a limited number of observations and caution must be given when extracting conclusions. However, it is also recognised that this application does not depart from the usual practice, and the extension on the sample to the past may not be convenient.

In one of the series under study, the quantity of exports of soybeans, we can see that around the end of each year, there are no exports of this commodity, pointing to the case of zero values in the series we are reporting in this chapter. Moreover, in the quantity of exports of maize and wheat, despite not having zero values, does have extremely low values in some seasons compared with the peaks observed in the rest of the seasons. This application can be extended to other cases where data present similar characteristics.

In Table 3.9, we present the results of the tests of seasonal unit roots on the six series considered. Columns 2 to 8 give the values of the tests statistics. In columns 9 and 10, the chosen specification is detailed in terms of lag structure and the presence of deterministic elements of the equation. The precise specification was determined by the statistical significance of intercept and trend terms in the regression equation. However, dummy variables were always included in the regressions to capture the deterministic pattern of the seasons.

Given its popularity, statistical properties and convenience, the maximum lag length was determined using the general to specific approach explained above. We have also included and excluded some lags according to the 10% level of significance (within this maximum limit set by the general to specific approach) to help to obtain well behaved residuals. This is because using the above-mentioned approach for maximum lag selection (including all lags until lag ρ-1) did not yield the desired properties on the residuals as well as including non-significant parameters in the model that could reduce the power of the tests. Therefore, a mixed approach has been followed by combining the maximum lag length selection using the general to specific approach and some testing on the residuals' properties to determine which lags should be included within the maximum selected. The last two columns give the statistics for the Breusch-Godfrey test for serial correlation and the Bera-Jarque test for normality of the residuals.

Table 3.9 Results of test for seasonal unit roots in monthly series

	0 π_1	π π_2	$\pi/2$ $F_{3,4}$	$2\pi/3$ $F_{5,6}$	$\pi/3$ $F_{7,8}$	$5\pi/6$ $F_{9,10}$	$\pi/6$ $F_{11,12}$	Lags	Deterministic elements	B-G	B-J
qesoy	-2.23	-4.89	16.95	8.87	9.41	9.27	0.93	6,12,15	T	16.5	1.80
qdsoy	-1.80	-3.73	7.87	8.48	11.40	6.51	9.33	19,20,36	T,C	15.1	2.05
qemaz	-1.49	-3.39	22.61	18.06	17.28	13.35	3.24	29,30		6.3	2.33
qdmaz	3.20	-5.69	9.94	13.98	16.17	14.17	9.58	11,12		5.4	0.75
qewht	-1.41	-5.32	18.61	22.49	30.76	16.61	25.71	12,13	I,C	12.3	4.76
qdwht	1.99	-5.67	19.2	7.97	10.84	7.01	7.59	3,4,15,33		18.3	5.02

Notes:
I) B-G refers to the Breusch-Godfrey test for serial autocorrelation on the residuals with 12 lags. Critical value 21.09
II) B-J refers to the Bera-Jarque normality test on the residuals. Critical value 5.99
III) Deterministic elements identifies if a trend (T), an intercept (I) or the conflict (C) dummy has been included in the testing equation. Conflict is a dummy variable reflecting the conflict between the Government and farmers between March 2008 and August 2008.

Source: Own estimations

By selecting the appropriate critical values, according to the specification chosen, we can conduct the tests of unit roots at the zero and at the rest of the frequencies. Given that the critical values obtained in our simulation exercise, considering zero values in the DGP, do not differ from those found in the literature, it is indistinctive to use the one found here or those already tabulated in the literature, even though only some series present zero values.

In none of the series is it possible to reject a unit root at the zero frequency (or a unit root *a la Dickey-Fuller*) at the conventional levels of significance, which suggests that the series may possess a stochastic trend. Unit roots at frequency π can be rejected in all the cases. For the rest of the frequencies, it is possible to reject unit roots in all series with the exception of the quantity of exported soybeans and the quantity of exported maize at the bi-annual frequency ($\pi/6$). Using a 97.5% level of confidence, we cannot reject a unit root at the four-monthly frequency ($5\pi/6$) in the quantity of domestic supply wheat and in the quantity of domestic supply of soybeans. The residuals seem to present the desired properties. Therefore, in addition to the long run unit root present in all series, some series seem to be affected by seasonal unit roots or stochastic seasonality.

This suggests that the stochastic elements may have permanent effects on the seasonal pattern of some of the series analysed and they cannot be considered as part of the general stochastic component of the series. These stochastic elements do not die out fast enough and their effects are transmitted to future values of the seasons. Therefore, for some of the series, there seems to be some mild evidence that suggests that they might require a treatment beyond the deterministic approach. Weather, market and other technological effects may be affecting the pattern of seasonality in some of the series analysed in a permanent way. However, this mild evidence and the fact that the deterministic approach provided a good explanation for the variation of the dependent variable, do not allow us to extract emphatic conclusions about the general presence of stochastic seasonality. More and deeper analysis is required, especially on the possibility that some of the unit roots found might have been confused with structural breaks. This will be analysed in the following sections.

3.5. STRUCTURAL BREAKS AND SEASONAL UNIT ROOTS

As we have discussed, structural breaks may also be present in agricultural series. These structural breaks may not only have general effects on the levels and trends of the series, but also they may change the observed pattern of seasonality. Rumours on changes and the changes themselves in legislation and the commercialisation regime, for example, may introduce conjunctural effects in the series. Generally, in these cases, agents may speed up or postpone their decisions following a “wait and see” strategy, affecting the pattern of seasonality. However, it is convenient to distinguish between a permanent change in the pattern (which eventually must be included in the deterministic component) and an innovation that introduces some noise in the pattern. Since it is hard to identify the effect of these structural changes on the series, there exists the possibility that a seasonal unit root found might be explained by this change rather than stochastic seasonality *per se*.

As in the standard non-seasonal unit root tests, if structural breaks are present in the series, the HEGY approach tends to find too many unit roots, or it suffers from low power. In other words, structural breaks could be disguising otherwise stationary processes as containing unit root processes. This calls us to consider the possibility of breaks when testing for seasonal unit roots.

Originally, this potential problem of the unit roots was considered in the non-seasonal case. If the date of the break is known beforehand in the non-seasonal case, the adjustment necessary is relatively straightforward. It entails the addition of dummies to capture the different segments in which the series is divided (before and after breaks). Critical values for the tests are non-standard and they have been tabulated by Perron (1989). It is important to remark that these tests are not intended to ascertain the presence of a non-zero drift *per se*, since in both null and alternative hypothesis the process is assumed to possess a structural break. When the objective is to determine the presence of a structural break *per se*, the procedure outlined by Chow (1960) can do the job. However, the technique suggested by Perron (1989) has its focus on testing for unit roots under the presence of structural breaks and not for the presence of the structural break itself.

Things are a little more complicated if the researcher does not know where (or when) the break is located. In this case, rather than being exogenous information, like a change in regime or an independent event whose location is known with certainty, the date of the break is unknown and must be estimated. However, this does not mean that the process originates the break in the sense that its appearance can be modelled or be part of the deterministic component of the series. On the contrary, the break is still exogenous (affecting the process) but its location in the series becomes part of the estimation process.

This problem was addressed initially by Perron and Vogelsang (1992), Zivot and Andrews (1992) and Banerjee, Lumsdaine and Stock (1992) in a non-seasonal framework. Two models have been developed to consider how a break could affect a given series. The additive outlier (AO) model allows for an instantaneous shift in the intercept of the deterministic trend of the series. The innovative outlier (IO), on the other hand, allows changes in the series to have a gradual effect.

In the first model, the effect of the change on the level of the series is not affected by the dynamics of the correlation structure of the series. In the second model, it is assumed that the series reacts to a change in the mean in the same way that it responds to other shocks. This implies that there is a transition period in the adjustment of the series. Operationally, the main difference between both procedures is that in the IO the estimation is conducted on a single equation, while the AO requires an auxiliary regression. In addition to the simplification on the estimation, the type of phenomena that could affect agricultural production (new techniques of production, for example) has a gradual effect (the implementation period) on the series that can be more accurately modelled using the IO specification rather than the AO specification. Therefore, we will focus our attention on this latter procedure. For further details about the AO model, see Harris and Sollis (2003).

The procedure for testing unit roots when the date of the break is unknown requires the estimation of the regression equation considering all the potential break dates. Therefore, the model must be estimated as many times as potential breaks are considered. In essence, the procedure lies in using the Perron (1989) test as many times as breaks are considered. The main point lies in selecting the appropriate break time among all possible dates. In addition, the statistics resulting from that particular specification are eventually those considered to perform the test.

Extensions to the seasonal case using the HEGY approach when the time of the break is unknown have been applied on quarterly data by Franses and Volgelsang (1998) using GDP data on several countries, whilst Ghysels and Osborn (2001) have followed a more theoretical approach analysing the properties of the procedure. A more technical discussion of the procedure is contained in Harris and Sollis (2003). Additional extensions can be found in Balcombe (1999) on US quarterly food prices indexes and different US quarterly macroeconomic aggregates. Nevertheless, the application of the procedure on monthly data has been particularly scarce.

If we assume that there is a single break that occurs at time T_b (being $1<T_b < T$), it is possible to test the null of an IO break in monthly data by estimating the following equation

$$\begin{aligned}\Delta_{12}y_t = \mu + \beta t \\ &+ \sum_{k=2}^{12} \delta_k D_{kt} \\ &+ \sum_{k=1}^{12} \pi_k Z_{k,t-1} + \sum_{i=1}^{\rho-1} \psi_i \Delta_{12} y_{t-i} + \sum_{k=2}^{12} \theta_k S_{kt} + \sum_{k=2}^{12} \eta_k \Delta_{12} S_{kt} + v_t\end{aligned} \tag{3.12}$$

where

$$S_{kt} = \begin{cases} 1 & (t > T_b) \\ 0 & (t \geq T_b) \end{cases} \qquad k = 1, \dots, 12$$

Therefore, S_{kt} is a standard seasonal dummy that starts to be active at the time of the break. Similarly to the non-seasonal case, the procedure requires us to fit the equation above to all the potential break-dates in data. It is advisable to restrict the range of possible breaks to $T_b^*, \dots, T - T_b^*$, where $T_b^* = \lambda T$. The value of λ is called the amount of trimming, and it excludes from the potential break dates some observations at the beginning and end of the series. This is done in order to assure that the results are asymptotically valid as recommended by Franses and Volgelsang (1998). Moreover, if a series has a break either at the beginning or end of the series, it would make very little sense to consider the break, or it might be advisable to exclude those observations (those before or after the break) from the calculations.

It is important to highlight that this procedure allows for testing for the presence of seasonal unit roots under the presence of a single unknown structural break. This means that this procedure will identify, probably, the most important (in terms of its effects on the series) of the present structural breaks whilst still leaving the influence of other structural breaks that may be present in the series. Whilst Tasseven (2008) developed a procedure for testing for the presence of two structural breaks, this approach assumes the knowledge of the place or time of both breaks. The possible presence of more than one structural break is an additional warning and recommendation against the use of very long time series.

In order to select the break date, there are two approaches. The first method involves minimizing the $t_{\pi i}(T_b)$ and maximizing the $F_{odd,even}(T_b)$ statistics over all possible break dates, or select the break when the statistics are least favourable to the null hypothesis. Note that we are not selecting break dates but selecting statistics values for the unit root tests given all the possible breaks in the series. We can define this method by

$$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b) \qquad i = 1,2 \qquad (3.13)$$

$$\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b) \qquad (3.14)$$

The use of this criterion will identify as many break dates as unit roots are considered, given that for each frequency, the selection is based on the statistic that is least favourable to the null hypothesis. This is precisely the result that Franses and Volgelsang (1998) obtained when using this criterion. In their application, breaks occur in different periods depending on the frequency analysed and the series.

This is not suggesting that there are multiple breaks affecting the series, but that the manifestation of a single break is captured in different periods depending on the seasons of the series. It should be recalled that the specification of the IO model precisely captures the gradual effect of a structural break. At the end, the break date in this criterion is selected by identifying the unit root test's statistic more favourable to the rejection of the null hypothesis and not by the statistical significance of the break. Moreover, the tests are designed to determine if, eventually, the seasonal unit root tests performed before have been affected by the presence of a structural break. This suggests that the use of this procedure to identify unknown structural breaks in a more general context may not be appropriate.

The second method addresses this issue by selecting the break date, based on the maximization of the significance of the seasonal shift dummy variable or

$$\hat{T}_b = argmax\ F_{\theta}(T_b) \tag{3.15}$$

In this case, a unique break date will be identified and the trimming of the series is necessary. In essence, the trimming on the data is not necessary if the first method (equations 3.13 and 3.14) is employed. It has been shown by Perron and Vogelsang (1992) that this second method has more power than the first method and its use has been widely recommended. However, their recommendations are based on its statistical properties and not on the grounds of its capability to select breaks dates.

The use of this second criterion for the selection of the break date entails an additional complication. When using the second method for identifying the break, equation (3.15), and assuming an IO model, Harvey, Leybourne and Newbold (2001) found that in the non-seasonal case, there is a tendency to anticipate the break by one period. This is exacerbated particularly if the size of the break is particularly large. Alternatively, basically, the *t*-statistic associated with the shift dummy variable has a distribution whose mean is maximised at $\hat{T}_b - 1$. On the other hand, using quarterly data with seasonality, Harvey, Leybourne and Newbold (2002) find that the rule tends to anticipate the break by four quarters ($\hat{T}_b - 4$), and they suggest adjusting the second method by

$$\hat{T}_b = 4 + argmax\ F_{\theta}(T_b) \tag{3.16}$$

They show that the incorrect selection of the break date has statistical implications. Not only will the tests be done with incorrect statistical values, but also they will increase the test size, leading to a spurious rejection of the null hypothesis.

No applications of seasonal unit root test with breaks applied to monthly data have been found that have verified this misplacement. One would expect that in this case, and following a logical inferential process, the rule given by equation (3.15) should anticipate the break by around twelve months in the monthly case (a modification of the rule established by equation (3.16)). We will return to these aspects below.

3.5.1. Zero values and their effects on critical values of seasonal unit roots under the presence of structural breaks

As for the non-seasonal case, the statistics for testing for the presence of unit roots when breaks are allowed have specific distributions. Franses and Hobijn (1997) and Smith and Otero (1997) have tabulated the critical values for both types of methods to select the break on quarterly data. However, no critical values are available for monthly data for any type of data, and not just for the case of series that contain zeros in their domain.

This presents a practical problem not only for the exercise we are carrying out on seasonal unit roots on agricultural commodities, but also for any other exercise where monthly data are used. Since we have seen that zero values do not affect the critical values of the HEGY test without breaks, the tabulation of critical values for the monthly case, even when the variable can observe zero values, might be an avenue for future research work in this or other fields.

In order to solve this lack of reference values, we have run some Monte Carlo experiments to obtain appropriate critical values for monthly data when breaks are allowed, and we have considered the possibility of zero values in the seasons. The DGP is similar to the one presented before when we introduced the critical values without the structural break. However, it contains a shift in the level of the series. While the time of the break was fixed in the DGP, we allowed the procedure to select the moment of the break by using the two methods highlighted in equations (3.13), (3.14) and (3.15).

We consider two possible sizes for the break. In one case, the size of the break follows a standard normal distribution and in the second case, the size of the break was given exogenously and subjectively a value of 10 that can capture what would be a large break, given the level of the values. The idea is to verify in this latter case (in the monthly case), the anticipation in the break selection rule found by Harvey, Leybourne and Newbold (2002). The simulations were performed on a sample size of 200 observations for different specifications of the model in terms of the presence of intercept, trends and deterministic seasonal dummies. For each simulation, an equation similar to (3.12) is fitted considering all the potential break dates in the sample, excluding the trimming (λ=0.1 was used). Given computational times, simulations were limited to 10,000. With 200 observations and a trim of 0.1, it is necessary to run 1.6 millions of regressions for each model specification. In a very new and powerful computer, the seven model specifications took more than four days working full time to be run.

Table II.2 in Appendix II presents the critical values when the monthly DGP contains a structural break. The size of the break follows a standard normal distribution and zero values are considered. In the top panel, we can see the values when the break selection criterion is based on the minimization of the values of t and the maximization of the values of F (equations (2.3) and (2.4)). In general, when structural breaks are allowed, the critical values for t are shifted to the left while the values of the F are shifted to the right (favouring the non-rejection of the null hypothesis). Moreover, the values for testing for unit roots at the zero frequency are not substantially different from those found by Perron and Vogelsang (1992) in the non-seasonal case. A result already found by Franses and Volgelsang (1998) in the quarterly case. This implies that the criterion for evaluating unit roots at the zero frequency when unknown breaks are considered, is not substantially affected by the seasonality as Ghysels, Lee and Noh (1994) also have found on quarterly data contexts. In the bottom panel, we present the critical values when the break is selected by maximizing the significance of the seasonal break dummy (equation (3.15)). This second criterion is, in general, preferred in terms of its size and power properties as suggested by Harris and Sollis (2003).

Given the lack of literature using monthly data, we can only compare these values with the quarterly data results. Using quarterly data and a substantially longer sample, (Franses and Volgelsang, 1998, p. 23) found almost identical critical values. They have found slightly larger values compared to the ones found here, favouring the non-rejection of the null. This is explained, as was mentioned, by the different sample sizes. However, they have only considered one specification (deterministic seasonal dummies, no trend and no intercept), whilst we have considered a more complete range of model specifications.

It is interesting to see that Franses and Volgelsang's (1998) critical values (as well as those found in this chapter) also differ from those found by Franses and Hobijn (1997) in quarterly data. Nevertheless, it must be stressed that the latter has set their experiments with specified breaks (or known breaks), and the model was not allowed to choose the break date. Therefore, their comparison cannot be made directly since it has been seen that, even in non-seasonal cases, critical values differ depending on the knowledge on the location of the break.

Therefore, it seems that unknown breaks in monthly data when zero values are considered, do not affect the critical values to be used in the testing. If we considered that this was also the result obtained when we considered the critical values without a structural break, it can be said that the presence of zero values does not invalidate the critical values generally used in the HEGY test, with or without breaks. This implies that the critical values obtained here using monthly data, when unknown structural breaks are considered, can be used without problem in any general exercise.

3.5.2. The misplacement in the identification of the break using monthly data

As mentioned previously, we have also considered the possibility of a large break to see how sensitive the critical values are, and to see when the model is selecting the break in order to determine if an adjustment is necessary in the selection rule of the second criterion. Table II.3 in Appendix II presents the results. The critical values for the unit root at the zero frequency are not affected but the F values are substantially shifted to the right. Using quarterly data, Harvey, Leybourne and Newbold (2002) suggest that the HEGY procedure could be severely over-sized when there is a large break and that the reason lies in the incorrect selection of the break date. They suggest that the maximization of the significance of the break criterion (equation (3.15)), as mentioned, tends to anticipate the time of the break by four quarters and effectively, they corrected for this problem (equation (3.16)).

Our results using monthly data, shown in Table II.4 in Appendix II, suggest that, despite the distribution being left skewed and, effectively, anticipating the time of the shock around 10 months before the break, it will be extremely unadvisable to suggest, as they did, a modification to the selection rule. The reason is that it was not possible to observe a clear and unambiguous pattern for the anticipation of the break date as they have found.

Although the model selects the break 10 months before around 20% of the time (according to the specification), this is very low compared with the percentage of times those authors have found using quarterly data (around 80% of the time). It is important to bear in mind that this analysis is using a short sample (about half the size they used) and, furthermore, their findings, in terms of the anticipation of the break date, may not be evident in this case. However, it seems that, effectively, the selection criterion tends to anticipate the time of the shock (without a clear pattern) and, eventually, they might generate more problems of size and power of the tests as these authors have suggested if the correction is not done. More evidence is required, by using a larger sample, to shed some light on this issue.

However, there is the possibility that the DGP used in this analysis and the one used in Harvey, Leybourne and Newbold (2002) might not be compatible. The form that the DGP has taken in their Monte Carlo exercise has not been specified, so there is a possibility that the processes used in this and their paper might not be comparable. On the other hand, it might be that the sizes of the breaks differ between both approaches. Since the misplacement phenomenon observed by those authors is particularly acute when the shock is high, their effect and manifestation will depend on the size of the break. It may be possible that the break considered here was not strong enough to verify it.

Given the short sample used and the possibility of low power in the tests, we have also considered the possibility of modelling the structural break using a single identification dummy in our Monte-Carlo exercise. In essence, we will be capturing the possibility of a break in the level of the series and not in the seasonal pattern. However, it might be useful since, when using short samples, the tests may have more power and be more reliable. Therefore, when there is not *a priori* evidence that the seasonal pattern might have changed, (i.e., production processes affected by natural seasons) it might be convenient to use this approach if samples are short.

Table II.5 in Appendix II presents the results of this exercise. It can be seen that the critical values for rejection of unit roots at the zero frequency do not differ substantially from those found when equation (3.11) is fitted. Moreover, the criterion for selecting the break, either equations (3.13)-(3.14) or equation (3.15), does not substantially affect the critical values: they are similar to each other. This suggests that the gain in terms of size and power that the use of equation (3.15) seems to yield is not as big when we use this approach.

We have also simulated critical values when a large shock is considered and a single dummy is used to identify the break. The results can be seen in Table II.6 in Appendix II. In general, critical values tend to be larger when a large shock is considered. However, when using equation (3.15) to identify the break date, the critical values are not displaced too much, compared to the case when minimizing t and maximizing F values to identify the break. These large values could bias the conclusions of the tests on unit roots towards the non-rejection of the null; this could be evidencing some lack of power when using the first criterion. Therefore, it may be advisable, as the literature has suggested, to base conclusions on the second criterion (i.e. maximizing the significance of the break dummy).

The misplacement of the break date has also been analysed using a single dummy for identification of the break. In the context of non-seasonal data, Harvey, Leybourne and Newbold (2001) found that using the IO model when the break is selected using equation (3.15), anticipates the break date by one observation. In our application, as seen in Table II.7 in Appendix II, we have found a similar result to the one found by those authors. Given the form of the DGP used, a $t+2$ is equivalent to the $t+1$ in Harvey, Leybourne and Newbold (2001). However, rather than anticipating the break, our model places the break in a later period. This may arise given differences in how the DGP has been specified, however, the results tend to support those already found. Therefore, when using this criterion, it is necessary to adjust the selection of the break date to consider this fact.

The misplacement of the break using a single dummy variable for identification is verified using seasonal data, although after the break rather than before. This leads us to suggest that this misplacement is the result of the way structural breaks are considered in the HEGY test under unknown structural breaks and not the result of seasonality *per se*. Non-seasonal and seasonal data (used in this test) verified this misplacement by one period when a single dummy variable is used. The fact that this misplacement has not been verified before (using the complete set of dummy variables to identify the break) requires further research.

3.6. EVIDENCE OF THE EFFECTS OF STRUCTURAL BREAKS IN SEASONAL UNIT ROOTS IN AGRICULTURAL COMMODITIES

Using the data presented above, and considering only those series for which we could not reject seasonal unit roots, we have applied the HEGY approach when breaks are considered. This is explained by the fact, as discussed, that the HEGY test under structural breaks can only confirm if the unit roots found previously are effective or have been the result of a structural break.

With the help of the critical values obtained before, we test for the presence of seasonal unit roots when breaks exist in the series. From the inspection of the series, we can confidently see that, if breaks exist, they tend to be not particularly large. We cannot observe that series tend to jump to extremely high or low values. Therefore, the set of critical values we will use are the ones that consider a break size that follows a standard normal distribution (or with one standard deviation). If we have evidence that suggests that the break could be large, these critical values could lead to a spurious rejection of the null. In that case, it may be appropriate to use the critical values for large breaks. However, in this case the date selection correction as explained above should be considered.

The number of observations and the length of the data in the estimation of the model with several parameters are problematic since the power of the tests is reduced. We have already discussed the number of observations in the context of the application presented here. This means that the implications on this application considering unknown structural breaks are even more severe.

Table 3.10 presents the results. In the left panel, we present the results when the break selection is based on the least favourable to the null hypothesis (minimising t and maximising F values). Under each statistical value, we present the particular time of the break found. Therefore, we have obtained different break dates, each associated with a particular root. The model specifications (in terms of inclusion of deterministic dummies, trends and intercept) are the same as the ones used in Table 3.9. Therefore, only a break is considered and no other elements in the specifications of the model. We still cannot reject the null of a unit root at a zero frequency in any of the three series. At a 95% confidence, we cannot reject the null of unit root at the bi-annual frequency in the quantity of exported soybeans and in the quantity of exported maize. However, we can now reject the unit root at the six monthly frequencies in the quantity of domestic supply of wheat. This suggests that the result found in the previous exercise without structural breaks, in this frequency on wheat has been the result of the presence of a structural break and not a seasonal unit root.

In the right panel of Table 3.10 we present the results when the break time is selected by maximising the significance of the seasonal break dummies[29] (equation (3.15)). In this case, the selection process can identify a single break period. This can be found at the bottom of the right panel. Again, we cannot reject unit roots at the zero frequencies in any of the series. The conclusions in terms of the unit roots for the quantity of exported soybeans (qesoy) and the quantity of exported maize (qemaz) at the bi-annual frequency remain unchanged. However, when using this criterion for the selection of the break, is reached a different conclusion on the six-month frequency in the case of the domestic supply of wheat. In the case of the first criterion, we have rejected a unit root, whilst here we are confirming its presence. Since this second method is generally preferred, given its power properties, this would be the conclusion of our test.

The fact that, when structural breaks are considered, new seasonal unit roots seem to appear in the series seems problematic. For example, using the second criterion for selecting the break, a unit root seems to be present in the quantity of domestic wheat at the quarterly frequency ($\pi/3$). Therefore, rather than helping to confirm results, considering structural breaks in series seems to complicate our judgement. However, it is important to remember that this test will have less power than the test without the structural break. More parameters are estimated using the same length of data or number of observations. Therefore, it is possible that new seasonal unit roots will appear.

[29] For this method, a trimming factor (λ) of 0.1 was used.

Table 3.10 Test of seasonal unit roots in monthly series under the presence of unknown structural break

Frequency	Statistic	$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b)$ $i = 1,2$ $\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b)$			$\hat{T}_b = argmax\, F_\theta(T_b)$		
		qesoy	qemaz	qdwht	qesoy	qemaz	qdwht
0	π_1	-2.73 *Aug-06*	-2.44 *Apr-96*	-1.56 *Jun-05*	-2.07	-2.17	-1.72
π	π_2	-5.81 *Apr-05*	-3.58 *Mar-05*	-6.66 *Jun-06*	-4.35	-2.90	-6.24
π/2	$F_{3,4}$	25.26 *Oct-98*	27.98 *Oct-02*	16.48 *Jul-06*	14.88	20.11	23.59
2π/3	$F_{5,6}$	18.57 *Aug-98*	21.81 *Aug-97*	13.75 *Aug-06*	10.40	17.73	20.61
π/3	$F_{7,8}$	19.29 *Sep-02*	24.93 *Feb-04*	18.36 *Jul-06*	12.17	17.71	6.87
5π/6	$F_{9,10}$	20.40 *Aug-98*	14.36 *Apr-03*	14.58 *Aug-06*	11.88	11.44	10.17
π/6	$F_{11,12}$	3.76 *Oct-99*	9.40 *Mar-98*	14.93 *Oct-03*	2.84	4.35	12.69
Break date					*Apr-00*	*Apr-96*	*Sep-06*

Source: Own estimations

This result reminds us of the effects of structural breaks and the proper use of this extension of the HEGY test. Since structural breaks may disguise an otherwise stationary process as one presenting a unit root, the HEGY test considering structural breaks would eventually only confirm if the unit root found is a real case when the test was applied without considering structural breaks. If the HEGY test (without structural breaks) suggests that the series do not contain a unit root test, the extended HEGY (consider structural breaks) test should not be carried out, given that the former has more power.

We have also considered the possibility of modelling the structural break by using a single dummy variable. This will help to reinforce or confirm some results found. The results are presented in Table 3.11. In terms of unit roots at the zero frequency, this analysis ratifies the results already found when the test was performed without considering breaks. Using both selection criteria, we cannot reject a unit root at the zero frequency. In terms of unit roots at seasonal frequencies, for the quantity of exported soybeans and exported maize, we still cannot reject the null of a unit root at the bi-annual frequency using both criteria, even at high levels of significance. Therefore, with some degree of confidence, we can consider that these two series present unit roots at this frequency and that the presence of structural breaks has not disguised as a unit root process an otherwise stationary one.

Using a more conservative approach, we can rule out the possibility of a unit root at the quarterly frequency in the quantity of domestic supplied wheat. Either in the minimizing *t* and maximizing *F* criterion or in the criterion that maximizes the significance of the break, we can reject unit roots at that frequency at standard level of confidence. Therefore, the problem we saw above of a new unit root when introducing structural breaks, might be the result of an overparameterisation, intrinsic when using several dummies to capture the break, of a model using a very short sample. Nevertheless, the fact that we have used such a short sample requires care not only in this case but also in the whole analysis.

Table 3.11 Test of seasonal unit roots in monthly series under the presence of unknown structural break. Single dichotomic variable

Frequency	Statistic	$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b)$ $i = 1,2$ $\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b)$			$\hat{T}_b = argmax\, F_\theta(T_b)$		
		qesoy	qemaz	qdwht	qesoy	qemaz	qdwht
0	π_1	-2.33 *Apr-96*	-2.72 *Dec-96*	-0.77 *Jun-06*	-2.1	-2.17	-0.93
Π	π_2	-5.42 *Apr-96*	-3.10 *Aug-97*	-5.51 *Apr-06*	-4.94	-2.9	-5.06
π/2	$F_{3,4}$	22.65 *Aug-03*	24.55 *Apr-96*	16.46 *Aug-97*	15.41	20.11	14.69
2π/3	$F_{5,6}$	11.64 *Apr-96*	21.01 *Jan-97*	8.85 *Mar-05*	10.51	17.73	7.5
π/3	$F_{7,8}$	13.43 *Mar-96*	23.58 *Dec-96*	15.39 *Oct-99*	11.37	17.71	12.41
5π/6	$F_{9,10}$	13.36 *Apr-96*	12.81 *Apr-96*	11.18 *Feb-97*	13.56	11.44	7.87
π/6	$F_{11,12}$	1.63 *Mar-96*	5.45 *Apr-97*	12.64 *Nov-03*	2.03	4.35	10.94
break date					*Apr-96*	*Dec-96*	*May-06*

Source: Own estimations

Controlling for the presence of structural breaks has led us to confirm or reject some of the results found before when we carried out the HEGY test without structural breaks. The quantities of exported maize and soybeans seem to present some seasonal unit roots, and these results have not been explained by a structural break that has affected the test. This means that some events such as weather, economic, technology or other institutional aspects have changed the pattern of seasonality on these series. In the rest of the series, however, the evidence points to characterize deterministic seasonality as a more appropriate approach for the control or modelling of seasonality.

3.7. CONCLUSIONS

Time series based on agricultural process may exhibit important seasonality. The limited storage capacity of annual crops, among other factors, may exacerbate the typical seasonality by introducing seasons where no exports or domestic supply is observed. This phenomenon cannot be explained by missing values nor the lack of registration.

Whilst temporal aggregation may help to reduce the impact of seasonality, the implications for the data generation process and the inference based on it could be serious. Monthly data present some advantages since it keeps relevant and useful information that tends to be hidden when data are aggregated into lower frequencies. Consequently, if the data are originally available in monthly data, temporal aggregation may not add any value and might be counterproductive.

Whilst in general seasonality in agriculture is seen as stable and predictable, stochastic events such as weather, economic decisions, technology and other institutional changes may affect the stability of the seasonal pattern, making the use of a deterministic approach to treat seasonality inadequate. As long as the stochastic elements do not have permanent effects on the seasonal pattern, a deterministic approach may be appropriate, but if these stochastic elements are transmitted or have permanent or cumulative effects on the seasons, a specific treatment should be attempted at the time of dealing with these series. Therefore, it is necessary to test if the stochastic elements present in seasons affect their pattern, and the HEGY test is suggested as the appropriate tool.

Nevertheless, the HEGY test has never been applied on agricultural time series where zero values are part of the domain of the series and give strong seasonal patterns to this type of series. A Monte-Carlo simulation exercise has been attempted to verify if the critical values used in this test are affected by the phenomenon. It was found that the presence of zero values in seasons in monthly series does not seem to affect the distributions of the test statistics used in the HEGY approach. The critical values obtained for these cases in monthly data do not differ substantially from the ones already found in the literature without this characteristic of the data. Therefore, the tests of seasonal unit roots can be still applied and the critical values remain valid. Given this fact, the additional specifications in the data generation process considered in this analysis can be seen as augmenting the cases already analysed by the literature.

The possibility that a stationary process could be wrongly deemed as a non-stationary one because of the presence of structural breaks has also been considered. However, no suitable critical values are available for any type of monthly data when considering the presence of structural breaks. Therefore, Monte-Carlo experiments have been run in order to obtain the appropriate ones, specifically for the short sample case. As in the quarterly case, it has been found that the presence of breaks in seasonal contexts do not affect the critical values for testing unit roots at the zero frequency. However, given the additional parameters, critical values for the rest of the seasonal frequencies tend to be larger in absolute value with respect to the cases where no breaks are considered.

Additionally, it has been seen that the presence of zero values does not seem to affect the critical values. Only the structural breaks generate differences, as in the non-seasonal case. Consequently, the critical values tabulated here can be used in other general contexts. Since critical values for monthly data for seasonal unit roots under the presence of structural breaks are not available, this constitutes an important contribution for applied work.

The possible misplacement of the identification of the break date has been analysed and, despite the fact that there exists some evidence of anticipation in the time the break date is selected, this evidence cannot be seen as conclusive. It was not been possible to identify clearly by how much the rule anticipates the break. This phenomenon may not be captured because the sample used is very short, and an analysis with larger samples may be necessary. Moreover, differences in the way the DGP used in the Monte Carlo experiment, as well as the size of the break, may be behind these differences.

A simpler model for considering unknown breaks in the HEGY approach using a single dummy variable has been also evaluated. Appropriate critical values were also calculated revealing no important departures from the ones found when using the standard HEGY approach under the presence of breaks. The advantage of this approach is the gain in the power of the tests because of the simpler and more parsimonious specification of the estimation equation. The disadvantage is that this approach does not allow changes in the seasonal pattern to be captured but just standard shifts in the intercept of the series. It has been observed that, as in the non-seasonal case, the break tends to be misplaced by one period when using monthly seasonal data. Therefore, rather than being the seasons, it is the way the structural break is captured that generates the misplacement.

An application of different techniques for the inference on the existence of deterministic and stochastic seasonality has been performed using data on monthly exports and monthly domestic supply of soybeans, maize and wheat in Argentina between 1994 and 2008. The nature of the seasonality affecting the series under study has been analysed by the inspection of the plots of the series. Additionally, the ACFs and the PACFs have been compared to their theoretical counterparts for pure seasonal processes. A simple deterministic approach to model seasonality has been applied to the series. This exercise revealed that the deterministic approach might provide an adequate explanation for the seasonal variation, especially in the case of exports. This suggests that exports tend to be more affected by seasonality than the domestic supply as a result, possibly, of the availability of vessels in given parts of the year.

On the other hand, the HEGY approach has not rejected unit roots in all series at the zero frequency, and, in some series, at other seasonal frequencies as well. Some of these unit roots have been confirmed when the test has been performed using a HEGY test under unknown structural breaks, suggesting that those unit roots are real and not the effect of a structural break.

Therefore, although stochastic events may have permanent effects on the seasonal pattern of some agricultural series that might require some specific econometric treatment, and that some of the series could not reject seasonal unit roots in some of the series, the evidence is not very strong. This suggests that deterministic seasonality might provide a sufficient and adequate approach to the modelling of seasonality. This, however, does not completely exclude the possibility that some of the series might need an approach in their modelling that considers the stochastic seasonality such as seasonal cointegration.

CHAPTER FOUR

THE EXPORT AND DOMESTIC SUPPLY OF COMMODITIES IN ARGENTINA: A COINTEGRATION EXERCISE

Summary

The estimation of the export and domestic supply equations developed in Chapter 1 is attempted using the Engle, Granger, Hylleberg and Lee (1990) seasonal cointegration or HEGY approach. This approach has never been applied to monthly data, and the seasonal cointegration relationships for these frequencies are obtained here. Seasonal unit roots tests are performed on a series of prices of agricultural commodities between 1994 and 2008 in Argentina, with and without unknown structural breaks. The seasonal unit roots found in these prices do not match the frequencies of the seasonal unit roots found in the quantities of export and domestic supplies in Chapter 2; this suggests that only Engle–Granger cointegration, given the presence of long-run unit roots, could be applied. Augmented Dickey–Fuller tests on the residuals of the long-run cointegration relationships suggest ambiguous evidence of cointegration, given that normality assumptions are violated by the data and that the different critical values suggested by the literature to perform the inference indicate mixed evidence. Cointegration is rejected once insignificant error correction terms are found in the error correction model. The quantity of data used, the data themselves, the cointegration technique approach used and the rigid requirements of the cointegration technique are all identified as the main explanations for these results.

4.1. INTRODUCTION

Technological risk in production, and the volatility in agricultural commodity markets, affect the way producers and traders make their output and commercialisation decisions. The export and domestic supply of agricultural commodities is affected not only by the level of prices in both markets but also by the expected variances and covariances of the prices of both products. As a consequence, assuming that traders are interested in the reduction of volatility in their profits, a cross hedging strategy may be advisable by increasing the supply of the product that exhibits the lower expected variance and/or presents a lower or negative covariance with the price of the input. Additionally, futures markets can increase the certainty about the price of the input a trader will face at the time of making their commercialisation decisions. Consequently, the prices of both commercialised products, the futures prices, and the input price affect the supply of both exports and domestic products.

In the first chapter of this research, a model addresses these issues, assuming a storable agricultural product subject to additive output risk, where the export and supply decision are made by a trader, and where a futures market is available to hedge against fluctuations in the price of the product. The development proved to be particularly hard to tract analytically and econometric validation was suggested as a way of gaining more insight about the behaviour of the model. Two supply equations, one for the exported and one for the domestically supplied product, have been identified that replicate the characteristics in the supply of agricultural commodities outlined above, that are linear and subject to econometric estimation.

In the second chapter, the econometric validation procedure started by discussing important issues related to the nature of agricultural time series, specifically with respect to their seasonality and the stochastic elements that can affect it. Seasonal unit root tests using the HEGY approach were carried out on the exports and domestic supply of three agricultural commodities in Argentina, as well as discussing other general aspects related to testing procedures in agricultural time series data. The results obtained there suggested that the series analysed, when used in the estimation of the equations of the model developed in Chapter 1, require a treatment that considers the stochastic seasonality present in some of the series. Seasonal cointegration is suggested as a viable alternative.

Integration and cointegration analysis have received major attention for a number of years. It has become a mature technique widely accepted and used in economics, particularly in macroeconomics. Its extension to series that present seasonal unit roots is more recent, but has also received important attention. However, its focus has been primarily on series that are observed quarterly, with substantially less focus on series with higher frequencies, such as monthly data.

One of the main advantages of the cointegration technique lies in the less restrictive conditions in terms of the exogeneity of the explicative variables. In alternative estimation techniques, such as transfer functions through autoregressive distributed lag models (ADL), there are assumptions about the unidirectional effect of the independent variable on the dependent variable. Cointegration analysis presents, in this sense, a less restrictive framework that facilitates the analysis as well as reduces the additional risk associated with extracting a false conclusion about the exogeneity that might be transmitted to the proper estimation. Additionally, the fact that, under given conditions, it is possible to use data without performing previous transformations, such as differenced data or the application of filters, allows the data to be used more efficiently and take advantage of all the information contained in series.

However, and at the same time, the cointegration analysis may be seen to be a little rigid in some cases. The conditions imposed on the concurrent integration of the series involved in the estimation may be too restrictive to be met, and can lead to making conclusions about the non-existence of relationships between the variables, when in fact, it was just the requirements of cointegration that failed to be met. (i.e., the fact that two variables are integrated of a different order does not imply the absence of a relationship between them).

Given its advantages, and despite its disadvantages, cointegration analysis has become very popular in the analysis of time series data and received substantial attention, improvements and additions since its inception in economics. One of them has been the analysis of cointegration with seasonal data and particularly, when seasons are stochastic.

Probably given its simplicity, the seasonal cointegration analysis has been based mainly on quarterly data. Whilst its extension to monthly data may be almost direct, there have been very few research applications with data of this frequency, even though working with monthly data has been identified as preferred over aggregation or, alternatively, the use of original quarterly data. One of the first purposes of this chapter is to make a deeper exploration of the seasonal cointegration on monthly data by obtaining seasonal cointegration relationships ready to be estimated using OLS on monthly data based on the Engle *et al*. (1993) approach or EGHL.

This chapter is a natural extension to the work done in the third chapter on seasonal unit roots on monthly data, where a comprehensive analysis of the seasonal unit roots was performed on a time series of monthly quantities of exported and domestically supplied grains in Argentina between 1994 and 2008. The particularity of the work done in the third chapter with respect to other similar contributions is that some series under study presented zero values. Therefore, facing the possible inadequacy of the critical values generally used to perform the seasonal unit root tests, new critical values were obtained for this particular series. This chapter also considered the possibility of unknown structural breaks that may affect the data generation process, and the tests were performed under that case.

The model developed in the first chapter postulates that the export and the domestic supply decisions are explained by the future price, the export and the domestic price, and the primary product price. Therefore, cointegration analysis is applied between quantities of exported and domestically supplied grains in Argentina, whose stochastic seasonality was analysed in the third chapter, and the respective prices presented in the model, whose stochastic seasonality will be briefly analysed in this chapter, with the idea of empirically verifying the model developed.

The mathematical and theoretical complexities of the model developed in the first chapter have complicated the analysis on the sign of the coefficients of the equations. These coefficients depend on the interaction of the expected variances and covariances of all the prices involved, and only through assumptions about their values is it possible to determine their sign and size. Nevertheless, the assumptions will eventually determine signs of the coefficients that might not be realistic and simply the result of the assumption made about them. Therefore, a second purpose of the current chapter is to obtain an idea on how this model behaves, particularly what are the signs and sizes of the coefficients of the specified equations. The objective is to determine how adequate the model is in its efforts to explain the export and the domestic supply of storable goods.

This presents an additional serious problem since this is the first empirical validation of this model, which means the absence of other applications with which to compare these estimates. In general, a researcher uses a comparison of its estimations against previous and similar exercises to obtain some security about the procedure or the quality of the results obtained. Unfortunately, in this chapter we have not such an assistance given the limitation of being the first empirical application.

This has important implications. Given these limitations, a researcher needs to rely on intuition or previous knowledge to judge how the estimation obtained during the empirical validation make sense in economic terms. This is, of course, unsatisfactory since the explanation of a particular economic aspect should be clear from the model under study without the need to refer to other theories or previous knowledge.

On the other hand, however, it is also true that every theoretical model had a moment where its empirical validity or suitability was not completely certain. It has been through successive empirical validation and the capability of being subject to continuous exercises of falsifiability that makes theories conform to the standard scientific principles. The practical implication of this, in light of this current exercise, is that this empirical application constitutes, probably, the first reference for future empirical applications in other contexts.

This chapter is organised as follows. In the first section, we will provide a short review on the effects of stochastic seasonality on the estimation of econometric models. Immediately, we will present an analysis of the main features of cointegration and its seasonal extension. In the third section, the seasonal cointegration relationships for monthly data are presented. In the fourth section, we will reintroduce the problem of the trader supplying exported and domestic products, as presented in the first chapter, as well as define some elements and considerations necessary for the estimation stage. In the following section, the price data that will be used in the estimation are presented and immediately, a seasonal unit root analysis using the HEGY approach is performed on them. In the seventh part, a cointegration analysis will be performed between prices and the quantities analysed in the second chapter. In the eighth section, an error correction model (ECM) will be estimated on the supply equations presented in the first chapter and after that, a discussion on alternative estimation procedures will follow. In this section, an analysis on the variables used for the identification of the supply functions is performed. Finally, some final comments with the main results found will be presented.

4.2. STOCHASTIC SEASONALITY AND THE ESTIMATION OF ECONOMETRIC MODELS

When time series present strong seasonality patterns, the possibility of exhibiting seasonal unit roots cannot be ruled out. Therefore, cointegration may occur not only (or instead of) at the zero frequency, but also at the seasonal cycles. If the cointegration relationship is thought to be a long run one (or at zero frequency), and seasonal unit roots present in the series are ignored or not properly treated, the relationship between two or more variables may give inconsistent estimates.

Nevertheless, cointegration or seasonal cointegration are not the only alternatives to conduct the econometric estimation of a model. Before the development of the cointegration technique, econometricians have been estimating models on time series also under seasonality. Jorgenson (1966) has contributed with the development of distributed lag models or, as he called them, rational lag models. In addition, Nerlove (1972), in his discussion about dynamics in production and investment, has extensively applied models with distributed lags.

The problem of stochastic trends in these models is solved generally by differencing the variables that present a unit root. If the series presents a stochastic trend, applying the difference operator to the series would assure, in general, that the series would observe the desired white noise properties, to apply for example the standard Box-Jenkins methodology. In this case, rather than estimate or forecast the level of the series, the changes in them, attributable also to some white noise variable, will be modelled.

The problem with this approach lies in that the difference operator applied to the series removes important and relevant information about the process to be modelled. The possibility of performing spurious regressions and its proposed solution can hide or specifically exclude the long-run relationship between the variables. This means that only short-run effects are modelled and the long-run relationship between the variables is missing.

A second problem with the ADL models lies in the required exogeneity of the explanatory variable. It is expected that past and current innovations of the dependent variable have no effect on the independent variable or that no feedback exists. When no doubts exist about the exogeneity of the independent variable, no autocorrelation will be observed in the residuals and the procedure is perfectly valid. However, problems may occur if the variables are jointly determined making difficult to disentangle the interdependence between them. This is particularly the case when supply or demand equations are estimated given that prices and quantities are generally jointly determined.

Additionally, the treatment of seasonal data in these types of approaches does not go further than the application of seasonal difference operators (additional to the difference operators used to remove unit roots), or a pure deterministic treatment. As we have seen in the previous chapter, the use of different seasonal filters is not advised given the changes they introduce in the data generation process that complicates the estimation as well as inference.

4.3. NOTES ON COINTEGRATION AND SEASONAL COINTEGRATION

Engle and Granger (1987) have addressed the issues with respect to the loss of relevant information on the process, the estimation of long-run relationships, and the requirements of the exogeneity of the explanatory variables. This approach, based on the cointegration findings of Granger (1981), allows in this sense a less restrictive treatment of the estimation of time series models.

According to this framework, if two (or more) time series combine to present a single unit root or they are integrated of the first order, I(1), there is a possibility that a linear combination of them may be stationary, or I(0). This implies that, whilst variables may not be stationary, a linear combination of them may be; permitting a direct estimation of a model, through the ordinary least squares method, including these non-stationary variables will provide properly behaved residuals and consistent estimates of the parameters.

By using series in their unaltered univariate dynamics, the cointegration approach does not discard important and relevant information about the data generation process. Therefore, the long-run equilibrium relationship between variables can be properly represented in the estimation. Additionally, through the error correction model (ECM) specification, short-run dynamics can be considered by the deviations about the long run equilibrium. This implies that the cointegration approach allows a more comprehensive treatment of the behaviour and relationships between series.

Additionally, the exogeneity requirements that standard dynamic models impose on the relationship to be modelled are particularly less stringent. The EG approach can be used when is not possible to identify a dependent variable and a set of independent variables. In this context, the assumption that a variable is exogenous or weakly exogenous[30], when it is not, does not affect the estimation or the inference properties of the procedure. This is not the case for the ADL, for example, where the violation of the exogeneity assumption may affect the results obtained during the regression.

The possibility that additional unit roots, besides the long-run one, may be present in the series under analysis have received special treatment within the EG approach. Particularly in the case of seasonality, the undetected presence of stochastic seasonality or seasonal unit roots in the cointegration analysis may lead to inconsistent estimates. However, as in the non-seasonal case, if the series involved in the analysis contain concurrent seasonal unit roots at the same frequencies, it is possible to estimate a model with these variables in the context of seasonal cointegration.

The Hylleberg, Engle, *et al*. (1990), or HEGY, approach provides the appropriate framework for testing seasonal unit roots by separating the different roots (zero and seasonal unit roots) present in a series and allow for testing each of them separately. The HEGY technique has been applied extensively in quarterly and monthly data. It provides an efficient and relatively simple way of testing for the presence of seasonal unit roots.

The relationship between seasonal unit root testing and seasonal cointegration is direct. In fact, Hylleberg, Engle, *et al*. (1990) also provide the framework for testing and estimating seasonal cointegration. The same variables created for testing seasonal unit roots are used also for testing for a cointegration relationship. In fact, the seasonal unit root test can be seen as a cointegration analysis of a variable with itself.

[30] Formally, a variable x_{it} is weakly exogenous for the variable y_{it} if the marginal distribution of x_{it}no contains relevant information for making inference on y_{it} (Enders, 2010)

Engle, Granger and Hylleberg *et al*. (1993) or EGHL, based on the HEGY approach for seasonal unit roots developed and provided the framework to analyse seasonal cointegration and the Seasonal Error Correction Model (SECM). This procedure mimics the two-step procedure for cointegration developed by Engle and Granger (1987) for non-seasonal data. A cointegration relationship is derived to test if the residuals of that relationship are stationary. Eventually, that relationship is introduced in an Error Correction Model (ECM) to permit a long run relationship and the deviations from the equilibrium.

Engle, Granger and Hylleberg *et al*. (1993) provide the tests to be used in the HEGY approach to treat seasonal cointegration and have constituted the analytical framework on this technique. This technique has also been discussed in extension by Osborn (1993), suggesting alternative specifications of the error correction model linking the analysis to the periodic cointegration approach.

As for the non-seasonal Engle-Granger procedure, the EGHL approach presents important drawbacks. The non-standard distribution of the statistics used to make inference is presented generally as one of the drawbacks. However, the fact that these distributions have been tabulated and are widely available should not be an impediment for the use of this procedure. Another drawback highlighted is the two-stage nature of the procedure and the possibility of carrying mistakes from one stage to the next, as Enders (1995) suggests. Any mistake committed at the time of identifying the seasonal unit roots, for example, may lead to identifying and concluding the existence of a cointegration relationship between two variables that do not present unit roots at the same frequency if the previous step has not been done properly.

Nevertheless, these are least important of its drawbacks. The issues related with the consistency of the estimates and the potential lower power of the tests seems to be particularly problematic. When series are cointegrated, the ordinary least squares (OLS) estimator will be super consistent; implying that as the sample size goes to infinity the estimates will converge to their true value much faster than the usual OLS estimator (Harris and Sollis, 2003). However, when samples are finite this presents a problem, as there is a bias between the estimator and the true parameter value, leading to the possibility of obtaining inconsistent estimators.

The second important problem is associated with the potential low power of the tests. The fact that the ADF tests require a very long lag length in order to generate white noise residuals, increases the likelihood of over rejecting a null hypothesis when it is valid. Moreover, when variables are fully seasonally cointegrated, an SECM will contain as many error correction terms as cointegrated frequencies, which notably increase the number of parameters to be estimated.

On the other hand, Johansen and Schaumburg (1999) reconsidered the concept in a multivariate framework by extending the Johansen (1988) approach of Vector Error Correction Models (VECM), and trying to address some of the issues of the Engle-Granger methodology. This approach works through the relationship between the rank and the eigenvalues of the matrix of coefficients to determine the presence of cointegration.

This approach seems to have received more attention since then. Bohl (2000) analyses the German M2 demand for money on seasonally adjusted (through the application of seasonal filters) and unadjusted data using a Johansen-type of analysis. Using unadjusted data, the author finds a long-run relationship between the variables (M2, GDP, and interest rate), whilst no relationship appears when data are adjusted for seasonality using filters, suggesting the effect that seasonal filters have on the estimation and inference. Also on German M3 demand, Herwartz and Reimers (2003) find seasonal cointegration between similar variables. In a different context, Perez-Pascual and Sanz-Carnero (2009), using the Johansen approach, find that the Spanish wheat market is becoming more integrated within regions.

The main advantage of the Johansen-type of approach lies in the unrestrictive specification of the parameters in the error correction model for the estimation procedure, in contrast to the Engle-Granger type of analysis. Instead, an alternative method is proposed based on maximum likelihood approach and a more general asymptotic theory for the seasonal cointegration model. However, on the other side of the coin, the unrestrictive nature of the Johansen-type of approach may lead to mis-specification problems as Clements and Madlener (1999) suggest. On the other hand, with respect to the issue on the staged nature of the EGHL procedure, the seasonal cointegration tests and the VECM can be estimated in a single step in the Johansen-type of approach. This is less to error prone than the EGHL method comprising two-stages.

However, Franses and McAleer (1998) highlight some problems with the Johansen approach. The first drawback is related to the interpretation of the estimated cointegrated vectors. According to the authors, this problem is particularly acute when monthly data are used. In contrast, the interpretation of the coefficients in the EGHL approach is more amenable to common economic interpretations, (i.e. elasticities); while the elements in the vectors of the VECM cannot be directly associated with known economic concepts.

The second drawback of the Johansen approach is the interpretation of the intercepts under seasonal cointegration, since intercepts imply expanding cycles, which are deemed economically difficult to interpret. Therefore, it is recommended that, if not too many variables are analysed and if it is known that some variables are of more interest than others are, the bivariate approach using the EG cointegration approach should be used.

Although the Johansen methodology fixes several of the issues present in the Engle-Granger methodology, the interpretational advantages of the latter are particularly superior. Despite its problems, the Engle-Granger methodology is not only simpler, but also clearer. It is more intuitive when analysing its results. Moreover, its estimation is easier, particularly when single equation models are involved.

It is important to remark that the selection of the estimation methodology has important implications and can yield completely opposite results. Beenstock, Goldin and Nabot (1999), analysing the demand for electricity in Israel, find that the EGHL approach rejected cointegration whilst a Johansen type of approach indicated that variables were in fact cointegrated. In addition, Huang and Shen (1999) found different results between the EGHL and a Johansen-type of approach in their analysis on the demand for international reserves in Taiwan, suggesting that the Johansen-type presented lower goodness of fit, less stable functional form and economically inconsistent estimates.

The bivariate approach, based on EGHL, has also received some empirical attention. For example, in this framework Franses (1993) and Osborn (2002) analysed the relationships between seasonal cointegration and periodic cointegration and provided alternative insights into this type of analysis. Bohl and Sell (1998), analysing the demand for cash balances in Germany, found no long-run cointegration relationship between the demand for money and consumption expenditures, but they were cointegrated at the annual and the biannual frequencies. Hamori and Tokihisa (2001), in their study on the Japanese money demand function, could not find evidence of cointegration between the demand for money and GDP. Huang and Shen (1999) found a lower speed of adjustment (seasonal error correction terms) to the long run, explained by Taiwan's high level of international reserves. Recently, Hasan (2011) could not reject the hypothesis of neutrality of money in the US by finding long-run cointegration between money and prices, and no seasonal cointegration between the real output and money supply. All these examples must be added to the seminal paper by Engle, Granger and Hylleberg *et al*. (1993).

All applications we have discussed so far have dealt with quarterly data. In fact, very little has been applied on monthly data. It is recognised by Franses and McAleer (1998) that the extension to monthly data could be complicated in practical grounds given the number of unit roots involved. Nevertheless, some applications can be found, primarily using a Johansen-type of approach.

In an analysis of different Spanish production indicators, Caminero and Diaz-Emperanza (1997) found little evidence of cointegration at seasonal frequencies. McErlean *et al*. (2003) found evidence, with monthly data, that the EU's direct payments to beef producers are cointegrated with the level of output, which cast some doubts about the non-distortionary nature of these subsidies. Darne (2004) found evidence of seasonal cointegration between stocks and retail sales in the US industrial sectors.

It is important to remark that the use of monthly or quarterly data is not simply a matter of choice by the researcher or data availability. As discussed in the previous chapter, high frequency data may exhibit superior estimation performance (Amemiya and Wu, 1972). However, it is recognised that measurement errors are more frequent in monthly data. Nevertheless, the use of the data in its original reported frequency tends to be preferred to temporal aggregation, since information about the data cycle tends to be lost (Rossana and Seater, 1995), and, by its smoothing properties, could lead to false conclusions (Silvestrini and Veredas, 2008).

As we mentioned, all the applications on monthly data found have been made using a Johansen-type of analysis. Whilst Osborn (2002) suggests that the extension to monthly data of the EGHL approach is direct, no applications have been found on this ground.

4.4. SEASONAL COINTEGRATION RELATIONSHIPS IN MONTHLY DATA

The idea of this section is to try to contribute to the identification of the bi-variate monthly cointegration relationships under the EGHL approach. As we will see, this development will prove to be unnecessary for the empirical application we present later. However, we consider it interesting *per se* and necessary to explain further the problems under consideration.

Let us consider an $n \times 1$ vector of time series, x_t, observed every month. The Wold representation of such a process can be described as

$$(1 - B^{12})x_t = C(B)\varepsilon_t \qquad (4.1)$$

where ε_t is an $n \times 1$ vector of normally and independently distributed variables with mean zero and positive definite matrix of covariances. $C(B)$ is an $n \times n$ matrix of backshift operators. A decomposition of the equation $1 - B^{12} = 0$ has 12 solutions given by

$$\theta_k = \left(\pm 1, \pm i, \frac{-1}{2} \pm \frac{\sqrt{3}}{2} i, \frac{1}{2} \pm \frac{\sqrt{3}}{2} i, \frac{-\sqrt{3}}{2} \pm \frac{i}{2}, \frac{\sqrt{3}}{2} \pm \frac{i}{2}\right).$$

The equation reveals the zero unit root and eleven seasonal unit roots. The matrix of polynomials $C(B)$ can be linearised to obtain, after reparameterisation (Franses, 1991).

$$\begin{aligned} C(B) = {} & -\pi_1 B\varphi_1(B) + \pi_2 B\varphi_2(B) + (\pi_3 + \pi_4 B)B\varphi_3(B) + (\pi_5 + \pi_6 B)B\varphi_4(B) \\ & + (\pi_7 + \pi_8 B)B\varphi_5(B) + (\pi_9 + \pi_{10} B)B\varphi_6(B) + (\pi_{11} + \pi_{12} B)B\varphi_7(B) \\ & + C^{**}(B)\varphi_8(B) \end{aligned} \qquad (4.2)$$

Where

$$\varphi_1(B) = (1 + B + B^2 + B^3 + B^4 + B^5 + B^6 + B^7 + B^8 + B^9 + B^{10} + B^{11})$$

$$\varphi_2(B) = (1 - B + B^2 - B^3 + B^4 - B^5 + B^6 - B^7 + B^8 - B^9 + B^{10} - B^{11})$$

$$\varphi_3(B) = (1 - B^2 + B^4 - B^6 - B^7 + B^8 - B^{10})$$

$$\varphi_4(B) = (1 - B + B^3 - B^4 - B^6 - B^7 + B^9 - B^{10})$$

$$\varphi_5(B) = (1 + B - B^3 - B^4 + B^6 + B^7 - B^9 - B^{10})$$

$$\varphi_6(B) = \left(1 - \sqrt{3}B + 2B^2 - \sqrt{3}B^3 + B^4 - B^6 + \sqrt{3}B^7 - 2B^8 + \sqrt{3}B^9 - B^{10}\right)$$

$$\varphi_8(B) = (1 - B^{12})$$

Replacing equation (4.2) into (4.1) and operating yields

$$C^{**}(B)\Delta_{12}x_t = \pi_1 x_{t-1}^{(1)} - \pi_2 x_{t-1}^{(2)} - (\pi_3 + \pi_4 B)x_{t-1}^{(3)} - (\pi_5 + \pi_6 B)x_{t-1}^{(3)} - (\pi_7 + \pi_8 B)x_{t-1}^{(4)} - (\pi_9 + \pi_{10} B)x_{t-1}^{(5)} - (\pi_{11} + \pi_{12} B)x_{t-1}^{(6)} + \varepsilon_t \tag{4.3}$$

where the variables $x_t^{(k)}$ are transformations of the variable x_t given by the expressions $\varphi_k(B)$ above. Equation (4.3) provides the basic estimation equation for testing for seasonal unit roots in the HEGY methodology, and it provides the basic equation for the SECM. The $\varphi_k(B)$ relates to the different zero and seasonal unit roots. Each variable is constructed to remove the influence of the $j \neq k$ roots and leaving the root k in their univariate dynamics. For example, $x_t^{(1)}$ relates to the zero frequency adjusted for the seasonal unit roots. Therefore, the influence of the seasonal frequencies is removed, and this transformed variable can be used to test for the presence of unit roots at the zero frequency. Similarly, for annual frequency unit root, the variable $x_t^{(2)}$ removes the influence of the zero and the rest of the seasonal unit roots. In the case of $x_t^{(3)}$, this variable is constructed to evaluate the complex pair of unit roots ($\pm i$); in addition, the rest of the variables are created to remove the zero frequency and the other seasonal frequencies, except the one under study.[31]

To simplify the analysis, let us assume that we have two monthly variables, x_t and z_t, and we want to model the first as a function of the second. If $x_t, z_t \sim SI(1)$, this implies that $x_t^{(k)}, z_t^{(k)} \sim I_\theta(1)$ or they contain a unit root in each of the frequencies. In this case, cointegration at the zero frequency implies that there exists a unique linear combination between the variables $x_t^{(1)}$ and $z_t^{(1)}$ that is stationary, or

[31] Beaulieu and Miron (1993) make the decomposition with twelve variables. However, it can be seen that their variables can be collapsed into the ones presented here.

$$\omega_t^{(1)} = x_t^{(1)} - \gamma_1 z_t^{(1)} \sim I(0) \quad (4.4)$$

where we have made use here of the usual normalisation of setting the coefficient on the dependent variable to unity to define the cointegration vector, and the coefficient $\gamma_1$is the long run relationship at the zero frequency. The variables $x_t^{(2)}$ and $z_t^{(2)}$ relate to the semiannual frequency, and cointegration at this frequency implies the existence of a unique stationary linear combination such that

$$\omega_t^{(2)} = x_t^{(2)} - \gamma_2 z_t^{(2)} \sim I(0) \quad (4.5)$$

As we have seen, the variables $x_t^{(3)}$ and $z_t^{(3)}$ are transformations of the respective variables to remove the real roots and the rest of the seasonal roots and leave the complex pair $(\pm i)$ in their univariate dynamics. Following Ghysels and Osborn (2001), we can build a cointegration relationship between these two variables by considering a cointegration complex coefficient $\gamma_R \pm i\gamma_i$. Factorising $(1 + B^2) = (1 + iB)(1 - iB)$ and applying the first factor to $x_t^{(3)}$ and $z_t^{(3)}$, we can work out the cointegration relationship between $x_t^{(3)} + ix_{t-1}^{(3)}$ and $z_t^{(3)} + iz_{t-1}^{(3)}$. This relationship implies the existence of a coefficient $\gamma_R^{(3)} - i\gamma_I^{(3)}$ such that the variable is formed as

$$x_t^{(3)} + ix_{t-1}^{(3)} - (\gamma_R - i\gamma_I)\left(z_t^{(3)} + iz_{t-1}^{(3)}\right) \sim I(0) \quad (4.6)$$

On the other hand, the complex conjugate pair of these variables must also follow the same relationship and it will be stationary.

$$x_t^{(3)} - ix_{t-1}^{(3)} - (\gamma_R + i\gamma_I)\left(z_t^{(3)} - iz_{t-1}^{(3)}\right) \sim I(0) \quad (4.7)$$

Adding these two expressions will also be a stationary variable since a linear combination of stationary variables will also be $I(0)$ and the imaginary terms will disappear

$$2x_t^{(3)} - 2\left(\gamma_R z_t^{(3)} + \gamma_I z_{t-1}^{(3)}\right) \sim I(0)$$

This variable is real and stationary. Therefore, we can define a cointegration relationship between the transformed variables $x_t^{(3)}$ and $z_t^{(3)}$ as

$$\omega_t^{(3)} = x_t^{(3)} - \gamma_3 z_t^{(3)} - \gamma_4 z_{t-1}^{(3)} \sim I(0) \tag{4.8}$$

where we have relabelled $\gamma_3 = \gamma_R$ and $\gamma_4 = \gamma_I$. Note that if we subtract equation (4.7) from equation (4.6), we will obtain another cointegration relationship that will also be stationary. We obtain two cointegration relationships since one variable is used to consider two unit roots. The cointegration relationships found so far are analogous to those found using quarterly data.

Franses and McAleer (1998) recognise that the analysis of seasonal cointegration in monthly data is expected to be much more complicated given the number of unit roots involved. In fact, the extensions to monthly data have been developed for the multivariate approach, or VECM, using the Johansen method (Caminero and Diaz-Emperanza, 1997; Darne, 2004). Two exceptions seem to be the work by Martin-Alvarez, Cano-Fernandez and Caceres-Hernandez (1999) and Cellini and Cuccia (2009). The former analyses monthly data seasonal cointegration using a different approach but it has not developed the cointegration relationships. It compares pairwise relationships between all the transformed variables. The latter, despite using monthly data, did not go into detail since they found that that the variables they were using presented unit roots only at the zero frequency, so invalidating the presence of seasonal cointegration. Therefore, we will try to develop the remaining cointegration relationships in order to apply the Engle, Granger and Hylleberg *et al*. (1993) cointegration approach to monthly data.

Variables $x_t^{(4)}$ and $z_t^{(4)}$ relate to the unit roots at $\left(\frac{-1}{2} \pm \frac{\sqrt{3}}{2} i\right)$. By continuing in the same fashion, we can factorise $(1 + B + B^2) = \left(\frac{1}{2} + \frac{\sqrt{3}i}{2} + B\right)\left(\frac{1}{2} - \frac{\sqrt{3}i}{2} + B\right)$ and apply the first term to the two transformed variables in question, we can consider cointegration between the transformed variables $x_t^{(4)}/2 + i\sqrt{3}x_t^{(4)}/2 + x_{t-1}^{(4)}$ and $z_t^{(4)}/2 + i\sqrt{3}z_t^{(4)}/2 + z_{t-1}^{(4)}$. As before, such cointegration implies the existence of a unique complex cointegration linear relationship, $\gamma_R \pm i\gamma_i$, such that the cointegration relationship is stationary or

$$\frac{x_t^{(4)}}{2} + i\frac{\sqrt{3}}{2}x_t^{(4)} + x_{t-1}^{(4)} - (\gamma_R - i\gamma_I)\left(\frac{z_t^{(4)}}{2} + i\frac{\sqrt{3}}{2}z_t^{(4)} + z_{t-1}^{(4)}\right) \sim I(0) \tag{4.9}$$

and

$$\frac{x_t^{(4)}}{2} - i\frac{\sqrt{3}}{2}x_t^{(4)} + x_{t-1}^{(4)} - (\gamma_R + i\gamma_I)\left(\frac{z_t^{(4)}}{2} - i\frac{\sqrt{3}}{2}z_t^{(4)} + z_{t-1}^{(4)}\right) \sim I(0) \tag{4.10}$$

Adding equations (4.9) and (4.10) we will obtain a first cointegration relationship,

$$\omega_t^{(4)} = x_t^{(4)} + 2x_{t-1}^{(4)} - \left(\gamma_5 + \sqrt{3}\gamma_6\right)z_t^{(4)} - 2\gamma_6 z_{t-1}^{(4)} \sim I(0) \tag{4.11}$$

While subtracting equation (3.9) from obtain yields

$$\omega_t^{(4)} = -\sqrt{3}x_t^{(4)} + \left(\sqrt{3}\gamma_5 - \gamma_6\right)z_t^{(4)} - 2\gamma_6 z_{t-1}^{(4)} \sim I(0) \tag{4.12}$$

Where we have relabelled $\gamma_5 = \gamma_R$ and $\gamma_6 = \gamma_I$. Both cointegration relationships are real and stationary. It can be seen that "duality" exists between them. Cointegration between variables $x_t^{(5)}$ and $z_t^{(5)}$ requires that we factorise the polynomial $(1 - B + B^2)$ as $\left(-\frac{1}{2} + \frac{\sqrt{3}i}{2} + B\right)\left(-\frac{1}{2} - \frac{\sqrt{3}i}{2} + B\right)$, and working in the same fashion, we obtain

$$\omega_t^{(5)} = -x_t^{(5)} + 2x_{t-1}^{(5)} + \left(\gamma_7 - \sqrt{3}\gamma_8\right)z_t^{(5)} - 2\gamma_7 z_{t-1}^{(5)} \sim I(0) \tag{4.13}$$

In addition, its dual cointegration relationship is

$$\omega_t^{(5)} = -\sqrt{3}x_t^{(5)} + \left(\sqrt{3}\gamma_7 + \gamma_8\right)z_t^{(5)} - 2\gamma_8 z_{t-1}^{(5)} \sim I(0) \tag{4.14}$$

The cointegration relationships between the remaining pairs of transformed variables, $(x_t^{(6)}, z_t^{(6)})$ and $(x_t^{(7)}, z_t^{(7)})$ are obtained in the same way and are represented by

$$\omega_t^{(6)} = \sqrt{3}x_t^{(6)} + 2x_{t-1}^{(6)} - \left(\sqrt{3}\gamma_9 + \gamma_{10}\right)z_t^{(6)} - 2\gamma_9 z_{t-1}^{(6)} \sim I(0) \tag{4.15}$$

$$\omega_t^{(6)} = -x_t^{(6)} - \left(\gamma_9 - \sqrt{3}\gamma_{10}\right)z_t^{(6)} - 2\gamma_{10} z_{t-1}^{(6)} \sim I(0) \tag{4.16}$$

$$\omega_t^{(7)} = -\sqrt{3}x_t^{(7)} + 2x_{t-1}^{(7)} + \left(\sqrt{3}\gamma_{11} - \gamma_{12}\right)z_t^{(7)} - 2\gamma_{11} z_{t-1}^{(7)} \sim I(0) \tag{4.17}$$

$$\omega_t^{(7)} = -x_t^{(7)} + \left(\sqrt{3}\gamma_{12} + \gamma_{11}\right)z_t^{(7)} - 2\gamma_{12} z_{t-1}^{(7)} \sim I(0) \tag{4.18}$$

The procedure for testing cointegration at each of these frequencies entails running an OLS regression on each of the cointegration relationships and testing for the presence of unit roots in the residuals using a Dickey-Fuller type test. As in the Engle-Granger two-step approach, superconsistent estimates will be obtained. For example, cointegration at the zero frequency can be tested using the following equation

$$\Delta\widehat{\omega}_t^{(1)} = \delta_1\widehat{\omega}_{t-1}^{(1)} + \sum_{i=1}^{\rho-1}\psi_i\,\Delta_{\mathrm{i}}\widehat{\omega}_{t-i}^{(1)} + \mu + \zeta t + v_t^{(1)} \quad v_t^{(1)} \sim IID(0,\sigma^2) \qquad (4.19)$$

where ψ_i are the coefficients of the lagged dependent variable included to ensure the desired white noise properties of the residuals, μ and *t* are the intercept and the trend that may or not be present. The test statistic is a t-test type of $H_0\colon \delta_1 = 0$ against $H_1\colon \delta_1 < 0$. If the null is not rejected, the residuals of the cointegration relationship, $\widehat{\omega}_t^{(1)}$, are not stationary and therefore, there is not cointegration at the zero frequency.

Testing for cointegration between variables $x_t^{(2)}$ and $z_t^{(2)}$ must be done by running a regression using equation (4.5) and analysing the behaviour of the residuals, $\widehat{\omega}_t^{(2)}$. The procedure is similar to the one sketched above, but the Dickey-Fuller equation is modified slightly to reflect the transformation operation in these variables. Therefore, the equation to be used is

$$\left(\widehat{\omega}_t^{(2)} + \widehat{\omega}_{t-1}^{(2)}\right) = -\delta_2\widehat{\omega}_{t-1}^{(2)} + \sum_{i=1}^{\rho-1}\psi_i\left(\widehat{\omega}_{t-i}^{(2)} + \widehat{\omega}_{t-i-1}^{(2)}\right) + \mu + \sum_{i=1}^{11}\zeta D_{it} + v_t^{(2)} \qquad (4.20)$$

where D_{it} is a dichotomic variable (to allow for determistic seasonality) corresponding to month *m* that may be or not be present. Again, the test is done on the $H_0\colon \delta_2 = 0$ against $H_1\colon \delta_2 < 0$. Given the prior estimation of the cointegration relationships, the standard critical values cannot be used, and those suggested by MacKinnon (1991) or Engle and Granger (1987) must be used. However, as we will see later, the election of these critical values may have important implications for conclusions regard the cointegration test.

To test for cointegration between $x_t^{(3)}$ and $z_t^{(3)}$ a regression is run using equation (4.8), and a unit root test is run on the residuals using

$$\left(\widehat{\omega}_t^{(3)} + \widehat{\omega}_{t-2}^{(3)}\right) = -\delta_3\widehat{\omega}_{t-1}^{(3)} - \delta_4\widehat{\omega}_{t-2}^{(3)} + v_t^{(3)} \qquad (4.21)$$

Potential additional dichotomous variables, lagged values of the dependent variable and intercept have been omitted for simplicity. By analogy with the HEGY test, the null hypothesis of no cointegration at this frequency requires that both, $\delta_3 = \delta_4 = 0$ in this regression.

Following Engle, Granger and Hylleberg *et al*. (1993) it is possible to derive the appropriate formula for the unit roots tests for the rest of the cointegration relationships. In analogy with the HEGY test, the null hypothesis of no cointegration at the frequency $\frac{2}{3}\pi$ implies that both $\delta_5 = \delta_6 = 0$ in the ancillary regression

$$\left(\widehat{\omega}_t^{(4)} + \widehat{\omega}_{t-1}^{(4)} + \widehat{\omega}_{t-2}^{(4)}\right) = -\delta_5\widehat{\omega}_{t-1}^{(4)} - \delta_6\widehat{\omega}_{t-2}^{(4)} + \upsilon_t^{(4)} \tag{4.22}$$

Again, it is possible to add lagged values of the dependent variables plus other components. The idea behind equation (4.22) is to consider a similar transformation on the residuals to the ones applied to the original transformed variables that yield those precise residuals. In this case, the factor $(1 + B + B^2)$ has been originally considered to remove the unit roots $\frac{-1}{2} \pm \frac{\sqrt{3}}{2}i$ in the transformed variables. In the same way, it is possible to obtain the rest of the unit root regression for the remaining frequencies. These are

$$\left(\widehat{\omega}_t^{(5)} - \widehat{\omega}_{t-1}^{(5)} + \widehat{\omega}_{t-2}^{(5)}\right) = -\delta_7\widehat{\omega}_{t-1}^{(5)} - \delta_8\widehat{\omega}_{t-2}^{(5)} + \upsilon_t^{(5)} \tag{4.23}$$

$$\left(\widehat{\omega}_t^{(6)} + \sqrt{3}\widehat{\omega}_{t-1}^{(6)} + \widehat{\omega}_{t-2}^{(6)}\right) = -\delta_9\widehat{\omega}_{t-1}^{(6)} - \delta_{10}\widehat{\omega}_{t-2}^{(6)} + \upsilon_t^{(6)} \tag{4.24}$$

$$\left(\widehat{\omega}_t^{(7)} - \sqrt{3}\widehat{\omega}_{t-1}^{(7)} + \widehat{\omega}_{t-2}^{(7)}\right) = -\delta_{11}\widehat{\omega}_{t-1}^{(7)} - \delta_{12}\widehat{\omega}_{t-2}^{(7)} + \upsilon_t^{(7)} \tag{4.25}$$

As was mentioned, the test at the zero frequency and the biannual frequency follow non-conventional distributions and special tabulated critical values have already been developed. For the rest of the seasonal unit roots critical values for quarterly data have been obtained by Engle, Granger and Hylleberg *et al*. (1993). Therefore, it is necessary to tabulate critical values for the rest of the frequencies in the context of monthly data.

Finally, if unit roots are present at the zero and each of the seasonal frequencies, the SECM for this bi-variate case can be represented by

$$\Delta_{12}x_t = \beta_0 + \sum_{i=1}^{l} \Phi_i \Delta_{12}x_{t-i} + \sum_{i=0}^{P} \Phi_i \Delta_{12}z_{t-i} + \beta_1 \widetilde{\omega}_{t-1}^{(1)} + \beta_2 \widetilde{\omega}_{t-1}^{(2)} + \beta_3 \widetilde{\omega}_{t-1}^{(3)} + \beta_4 \widetilde{\omega}_{t-1}^{(4)} + \beta_5 \widetilde{\omega}_{t-1}^{(5)} + \beta_6 \widetilde{\omega}_{t-1}^{(6)} + \beta_7 \widetilde{\omega}_{t-1}^{(7)} + \varepsilon_t \quad (4.26)$$

This equation is the analogue in the bi-variate framework is the one developed by Darne (2004) for the VECM Johansen-type approach. Following Engle and Granger (1987), as long as both series cointegrate in every frequency, the expressions for the seasonal deviations present in the SECM can be replaced by the residuals $\widetilde{\omega}_{t-1}^{i}$ obtained in the cointegration equations presented above. If the series do not cointegrate at all frequencies, only the seasonal error correction terms where they cointegrate will be present in the SECM.

The extension to consider more than one cointegrating variable is straightforward. Seasonal cointegration might be present (as a necessary but not sufficient condition) as long as every variable considered exhibits unit roots at the same frequencies. In this case, the SECM would be similar to the one presented above, but the residuals used in the estimation should be those obtained from the cointegration relationships involving all the variables considered. This is the approach followed by Bohl and Sell (1998), Huang and Shen (1999) and Hamori and Tokihisa (2001) using quarterly data.

4.5. THE EXPORT AND DOMESTIC SUPPLY OF AGRICULTURAL COMMODITIES

In the first chapter of this research, a model was developed that intends to explain the behaviour of the exports and the domestic supply of commodities when futures markets are operating; there is processing or trading and storage is allowed. In this framework, the export or domestic supply decision is affected by the export price, the domestic price, the primary product price, and the futures price. As such, the exported and the domestically supplied product are seen as two different products manufactured by a processor or trader that uses the primary product, produced by a farmer, as an input. Additionally, a storage company physically transfers the primary product over time. All these agents operate in the futures markets to hedge against fluctuations in the price of the primary product or obtain a speculative gain.

The model proved to be highly complex and the parameters that fit the supply equations depend on the different variances and covariances between all prices. Despite the assumptions made on possible values for the expected variances and covariances, it was impossible to identify theoretically what the signs of those coefficients would be. Nevertheless, the equations to be estimated are linear in their parameters and are capable of being estimated econometrically. The second purpose of this chapter is to obtain estimates of the potential values of these parameters.

Specifically, the objective is to determine how suitable is the model in the explanation of the behaviour of the exports and the domestic supply of storable products in the specific context of three commodities in Argentina. Up to this moment, this is the only empirical application of this model, which implies the additional complication of no comparable estimations. Therefore, with a lack of theoretical guidance of the signs of the coefficients, one needs to add the absence of other estimations to be used as a reference to compare those obtained here.

This implies that we need to rest on the intuition or previous knowledge to judge the economic sense and meaning of the estimations obtained. This is definitely unsatisfactory since the objective of a theory, such as the one developed in the first chapter, is the explanation of a particular phenomenon without referring to other *ad hoc* models or theories to explain it.

However, it is also true that every model originally developed had a time when its empirical suitability was not clear given that no earlier references existed to compare the results. The practical implication of this, in the light of this current exercise, is that this empirical application constitutes, probably, the first reference for future empirical applications in other contexts.

As the parameters are affected by the expected variances and covariances of the prices of the model, variations in their values will affect their values, making impossible the estimation. Consequently, we will make the assumption that the system has been already “launched” at some point in time, and the current expected variances and covariances are stable enough not to be seriously affected by the variation in prices. This is a crucial assumption but it is almost implicit in any model since any elasticity or parameter is, in the long run, a variable. Consequently, simplifying notation, the equations to estimate are

$$Q_t^e = \alpha_1^e P_t^e + \alpha_2^e P_t^d + \alpha_3^e P_t^q + \alpha_4^e P_t^f + u_t^e \quad (4.27)$$

$$Q_t^d = \alpha_1^d P_t^e + \alpha_2^d P_t^d + \alpha_3^d P_t^q + \alpha_4^d P_t^f + u_t^d \quad (4.28)$$

where Q_t^e and Q_t^d are the exported and the domestic supplied quantities, P_t^e is the export price, P_t^d is the domestic price, P_t^q is the price of the primary product, P_t^f is the futures price and u_t^e and u_t^d are error terms with the standard properties. These are similar to equations (1.20) and (1.21). Given the lack of reliable data about it, the demand for futures has been left out of the analysis.

The definitions of the coefficients for the export supply function (the coefficients of the domestic supply function are analogous) were given in Chapter 1 by,

$$a_1^e = \frac{d\sigma_P^2 + a_p k_{PPd}}{a_p\left[d(k_{PPd} + 2d + k_{PPe}) + a_p\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]}$$

$$a_2^e = \frac{-d\sigma_P^2 + a_p\delta}{a_p\left[d(k_{PPd} + 2d + k_{PPe}) + a_p\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]}$$

$$a_3^e = \frac{d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) - a_p(k_{PPd} + \delta + \lambda_e)}{\left[d(k_{PPd} + 2d + k_{PPe}) + a_p\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]}$$

$$a_4^e = \frac{-d\left(\sigma_{P,Pe} - \sigma_{P,Pd}\right) + a_p\lambda_e}{\left[d(k_{PPd} + 2d + k_{PPe}) + a_p\left(\sigma_P^2 m - \sigma_{Pe}^2\sigma_{P,Pd}^2 + 2\sigma_{Pd,Pe}\sigma_{P,Pd}\sigma_{P,Pe} - \sigma_{Pd}^2\sigma_{P,Pe}^2\right)\right]}$$

These definitions reveal the nature of the hedging strategy followed by the trader. It can be seen that part of the coefficient of the primary product price is repeated with the opposite sign in the coefficient of the future price, reflecting the hedging effect of future prices against the volatility of the price of the primary product. However, only a part of the primary product price effect is hedged, since speculation with the future price adds its effect. The separation result between output and hedging decisions, as seen in Chapter 1, does not hold.

Intuition suggests in this case that, during the estimation, it is expected that these two coefficients will have opposite signs, to reflect the hedging. However, it is possible that either both coefficients may have the same sign, by the effect of the speculation or because the operation in futures may be used to hedge against volatility in the export and domestic price if a cross hedging strategy is used. If this is the case, the coefficient of the future price should be smaller in absolute value than the coefficient of the primary product, given the composition of its parameter.

On the other hand, part of the coefficient of the export price is repeated in the coefficient of the domestic price, suggesting another component of the cross hedging strategy at the time of supplying in both markets. Intuition suggests that it is expected that the coefficient of the export price should be positive, and the coefficient of the domestic product be negative (a higher price in the domestic market should increase the supply in the export market). However, the possibility exists, depending on the expected variances and covariances, that this coefficient could be positive as well. This might be the case if, for example, the expected variance of the domestic price is high or the covariance between the domestic price and the input price is too high. In both cases, the trader might find it more convenient to increase the export price, even with increasing domestic prices.

On the other hand, even plausible assumptions made cannot guarantee the analytical identification of the sign of coefficients. If the coefficients of the export and domestic prices are positive and negative, respectively, the coefficients of the primary product and the future prices can still observe any sign depending on the expected variances and covariances in the definition of these coefficients. Therefore, only through an empirical exercise can we shed some light about the value of these coefficients.

It is important to remark that these specifications are not entirely in line with the equations of the model developed. In the original specification, conditional expectations of the export, domestic and primary product, together with the futures price, are the explanatory variables. Instead, in this specification, the variables appear in their current levels.

The reason for performing such a change lies in the definition of the conditional expectations on prices. The use of the expectation operator tends to smooth the variables and makes them insensitive to the latest and more relevant information. In particular, the latest values receive the same weight as the oldest ones. This makes the expected variable smoother, reflecting long-run movements, and comparatively insensitive to recent strong changes in the values of the original variable.

This might be solved partially by using a sample adjusting process where only the latest observations are considered effectively in the formation of the expected value. Such a "moving-average" procedure assigns all weight to the latest observation and removes from the memory past events so that the more recent events are more relevant. However, a question appears related to the precise definition of the length of the relevant process and on the weights assigned within this "relevant" period for the formation of expectations.

Whilst this may be addressed by a proper estimation of the length and the weights of the formation of expectations, the procedure is clearly *ad hoc*, particularly since it is unclear if agents form expectations in such a way. Moreover, no consideration exists for cases when agents wish to adjust their expectations when they are seen as incorrect.

On one side, using the standard expectations operator implies the smoothing of the variable and its estimation consequences as we have seen before. On the other side, assigning a different mechanism for the formation of expectations may not be theoretically compatible with the model presented and maybe inaccurate.

As a consequence, and understanding the implications of the change introduced in the equations, it is preferable to use the variables at their original levels rather than perform additional transformations. However, past realisations of the variables will eventually be included in the specification at the time of the estimation of the ECM as a requirement of the estimation procedure and not by the imposition of *ad hoc* elements.

On the other hand, for the purpose of identification of the supply function, it is necessary to consider the addition of variables that can help to locate the function in the price-quantity space. These variables might be of different nature but they must be associated to the supply side of the relationship. In this sense, variables associated to cost of production or commercialisation, weather variables such as rain that can be of importance in explaining the crop production, or variables associated to the competitivity of the activity may also help. Additionally, dummy variables that help to capture the deterministic seasonal behaviour of the supply also need to be considered.

To estimate the model presented above we require the quantities of export and domestic supply as well as the different prices involved. The quantities variables have been presented in the third chapter. Additionally, a deep analysis of seasonal unit roots using monthly data when variables are allowed to take zero values was performed, also considering the possibility of unknown structural breaks that may affect the unit root tests. From that analysis, we know the order of seasonal integration of the quantities exported and domestic supplies of those three products.

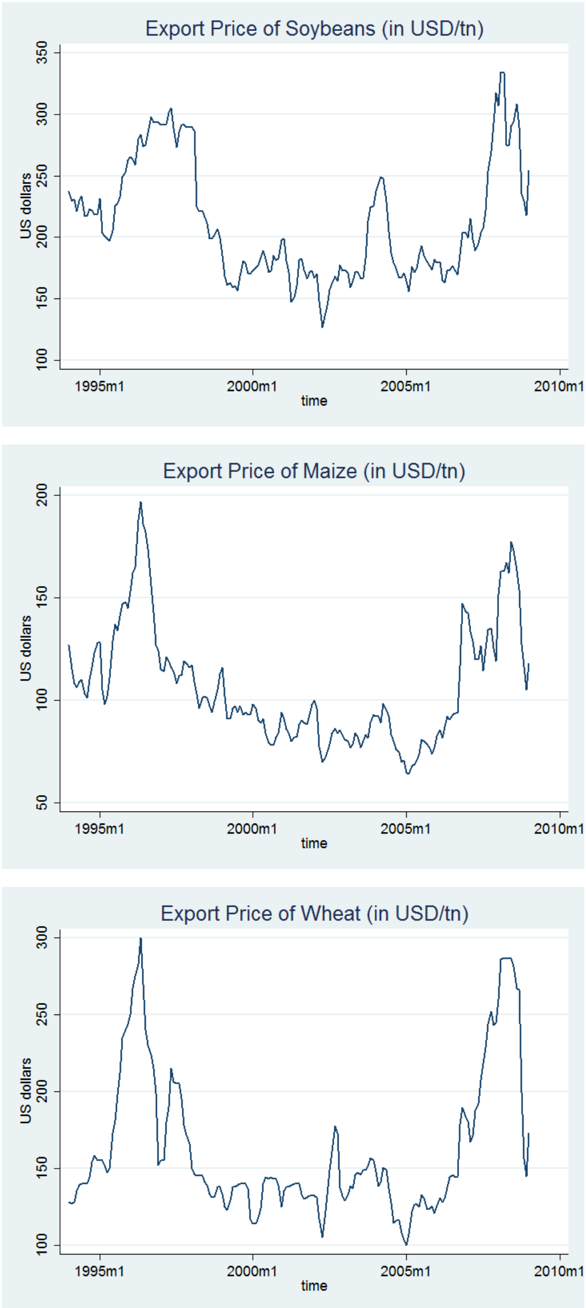
4.5.1. Prices of agricultural commodities

It remains to define and analyse the prices used in the cointegration analysis. The seasonal unit root analysis of the prices is performed in the next section. In the meantime, we will describe the variables used as well as discuss some general considerations. Unfortunately, not all prices required are available as reported series, which means that some of them have to be constructed using different assumptions about their compositions.

The export price is the FOB price in USD/t adjusted by the export tax. The export tax started to be applied in March 2002[32] [33]. This is an observed price and the export tax adjustment is necessary to keep compatibility with the evolution of the rest of the prices, since the effect of the export tax is immediate on the exported price but may have some delay on the rest of the prices. Export taxes affect, of course, the export price received by agents; however, it also affects, through arbitrage, the price received on domestic supplied products. A wedge would be introduced between the export and the rest of the prices if this adjustment were not made. Figure 4.1 presents the evolution of the export prices of maize, wheat and soybeans during the period under study. The distinguished moments of high prices can be seen in the late 1990s. There is another peak observed at the end of the period associated to the recent international prices escalade.

[32] Only soybeans had an export tax of 3.5% before that period. An additional duty of 20% was introduced in March 2002. This was increased to 30% in April 2002.
[33] Source: Ministerio de Agricultura, Ganadería y Pesca de la República Argentina.

Figure 4.1: Export price of Maize, Wheat and Soybeans



Source: Ministerio de Agricultura, Ganaderia y Pesca de la Republica Argentina

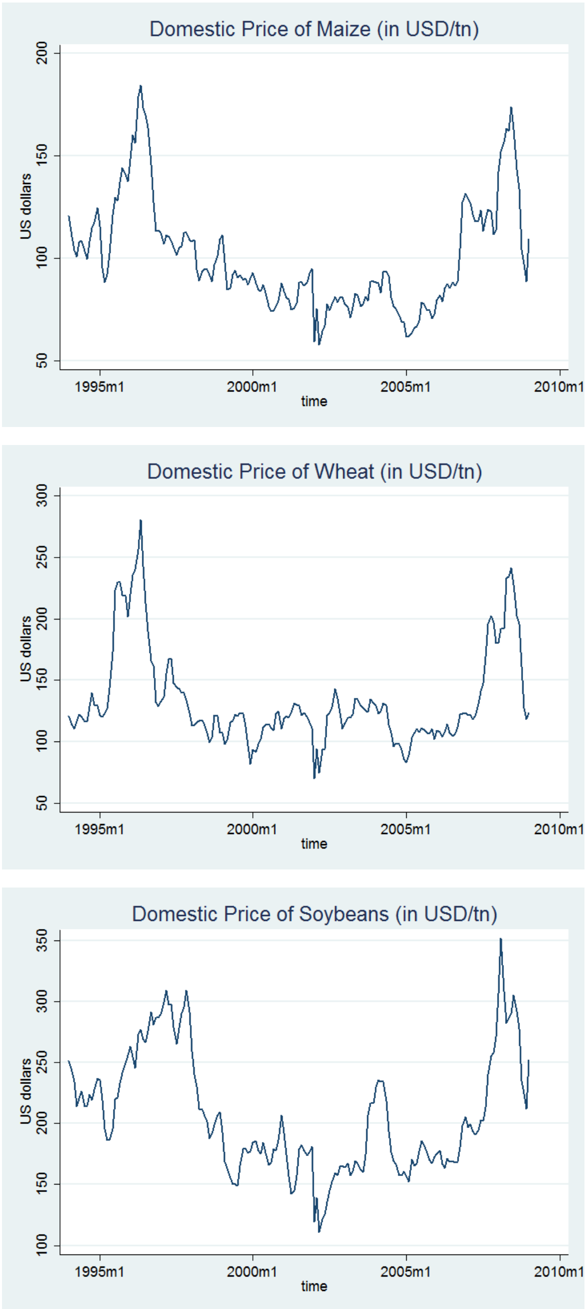
The domestic price for each product is the arbitrage chamber price in USD/t from the Buenos Aires Commodity Board of Trade. Prices of soybeans and maize are quoted at Rosario and prices for wheat are quoted at Buenos Aires. The domestic price comes from the daily observed traded spot positions paid by processors. Monthly prices have been obtained by taking averages of the daily values within a month. Figure 4.2 presents the evolution over the period studied of the domestic prices of maize, wheat and soybeans.

An observation and critique could be made on the temporal aggregation made of daily into monthly data for these prices after the discussion made in the previous chapter. The aggregation into lower frequency via averaging will smooth series and, consequently, will alter the data generation process. This will eventually affect the estimation of the cointegration relationship.

The estimation of the model with dependent and independent variables in their original frequency would avoid this problem. The main limitation comes from the fact that no data are available on daily exports and domestic supply of products, which makes it impossible to perform an analysis with data at different frequencies. However, whilst using monthly data in a seasonal cointegration context implies a higher level of complexity but still possible to perform, using a daily seasonal cointegration model is beyond the scope of this chapter. It is interesting to highlight that Andrade *et al*. (1999) and Tokihisa and Hamori (2001) have applied the HEGY approach to daily data in the analysis of stocks traded in the UK and Japan, respectively. However, they have applied this technique on daily data in the context of the analysis of seasonality within weeks and not on seasonal contexts such as the one presented here, where seasonality tends to be affected by elements with longer cycles.

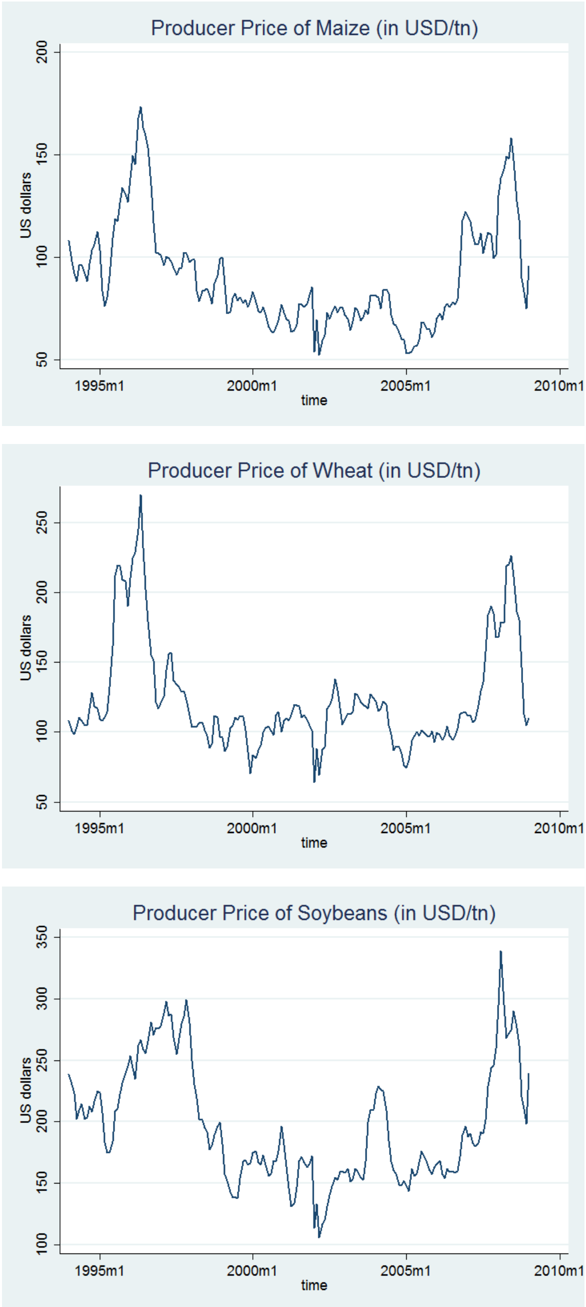
Producer price, or the primary product price, has been constructed since no data are available for the price of the producer at the farm gate. These prices are seen in Figure 4.3. Starting from the domestic price (the spot price); we subtract the transport cost to arrive at what would be a representative primary product price (frequently called "farm gate" price). We should have considered other costs (commissions, inspections, insurance premium rates, etc.), however, it is expected that they will not have too much variation in contrast to the transport costs that exhibit higher variability. Premium rates and commissions are expressed generally as *ad valorem* with respect to the price, and they will tend to observe the same variation of the price. Transport costs, in contrast, are affected by prices of fuel that tend to have higher variability and they also have their own seasonality (rail and road transport costs tend to rise at the time of harvest), giving the producer price a different variability from the domestic price. On the other hand, this treatment assumes that there is a wedge between the export and the domestic price, on one side, and the primary product price. In reality, that gap is narrower since the farmer receives the full price at the board of trade and he is the one that pays for the transport costs.

Figure 4.2: Domestic price of maize, wheat and soybeans



Source: Buenos Aires Commodity Board of Trade

Figure 4.3: Producer price of maize, wheat and soybeans



Source: Ministerio de Agricultura, Ganaderia y Pesca de la Republica Argentina

In constructing these series, we have identified what would be a representative producer. In a context of multiple producers and different distances between producers and the delivery place, this may be difficult to accept. An alternative may have been calculating different producer prices considering different distances and taking an average of them. However, this treatment is not substantially different from the one we are performing here. On the other hand, if producers' prices were available in different locations, it would be possible to obtain a more representative producer price. Nevertheless, in that case, we would have the price that we need. Instead, we have assumed that transport costs assume a short freight (20 km) by lorry/truck and long freight by train/truck (300 km)[34] and apply those distances to the cost per km/t. The cost per km/t differs depending on type of freight, the cost of short freights being proportionally higher. These distances are, on average, the distances that effectively exist between production areas and the different ports or industrial facilities. Moreover, the procedure we outlined here is a standard procedure generally applied by practitioners (agricultural engineers) when calculating indicative margins and costs of production of agricultural commodities.

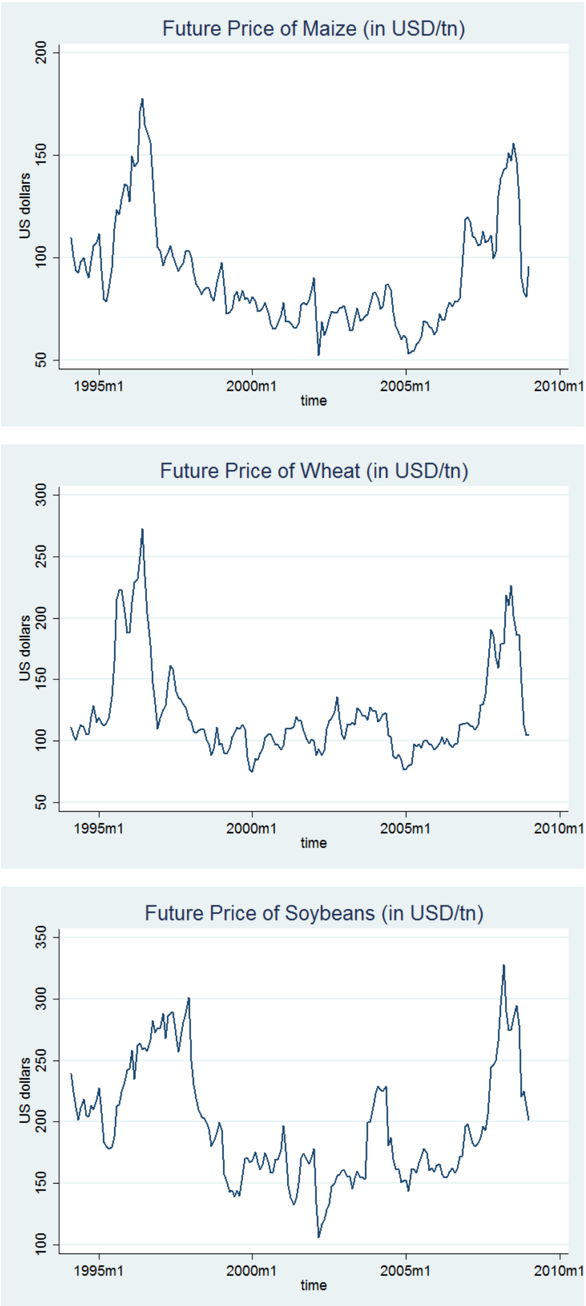
Futures prices are particularly problematic[35]. In some months, there are no futures prices. This is explained mainly by the non-existence of operations for a given position at that particular time. Markets tend to become particularly active when the time of the delivery approaches, or when there are changes in the price expectation. When no operations are recorded, it may imply that the expected price is similar to the future price, meaning that there are no gains to be made by speculating with them. Consequently, the future price should be equal to the expected spot price at delivery. Since we wanted to avoid the use of expectations on prices (replacing the missing future price with the expected spot price), we assumed that the future price replicated the price of the last period with trading positions. Figure 4.4 presents the future prices of maize, wheat and soybeans.

On the other hand, it is difficult to identify a particular period to highlight what the relevant future price is for a particular time of delivery. Futures markets for a particular position are generally opened several months in advance, so it is hard to identify if the export supply in a given month, for example, was made using the future price of the primary product for that position one, two or six months before. This does not imply that prices for a given position are available at any time during that period, as we have explained. Moreover, the existence of speculation complicates the analysis since futures markets operations are also affected by the changes in the expectation in the spot price that exist before the effective time of delivery.

[34] Source: Ministerio de Agricultura, Ganadería y Pesca de la República Argentina. Distances of short and long freights are of standard use in the calculations of costs of production and commercialization made by farmers.
[35] Source: Buenos Aires Commodity Board of Trade.

Given these difficulties, we have taken a simplifying approach and taken three different lead values of the future price in our analysis. This suggests that the future price in a given month was the future price for that time of delivery one, two or three months before. However, we will eventually only use the one that presents the best cointegration properties.

Figure 4.4: Future Price of maize, wheat and soybeans



Source: Buenos Aires Commodity Board of Trade

The number of observations may be problematic in our analysis. We are using 180 observations and the number of parameters to be estimated based on these observations will be large. Therefore, any inference based on them may have problems of power. In the previous chapter, we discussed extensively the issues related to the number of observations, as well as the length of the period under study. The same elements of the discussion can be translated to this chapter since we are using the same data.

Older data are not available and it comes from different sources that may present issues of compatibility with the data presented here. Moreover, older data may be affected by the presence of different commercialisation legislation (Grains Board) that makes older data irrelevant or incompatible with the analysis performed here. An extensive discussion on the different economic and agricultural policies followed in Argentina has been presented in the second chapter.

On the other hand, it is found that similar cointegration exercises have used a comparably similar number of observations. Kunst (1993) uses 64 observations in quarterly frequency for Finland in his analysis of macroeconomic models. Tiffin and Dawson (2000) use a sample of similar length to the one used here using monthly data. Martin-Alvarez, Cano-Fernandez and Caceres-Hernandez (1999), and McErlean et al. (2003) have used shorter samples. They have used 10 years of monthly observations in their respective studies.

A final note on the exogeneity of the variables is necessary to make. The model presented above relates quantities and prices; in particular, the direction of the effect seems to be unidirectional. We cannot exclude the possibility that the quantities supplied may have an effect on the prices. Argentina is an important supplier of the three commodities analysed and the possibility that some market power may exist, in such a way that the Argentine supply may have effects on the international price, cannot be ruled out. Of course, the domestic supply will have definite effects on the domestic price.

Consequently, the choice of the cointegration approach for the estimation of this model turns out to be particularly convenient. We could not have applied other more general dynamic models if prices were not exogenous. Of course, the possibility and probability that prices are indeed endogenous exists, but a violation of exogeneity will not affect the estimation procedure.

4.5.2. Seasonal unit root analysis on prices

Similarly, to the non-seasonal case, the necessary condition for seasonal cointegration is that series must be integrated of the same order. However, the condition is more specific as it requires that series under study must contain unit roots in exactly the same frequencies. Therefore, the first step in the seasonal cointegration, following Hylleberg, Engle *et al*. (1990), is the unit root testing.

By definition, prices cannot observe a zero value. Consequently, our analysis resembles the standard HEGY unit root tests. Critical values developed by Beaulieu and Miron (1993) for the monthly case can be applied. Nevertheless, as shown in the previous chapter, the critical values obtained for series that observe zero values can be used without loss of generality.

Table 4.1 presents the results of the HEGY analysis on prices. All prices reveal the presence of a zero frequency, standard or long run unit root. In the future price of maize one period ahead and in the future price of soybeans two periods ahead, it was not possible to reject a unit root in the biannual frequency (π). Additionally, a seasonal unit root was found in the future price of soybeans three periods ahead in the $\frac{\pi}{3}$ frequencies. Finally, unit roots could not be rejected in the export price of maize in the bimonthly and $\frac{5\pi}{6}$ frequencies. In the rest of the series, no evidence of additional seasonal unit roots is found.

Consequently, it will be only possible to find seasonal cointegration relationships if the quantities series analysed in the second chapter, and according to the model specification, present unit roots in the same frequencies. A close inspection of Table 3.9 (and Table 3.10 for the unknown structural break) in the previous chapter indicates that all quantity series present unit roots at the zero frequency. Whilst some unit roots have been found in some of the frequencies and series, they do not match the unit roots found in prices. Therefore, it is not possible to find seasonal cointegration between these series and only cointegration at the zero frequency (the standard cointegration) could exist. Additionally, as we have discussed in the previous chapter, the evidence on seasonal unit roots in the quantities is inconclusive.

However, the presence of structural breaks could substantially affect the results presented above. A structural break will make the HEGY approach lose power or find too many unit roots, as Perron and Vogelsang (1992) and Zivot and Andrews (1992) suggest. This means that structural breaks could disguise an otherwise stationary process as a unit root process. Therefore, if the researcher is unaware of the presence of structural breaks, he could identify false unit roots.

Of course, if the date of the break is known beforehand, the approach suggested by Perron (1989) for non-seasonal data of controlling for the two periods (before and after the break) at the time of performing the unit root tests can be applied. It is worthwhile to know that the critical values used to perform the inference are affected by the break, which means that those found by Perron (1989) can be used.

Table 4.1: HEGY test on commodity prices

Variable	Description	π_1	π_2	$F_{3,4}$	$F_{5,6}$	$F_{7,8}$	$F_{9,10}$	$F_{11,12}$	Lags	Model Specs.	B-G	B-J
pfwht_3	Future price of wheat 3 months before	0.43	-3.25	8.2	10.41	12.04	13.3	7.2	11, 23, 35, 34, 35	T	0.3367	0.0028
pfwht_2	Future price of wheat 2 months before	3.08	-4.56	20.19	25.14	19.64	17.89	15.11	1,12,13,24,25,31	T	0.4949	0.0009
pfwht_1	Future price of wheat 1 month before	3.02	-3.01	17.91	8.38	13.81	9.42	23.07	1-7		0.3664	0.0067
pfmaz_1	Future price of maize 1 month before	3.82	-2.26	15.96	16.31	9.16	9.67	9.64	1-12-13-24-25-31	T	0.6632	0.0037
pfmaz_2	Future price of maize 2 months before	4.85	-5.54	14.45	15.97	10.91	14.16	14.05	1-8-13	T	0.07	0
pfmaz_3	Future price of maize 3 months before	2.87	-3.13	12.74	29.47	7.67	15.15	12.73	9-10		0.2074	0.0016
pfsoy_1	Future price of soybeans 1 month before	3.97	-3.27	15.18	14.4	11.51	11.51	18.12	1-13		0.37	0.0001
pfsoy_2	Future price of soybeans 2 months before	2.49	-2.4	12.93	12.47	12.88	8.09	11.98	16-28-32	T	0.6528	0.0003
pfsoy_3	Future price of soybeans 3 months before	3.52	-3.24	9.83	11.59	5.58	12	9.38	10-11-12-22-28-32	T	0.201	0.0818
pewht	Export price of wheat	4.91	-4.71	27.92	27.73	19	17.74	30.17	13		0.255	0
pemaz	Export price of maize	1.34	-4.4	23.18	31.42	2.99	3.3	7.92	3,7,8,11,12,15,17,19,23,24,27		0.2283	0
pesoy	Export price of soybeans	5.1	-5	24.85	22.01	20.74	22.32	30.95	1,12,13,25,26,32,35	T	0.2083	0.0002
pqwht	Producer price of wheat	1.51	-3.8	18.52	19.9	36.18	7.9	11.51	14		0.3136	0.0047
pqmaz	Producer price of maize	3.32	-3.9	23.11	24.27	18.32	7.8	9.40	14	T	0.0129	0.0002
pqsoy	Producer price of soybeans	4.29	-3.96	15.03	18.41	10.52	13	20.22	12,23,24,31	T	0.2001	0.001
pdwht	Domestic price of wheat	3.29	-4.58	26.78	24.17	31.17	12.2	21.86	8		0.0487	0.0001
pdmaz	Domestic price of maize	4.37	-4.82	30.02	36.75	24.64	12.16	22.66	no lags		0.0244	0
pdsoy	Domestic price of soybeans	5.29	-5.44	19.28	24.13	19	16.76	25.12	12,23,24,29,31		0.6121	0.0009

Source: Own estimations

However, when the existence of the unit root is unknown to the researcher, its location must be estimated. Perron and Vogelsang (1992), Franses and Volgesang (1998) and Ghysels and Osborn (2001) have developed the technique for using the HEGY unit roots tests under the presence of unknown structural breaks. The procedure is well documented in those papers and has been explained extensively in the previous chapter.

Franses and Hobijn (1997) and Smith and Otero (1997) have tabulated critical values for testing seasonal unit roots under the presence of structural breaks for the quarterly case. The lack of applications of the HEGY tests under the presence of structural breaks is noteworthy. Even the latest applications found, such as Caporale, Cunado and Gil-Alana (2012), or the more complex analysis of seasonal unit roots tests under the presence of multiple structural breaks Tasseven (2008), have been made on quarterly data.

The only application made so far on seasonal unit roots under the presence of an unknown structural break on monthly data has been the one presented in the previous chapter, which also tabulated critical values for this case. Additionally, one could count the latest application to monthly data of the methodology of stochastic seasonality on panels of time series under the assumption of independence in the cross section dimension made by Kunst and Franses (2011). However, its association with the time series analysis is not direct given the additional cross-sectional dimension.

It is important to remember that this procedure only allows us to confirm if the roots found previously have not been affected by structural breaks or are spurious. Therefore, these tests will not change the conclusion that the quantities and price series are not seasonally cointegrated. This means that these tests are only performed for the sake of completeness on the analysis, and only applied to those series that presented unit roots in the case without a structural break.

Table 4.2 presents the results of the seasonal unit root tests under the presence of structural breaks when the standard technique of using multiple seasonal break dummies is used. In essence, two different tests are performed depending on how the break is identified. In the left panel we present the tests when the break is identified by minimising the $t_{\pi i}(T_b)$ and maximising the $F_{odd,even}(T_b)$ statistics over all possible break dates. This means the selection of the break when the statistics are least favourable to the null hypothesis, and in the left panel when the selection break date is based on the maximisation of the significance of the seasonal shift dummy variable. Further references on this procedure can be found in chapter 3.

Table 4.2: HEGY test on commodity prices under the presence of structural breaks using multiple dichotomic variables

	$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b)$ $i = 1,2$ $\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b)$				$\hat{T}_b = argmax\, F_\theta(T_b)$			
	pemaz	pfmaz_1	pfsoy_2	pfsoy_3	pemaz	pfmaz_1	pfsoy_2	pfsoy_3
0	-1.107 Mar-00	0.446 Apr-00	-0.192 May-99	-1.447 Mar-00	0.106	0.982	1.189	1.141
π	-4.486 Feb-04	-4.252 Dec-06	-4.092 Jan-05	-4.256 Mar-00	-2.717	-3.679	-3.299	-3.315
π/2	27.23 May-06	18.78 Dec-01	15.39 May-00	8.713 Dec-03	18	15.72	9.423	7.094
2π/3	22.2 Feb-05	18.92 Feb-06	22.84 Oct-03	23.05 Aug-03	14.79	18.92	15.47	13.53
π/3	8.6 Mar-06	22.45 Sep-05	12.29 Mar-97	8.07 Dec-03	6.429	20.49	12.26	3.679
5π/6	5.592 Sep-05	11.73 Jan-05	8.594 Feb-97	13.04 Mar-00	3.549	9.769	6.704	7.443
π/6	10.91 Nov-01	14.96 May-05	17.79 Jul-01	11.74 Aug-01	6.24	10.49	10.79	8.005
Break date					*Jan-06*	*Feb-06*	*Oct-07*	Dec-97

Source: Own estimations

According to the results, we are still not rejecting the null hypothesis of a unit root at the zero frequency in the four series analysed using both decision methods. The seasonal unit roots found in the export price of maize (pemaz) are still present when structural breaks are allowed for, as well as the root found in the future price of soybeans three periods ahead. Nevertheless, we can reject, under this framework, the presence of a unit root in the biannual frequency for the price of soybeans with two periods. Finally, we obtain a borderline case when considering the unit root in the future price of maize one period ahead, since the decision criteria indicate different conclusions.

Given the large number of parameters involved in the estimation of the equation used to perform the HEGY tests, a more parsimonious approach may be desired. Therefore, we conduct the HEGY test using a single dummy for the identification of the break, similar to how it has been done in the Chapter 3. Table 4.3 presents the results. There is no significant new evidence we can highlight with respect to the case presented above. The only new evidence seems to be that the borderline case we had in pfmaz_1 has disappeared, leaning towards the non-presence of a unit root in the biannual frequency for this variable.

Table 4.3: HEGY test on commodity prices under the presence of structural breaks using single dichotomic variable

	$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b)$ $i = 1,2$ $\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b)$				$\hat{T}_b = argmax\, F_\theta(T_b)$			
	pemaz	pfmaz_1	pfsoy_2	pfsoy_3	pemaz	pfmaz_1	pfsoy_2	pfsoy_3
0	0.405 May-97	2.324 Oct-99	0.971 Dec-98	1.963 Sep-98	0.561	3.798	2.507	3.111
Π	-4.435 Jul-98	-2.683 Dec-04	-2.548 Dec-97	-3.388 Aug-04	-4.276	-2.626	-2.232	-3.388
π/2	19.56 May-07	17.99 Sep-06	13.64 Nov-97	10.44 Mar-02	16.3	17.99	11.92	10.18
2π/3	26.05 Dec-06	18.48 Oct-06	15.55 Nov-97	12.22 Apr-99	25.91	18.42	15.03	10.86
π/3	3.15 Jul-98	12.41 Dec-04	12.44 Dec-97	5.987 Jan-98	2.013	9.927	9.667	5.656
5π/6	2.68 Aug-06	11.41 Dec-04	9.171 Nov-97	12.53 Aug-04	2.179	10.7	7.75	12.53
π/6	9.488 Jul-98	14.81 Dec-04	16.33 Nov-06	10.08 Nov-06	7.707	11.57	16.33	6.352
Break date					*Sep-06*	*Sep-06*	*Nov-06*	*Aug-04*

Source: Own estimations

The unit root tests performed in this section have allowed us to identify the seasonal unit roots in the prices. Since the seasonal unit roots found do not match the seasonal unit roots found in the previous chapter on the quantities exported and domestically supplied, the possibility of seasonal cointegration is ruled out. Only long-run cointegration could exist if sufficient conditions are met. This is the analysis we perform in the next section.

Moreover, the results obtained in the previous chapter suggest that a deterministic approach to address seasonality may be appropriate. This reduces even more the possibility of seasonal cointegration because the evidence on seasonal unit roots is not strong.

4.6. COINTEGRATION ANALYSIS

The following step in the cointegration analysis, following the Engle and Granger (1987) methodology and its extensions to the seasonal case, is the estimation of the cointegration relationships and testing of their residuals. The necessary condition for cointegration is that all series involved should be integrated of the same order. This also applies to the seasonal case.

Seasonal cointegration is evaluated in each of the frequencies when matching unit roots between dependent and independent variables can be identified. This requires the application of different transformations in order to remove the presence of the other seasonal unit roots. If, for example, the cointegration analysis is performed on the zero frequency and the series contained additional seasonal unit roots, the cointegration test is performed on the residuals of the regression given by equation (4.4) where both variables have been transformed by the application of the filter $\varphi_1(B)$. This transformation removes the effect of the additional seasonal unit roots and leaves the transformed variables only with the effect of the unit root under consideration. Therefore, since seasonal unit roots could not be ruled out in some cases considered in the previous chapter as well as in the unit root tests performed on the prices, in order to analyse cointegration at the zero frequency, it is necessary for those series to apply the above-mentioned filter in order to remove the effect of these additional roots.

The applications made by Engle *et al*. (1993) analysed seasonal cointegration based on the condition that every variable, in their quarterly case, has four different unit roots. However, the case where variables contain unit roots at different frequencies is more general. In this case, only cointegration will be possible where there is a match in the unit roots at the frequencies of the respective variables. In this sense, Bohl and Sell (1998), Martin-Alvarez *et al*. (1999), Bohl (2000), Tiffin and Dawson (2000), Hamori and Tokihisa (2001), Ouerfelli (2008), and Binet and Zaied (2011) have limited the cointegration analysis to the frequencies integrated of the same order.

Hamori and Tokihisa (2001) made an interesting application on their study the relationship between the money balances, real GDP and the interest rate in Japan. Whilst the first two series presented unit roots at the zero and other seasonal frequencies, the latter only presented unit roots at the zero frequency. Therefore, in their cointegration analysis, they only transformed the first two variables to remove the seasonal unit roots and left the interest rate unchanged. This was because only two variables contain seasonal unit roots whilst the rest only the zero frequency.

As we have seen, based on the unit root tests presented above, we have rejected the possibility of seasonal cointegration in our example since it was impossible to find matching seasonal unit roots between quantities and prices. Moreover, the results obtained on the seasonal unit root tests performed in the previous chapter were not strong. Therefore, we will analyse just the cointegration in the long run or at the zero frequency. Following the Engle and Granger (1987) two-step approach, in order to analyse the cointegration at the zero frequency it is necessary to estimate a model using equation (4.4). This equation can be augmented by the presence of constants, trends and seasonal dummies. If cointegration cannot be rejected at the zero frequency between the variables, the residuals of that regression $\omega_t^{(1)}$ should be stationary or I(0). The evaluation of stationarity can be made by using the Augmented Dickey-Fuller (ADF) test on those residuals by fitting

$$\Delta\tilde{\omega}_t^{(1)} = \rho\tilde{\omega}_{t-1}^{(1)} + \sum_{i=1}^{p} \sigma_i \widetilde{\Delta\omega}_{t-i}^{(1)} + \tau_t \qquad (4.29)$$

and test the significance of the coefficient ρ, using the critical values developed by Engle and Granger (1987). As usual, the equation could also contain a drift and trend to represent more accurately the true data generation process. Eventual testing for seasonal cointegration at a given frequency can be performed in a similar fashion, using the appropriate equation presented previously and running the ADF on their residuals.

Table 4.4 presents the results of our empirical example. The cointegration analysis for each of the regressands (the quantities exported and domestic supplied) and the different prices is computed. Additionally, the joint cointegration between the quantities and the prices is presented at the bottom. According to equations (4.27) and (4.28), this means that the cointegration is between the quantities and the respective prices. Only for the cointegration relationships involving more than two variables, Column 3 provides the identifier of the relationship. Column 4 indicates the specification used in the ADF unit root test of the residuals of the cointegration equation in terms of the inclusion of trend or constant terms. In addition, seasonal dummies have been included in all cases. Column 6 presents the t-ADF statistic, column 7 the number of lags used in the unit root equation of the residuals of the cointegration equation. The Durbin-Watson coefficient for first order autocorrelation is in column 8; in column 9, the Breusch-Godfrey statistic for the test of higher order autocorrelation in the residuals; and column 10 presents the Bera-Jarque test of the normal skewness and kurtosis. Seasonal dummies have in included in all cases.

Table 4.4: Pairwise and multiple cointegration tests

Regressand	Regressor	Identifier	Model	Adj R2	t-ADF	lags	DW	B-G	B-J
Qemaz (z1t)	pdmaz	1		0.576	-2.250	29	1.986	2.082	7.788
	pqmaz	2		0.581	-2.224	29	1.990	1.719	7.570
	pemaz (w1t)	3	T,C	0.747	-4.165***	24	1.886	5.564	32.550
	pfmaz_1	4		0.573	-2.436*	29	1.903	5.615	11.330
	pfmaz_2	5		0.579	-2.239	29	1.983	1.858	8.005
	pfmaz_3	6		0.577	-2.244	29	1.984	1.641	7.838
Qewht	pdwht	7	C	0.988	-2.670*	31	2.006	19.280	20.080
	pqwht	8	C	0.987	-2.770**	31	2.009	18.750	16.560
	pewth	9	T,C	0.446	-3.419**	24	1.978	18.840	4.436
	pfwht_1	10	C	0.983	-3.282***	24	2.050	13.800	10.960
	pfwht_2	11	C	0.975	-3.044***	24	2.009	13.620	5.560
	pfwht_3	12	C	0.972	-2.599*	24	2.126	14.270	11.990
Qesoy (z1t)	pdsoy	13		0.468	1.118	12	1.905	11.570	17.040
	pqsoy	14		0.467	1.111	12	1.907	9.398	17.470
	pesoy	15		0.498	0.759	16	1.926	24.020	21.920
	pfsoy_1	16		0.423	0.688	12	1.965	34.500	8.005
	pfsoy_2 (w1t)	17	C	0.562	-3.276***	16	1.962	6.083	19.070
	pfsoy_3 (w1t)	18	C	0.550	-3.290***	12	1.963	5.469	15.960
Qdmaz	Pdmaz	19		0.320	-1.833	23	1.935	2.098	11.830
	Pqmaz	20		0.298	-1.599	23	1.940	2.720	10.780
	pemaz (w1t)	21		0.298	-1.687	23	1.936	2.850	11.200
	pfmaz_1 (w1t)	22		0.315	-1.711	23	1.937	1.988	13.590
	pfmaz_2	23		0.316	-1.579	23	1.945	1.915	7.594
	pfmaz_3	24		0.311	-1.576	23	1.942	2.790	10.140
Qdwht	Pdwht	25		0.584	-1.981	28	2.020	0.326	14.020
	Pqwht	26		0.583	-1.873	28	2.018	0.287	14.630
	Pewht	27		0.599	-2.572*	28	2.047	0.643	20.810
	pfwht_1	28		0.584	-1.881	28	1.996	0.915	22.090
	pfwht_2	29		0.570	-1.954	28	1.996	0.867	19.720
	pfwht_3	30		0.560	-1.971	28	1.989	0.786	26.130
Qdsoy (z1t)	Pdsoy	31	T,C	0.516	-3.354**	28	1.965	2.582	7.599
	Pqsoy	32	T,C	0.518	-3.338**	28	1.961	2.697	7.648
	Pesoy	33	C	0.517	-2.684*	30	1.934	6.148	25.560
	pfsoy_1	34	T,C	0.408	-3.205*	28	1.965	12.300	14.380
	pfsoy_2 (w1t)	35	C	0.519	-2.922**	28	1.938	9.285	20.680
	pfsoy_3 (w1t)	36	C	0.571	-1.539	32	1.997	7.982	10.280

Table 4.4: Pairwise and multiple cointegration tests

Regressand	Regressor	Identifier	Model	Adj R2	t-ADF	lags	DW	B-G	B-J
Qemaz (z1t)	pdmaz, pqmaz, pemaz (w1t), pfmaz_1	37	T,C	0.533	-3.877***	35	2.082	0.230	28.730
	pdmaz, pqmaz, pemaz (w1t), pfmaz_2	38	T,C	0.441	-4.234***	25	2.033	0.090	14.200
	pdmaz, pqmaz, pemaz (w1t), pfmaz_3	39	T,C	0.272	-3.324**	12	2.024	4.366	14.230
	pdmaz, pqmaz	40	T,C	0.499	-3.120*	29	2.038	2.541	13.970
Qewht	pdwht, pqwht, pewht, pfwht_1	41	T,C	0.494	-3.997***	24	1.999	13.850	7.744
	pdwht, pqwht, pewht, pfwht_2	42	T,C	0.478	-3.884***	24	2.003	13.840	6.838
	pdwht, pqwht, pewht, pfwht_3	43	T,C	0.474	-3.844***	24	1.987	12.790	7.125
	pdwht, pqwth	44	T,C	0.469	-4.188***	24	1.989	22.110	5.750
Qesoy (z1t)	pdsoy, pqsoy, pesoy, pfsoy_1	45	C	0.593	-2.668*	36	1.869	7.315	18.050
	pdsoy, pqsoy, pesoy, pfsoy_2 (w1t)	46	C	0.300	-3.007**	12	1.978	2.639	4.045
	pdsoy, pqsoy, pesoy, pfsoy_3 (w1t)	47	T,C	0.293	-3.221*	12	1.977	4.550	6.737
Qdmaz	pdmaz, pqmaz, pemaz (w1t), pfmaz_1(z1t)	48	C	0.409	-2.448*	24	2.005	1.191	9.231
	pdmaz, pqmaz, pemaz (w1t), pfmaz_2	49		0.409	-2.392	24	2.008	4.710	7.519
	pdmaz, pqmaz, pemaz (w1t), pfmaz_3	50		0.431	-2.324	24	2.004	1.623	8.056
	pdmaz, pqmaz, pfmaz_2	51		0.361	-2.921**	24	1.935	6.065	14.260
	pdmaz, pqmaz, pfmaz_3	52		0.427	-2.390	24	1.988	2.588	4.592
Qdwht	pdwht, pqwht, pewht, pfwht_1	53	T,C	0.654	-4.46***	30	1.995	0.890	16.000
	pdwht, pqwht, pewht, pfwht_2	54	T,C	0.659	-3.967***	30	2.045	2.279	15.290
	pdwht, pqwht, pewht, pfwht_3	55	T,C	0.635	-3.672**	28	2.064	2.814	20.320
	pdwht, pewht, pqwth	56	T,C	0.632	-3.435**	28	2.068	3.255	15.890
Qdsoy (z1t)	pdsoy, pqsoy, pesoy, pfsoy_1	57	C	0.251	-3.094**	18	1.962	18.090	11.290
	pdsoy, pqsoy, pesoy, pfsoy_2(w1t)	58	T,C	0.286	-3.287**	28	1.913	8.288	13.920
	pdsoy, pqsoy, pesoy, pfsoy_3 (w1t)	59	T,C	0.284	-3.599**	28	1.884	9.005	13.130
	pdsoy, pqsoy	60	T,C	0.511	-3.445**	30	1.982	3.031	8.562

Note: *** Significance at 1%. ** Significance at 5%. * Significance at 10%. Seasonal deterministic dummies have been included in all the cases. T stands for trend and C for constant.

Source: Own estimations

Table 4.5 presents the definitions of the variables used. When (z1t) or (w1t) appear next to the variable, it means that particular variable received a transformation to remove the incidence of additional unit roots, and only contains the zero frequency unit root. The non-transformed variables only contain the zero frequency unit root.

Table 4.5: Variable definitions

Variable Name	Description
Qemaz	Exported quantity of maize
Qewht	Exported quantity of wheat
Qesoy	Exported quantity of soybeans
Qdmaz	Domestic supply quantity of maize
Qdwht	Domestic supply quantity of wheat
Qdsoy	Domestic supply quantity of soybeans
Pdmaz	Domestic price of maize
Pemaz	Export price of maize
Pqmaz	Producer price of maize
pfmaz_1	Future price of maize 1 month before
pfmaz_2	Future price of maize 2 months before
pfmaz_3	Future price of maize 3 months before
Pdwht	Domestic price of wheat
Pewht	Export price of wheat
Pqwht	Producer price of wheat
pfwht_1	Future price of wheat 1 month before
pfwht_2	Future price of wheat 2 months before
pfwht_3	Future price of wheat 3 months before
Pdsoy	Domestic price of soybeans
Pesoy	Export price of soybeans
Pqsoy	Producer price of soybeans
pfsoy_1	Future price of soybeans 1 month before
pfosy_2	Future price of soybeans 2 months before
pfsoy_3	Future price of soybeans 3 months before

As suggested by Perron (1988), we follow a sequential approach for the ADF test. We first estimate the testing equation including a drift and trend. If the null hypothesis of a unit root cannot be rejected, we move to a model including only a drift and, in case the null hypothesis cannot be rejected, we estimate the model without those terms and the conclusion of the test is reached in this instance. Of course, if the null is rejected in any of the previous steps, the test procedure concludes.

In order to achieve the white noise properties of the residuals of the estimation of equation (4.29), lags of the dependent variable are added. In order to identify the appropriate lag length, we have followed the general-to-specific approach by selecting the lag length using a 5% level of significance threshold. In general, and as we have done in the previous chapter, the level of significance used is 10%. However, by adopting a more stringent criterion, we tried to obtain a shorter lag length. Nevertheless, and despite the intention, the lag lengths found are relatively long that reduces the power of the tests by the reduction in the number of usable observations.

Alternatives for the definition of the lag length could have been used. As discussed in the previous chapter, the AIC/BIC could have been applied to determine lag length by selecting the model with the best fit. However, this does not guarantee that the residuals will also have the desired properties, and despite the fact that both criteria penalise the addition of explanatory variables, it was observed that the AIC/BIC criterion identified a similar lag length to the ones found here.

The hypothesis of cointegration could not be sustained in many of the pairs of variables analysed (the test could not reject the null of a unit root in the residuals). Only a few cases can be highlighted (notably the cases of qewht and qdsoy) where we can conclude that the variables involved might present a long run cointegration relationship.

More interesting for the purposes of this chapter are the results where cointegration between each quantity variable and the four prices according to the model developed is analysed. As we have seen, we have played with three definitions for the future prices. Therefore, we have chosen for the joint cointegration analysis those definitions that performed best in the pairwise analysis. When the previous criterion was inconclusive, we favoured the three lead definitions, given that, at the time of delivery, future prices and spot prices tend to be similar.

Additionally, we have also considered some intermediate cases and evaluated the possibility of cointegration between the quantity and some of the price variables. The cases have been selected via the analysis of the significance of the price variables in the full cointegration relationships. This implies the coverage of almost all the possible cointegration relationships between the quantities and the prices.

It is important to remark that when more than two variables are analysed, the D-F critical values are not valid, and those constructed using the response surface in MacKinnon (1991) and MacKinnon (2010) must be used. These critical values vary with the number of variables included in the regression equation used to obtain the residuals to be tested. When the number of cointegration variables is two (one explanatory variable in the regression), the D-F critical values and those obtained by MacKinnon (1991) do not differ. However, when more than two variables are included, the D-F critical values tend to reject more frequently the null of no cointegration. Therefore, we have compared the t-statistic of the cointegration test using the appropriate critical values depending on the number of explanatory variables. Table 4.6 presents the critical values for four explanatory variables used in the inferences[36]

Based on the MacKinnon critical values, cointegration is rejected in almost every one of the multiple variable cases considered, implying that there is no long run relationship between each quantity and the prices presented in the postulated equations. The only exception seems to be the quantity of domestically supplied wheat and their respective prices where we could narrowly reject the null of no cointegration. If, on the other hand, the DF critical values are used, or those that have been tabulated in Chapter 3, we could have rejected the null of no cointegration and conclude that there seems to be a long-term relationship between the quantities and the prices suggested.

In the intermediate cases analysed, no cointegration relationships are found. Only the case of the exports of wheat and the domestic supply and producer prices (relationship 44) seem to present some cointegration relationship based on the MacKinnon critical values for relationships with three independent variables.

The results in terms of the properties of the residuals demand some caution on the results. Durbin-Watson test of autocorrelation tends to reject the possibility of autocorrelation (For K=2, upper limit is 1.76. For K=5, upper limit is 1.79.). The Breusch-Godfrey test of higher autocorrelation tends, in general, to reject the null of higher order autocorrelation with some exceptions. This means that, with some caution, it can be said that residuals are not correlated.

More problematic are the results in terms of the normality of residuals. Whilst in general the null of normal skewness and normal kurtosis is upheld by the Bera-Jarque test, there are some cases where it rejected this null in favour of the alternative. This means that the results of the cointegration test (the ADF test) performed might be affected by the

[36] The response surface used to obtain the critical values is given by $CV = \beta_\infty + T^{-1}\beta_1 + T^{-2}\beta_2$; where T the sample size.

violation of the assumptions of this test. Consequently, the conclusions with respect to the cointegration of some of the relationships presented here might not be sustained.

Table 4. 6: Critical values for cointegration tests based on MacKinnon (1991)

Level	β_∞	β_1	β_2	CV
4 independent variables				
No trend				
1%	-4.649	-17.188	-59.3	-4.747
5%	-4.1	-10.745	-21.57	-4.16
10%	-3.834	-9.188	-4.85	-3.886
With trend				
1%	-4.97	-22.504	-50.22	-5.096
5%	-4.429	-14.501	-19.54	-4.511
10%	-4.147	-11.165	-9.88	-4.21
3 independent variables				
No trend				
1%	-4.298	-13.79	-46.37	-4.376
5%	-3.743	-8.352	-13.41	-3.790
10%	-3.452	-6.241	-2.79	-3.487
With trend				
1%	-4.668	-18.492	-49.35	-4.772
5%	-4.119	-12.024	-13.13	-4.186
10%	-3.834	-9.188	-4.85	-3.885
2 independent variables				
No trend				
1%	-3.9	-10.534	-30.03	-3.959
5%	-3.337	-5.967	-8.98	-3.370
10%	-3.046	-4.069	-5.73	-3.069
With trend				
1%	-4.326	-15.531	-34.03	-4.413
5%	-3.781	-9.421	-15.06	-3.834
10%	-3.496	-7.203	-4.01	-3.536

Source: MacKinnon (1991)

On the other hand, there exists the possibility that a pair of variables whose hypothesis of cointegration has been rejected, might eventually have cointegrated with another variable when considered altogether, as Enders (2010) suggests or multicointegration happens. Whilst the Granger representation theorem may not hold in this case, it is a possibility that should not be ruled out. However, we have found that variables that cointegrated in the pairwise analysis have failed to cointegrate when they were considered as part of a more complex relationship. This is not exactly the multicointegration case presented and suggests some problems in the procedure.

Additionally, it is important to consider that whilst the use of the MacKinnon critical values is recommended by several books specialised in time series analysis, such as Ghysels and Osborn (2001), Harris and Sollis (2003) and Enders (2010). However, empirical applications made by Hamori and Tokihisa (2001), Hasan (2011) and particularly Engle, Granger and Hylleberg *et al*. (1993) have used the standard D-F critical values to test unit roots at the zero frequency. This means that there is no clear recommendation against the use of the standard D-F critical values, which adds additional uncertainty about the way of proceeding and the results presented here.

Nevertheless, as we have seen, when we used the DF critical values, the tone of the conclusion changes and, with some confidence, we can say that, when considered in the model developed in Chapter 1, the variables effectively cointegrated. This suggests that the variables considered would seem to present, with some nuances and despite some violations in the assumptions, some evidence of a long-term relationship that could back the model developed in Chapter 1. Therefore, we will continue with the analysis and try to estimate the ECM in order to confirm the results obtained. The estimation of these models would also help to confirm or reject the existence of a cointegration relationship between the variables proposed.

4.7. ERROR CORRECTION REPRESENTATION

In Chapter 1, a theoretical model was developed that explains the export and domestic supply decision based on prices in both markets, input and futures prices. The objective of this section is to determine the existence of an empirical long-term relationship between these prices and the quantities that can provide validation to this model. Therefore, we will focus our attention on the analysis of the six different supply functions (three export supply and three domestic supply functions) and the effects of the prices on them given by equations (4.27) and (4.28). Therefore, we will focus on the "complete" specifications and not on the pairwise relationships.

Everything we have done so far was the necessary pre-estimation testing in order to determine the most appropriate technique to be employed. The following step in the Engle-Granger methodology is the estimation of the Error Correction Model. We have discussed some of its advantages and disadvantages before, and, at the same time, kept in mind that the conclusion reached about cointegration was not particularly strong.

We have seen in the previous section that the evidence of cointegration varies according to the critical values used. In addition, in some cases, there seems to be some violations of the statistical assumptions (particularly normality). If this assumption is violated, the inference made about the parameters of the equation of the ADF test on the residuals may not be valid. This would affect any conclusions about cointegration based on any critical values. Consequently, given the violation of the normality assumption, the conclusions about cointegration remain open. Whilst this would demand either a special treatment to solve the normality issue (through alternative estimation techniques that do not require the assumption of normality) or directly find an alternative estimation approach (since cointegration cannot be sustained), we will continue with the analysis by estimating the ECM, paying, for the moment, no attention to these aspects: they will reappear later.

Our error correction model will have, for the quantity of exported maize, this general form

$$\Delta_{12}Qe_t^M = \beta_0 + \sum_{i=1}^{l} \Phi_i \Delta_{12}Qe_{t-i}^M + \sum_{i=0}^{P} \delta_i \Delta_{12}Pe_{t-i}^M + \sum_{i=0}^{q} \psi_i \Delta_{12}Pd_{t-i}^M + \sum_{i=0}^{m} \varphi_i \Delta_{12}Pq_{t-i}^M + \sum_{i=0}^{n} \theta\Delta_{12}Pf_{t-i}^M + \beta_2 \widetilde{\omega}_{t-1}^{(1)} + \varepsilon_t \tag{4.30}$$

with similar specifications for the other dependent variables analysed. Where $\widetilde{\omega}_{t-1}^{(1)}$ are the residuals of the cointegration equation given by equation (4.29), or similarly depending on whether a drift or a trend was included, as we have seen in the previous section, β_2 is the coefficient of the error correction term and reflects the speed of adjustment of any deviation from the long term relationship between the variables.

Φ_i are the coefficients of the autoregressive structure. As Enders (2010) suggests, if all these elements are zero, there is no possible error correction representation because $\Delta_{12}Qe_t^M$, in our case, is not affected by the previous period's deviation from the long run path. This suggests an ultimate test for non-cointegration at the time of fitting the ECM. On the other hand, it is expected than the effect of the parameters of the autoregressive structure should decrease as time passes. Effectively, the explanatory effect of past values of the dependent variable should fade away.

The lag structure of the explanatory variables presents a practical problem. When only one variable is lagged, the approach is to estimate a sufficient long lag structure and pare down until the last lag is significant. This is the general-to-specific approach explained and applied before. However, in the cases under analysis each explanatory variable (as well as the dependent variable) presents its own lag structure making this approach hard to apply. The standard procedure is to apply the Akaike Information Criterion (AIC) to identify the most suitable model.

Nevertheless, an additional complication is added in our case with four explanatory variables. The quantity of potential specifications to analyse grows to the power of five. Considering the possibility of a maximum of four lags per variable and five variables (four prices and the quantity variable), this entails the estimation of 1,024 different specifications. If we consider the possibility of five lags for each variable, this would entail 3,125 models to analyse. This complicates the analysis as well as demands substantial computation time that grows at a similar rate as the quantity of lags.

We have limited our analysis to four lags per variable and chosen the most appropriate model using the AIC. This implies that there may be cases where the appropriate lag length could be longer than expected and the correct specification may be outside the set of specifications considered. One possibility to address this issue could be adding lags to those variables that, at the lag four, are significant. This means including additional lags to those variables that "require it". However, the other coefficients are sensitive to the lag length of the other variables since, at the end; the specification of the model is changed. On the other hand, the inclusion of additional parameters to estimate would reduce the number of observations available for estimation, which would reduce the power of the tests.

The most appropriate treatment would be to choose the right model from a larger set of potential models, whilst the problem of lower power remains, we would have the opportunity to be closer to the most suitable model. Unfortunately, as mentioned, this demands even higher computational time and the analysis would definitely become more complicated. Therefore, we will keep the limit in the set of specifications and choose the most suitable from this set.

The standard specification of the ECM demands the inclusion of a constant. We have run the model including and excluding the constant term for every lag length. In all cases, as we will see, the constant term has resulted to be insignificant and its exclusion has almost not altered the results. This means that the appropriate estimation strategy is to estimate the model including the constant, verify its significance and, in case it proves to be not significant, estimate the model without the constant for that particular specification. Seasonal dummies have been included in all specifications with the objective of controlling for the deterministic seasonality.

4.7.1. Identification variables

Additionally, for identification of the supply functions, we have used a series of variables that, in addition to the prices proposed, can also affect the supply and help to locate the function in the price-quantity space. They include multiple costs variables associated to agricultural production (for example, the index of the wholesale price of fertilizers, pesticides and insecticides, diesel, agricultural machinery and land); a variable associated with the general economic climate or competitivity (such as the real exchange rate), and a variable associated with the weather (such as the average monthly rain).

The fertilizer, insecticides and pesticides, and the agricultural machinery price indices correspond to Chapters 2412, 2421 and 2921 of the INDEC – Wholesale Basic Price Index (Índice de precios básicos al por mayor). They have been adjusted to US dollars and its base changed to 1999. The US dollar price adjustment has been made in order to avoid the complications associated with the rise in the nominal exchange rate after the devaluation that occurred in 2002. Moreover, as was discussed in Chapter 2, past events associated to hyperinflation and frequent devaluations had led agents to base investment and other economic decisions on the more stable prices in US dollars.

The Secretariat of Energy's average observed petrol station price for Diesel (Gas-Oil) for agricultural purposes in the Buenos Aires province (adjusted by the US dollar price) has been used to calculate the price index. The inclusion of this price is associated to the intense use of this fuel in production (machinery is used intensively in the planting, harvesting and in the application of fertilisers and pesticides), as well as in the transport from the farm to the different points of stockpile and demand (ports and factories).

The price of land is another important cost that might affect the production decision. Although it is possible (and frequent) to rent the land, the price for rent tends to be closely related to the selling price. Therefore, either price could be used to identify the function. Land prices correspond to the average price per hectare of land observed in the Núcleo

Zone (North of Buenos Aires Province and south of Santa Fe provinces) and they were originally denominated in US dollars. These prices have been obtained from private sources, particularly from the Asociación Argentina de Consorcios Regionales de Experimentación Agricola (Series de Precios Agropecuarios) and from Compañía Argentina de Tierras S. A.

As it was discussed in Chapter 2, experiences in contexts of high inflation and frequent devaluations have generated the general practice of denominate domestic transactions in US dollars. Consequently, changes in the nominal exchange rate tend to have minimal effects in production and they may be of little use in the identification. However, a general measure of profitability or competitivity that could address the relative changes between domestic and international prices is desired. Therefore, the real exchange rate is suggested to address the goal. Moreover, the effects of competitive real exchange rates in output and economic activity in general have been well discussed in Frenkel and Rapetti, (2008) and Arslan, Rapetti and Skott, (2012). The index of the real exchange rate has been constructed using Banco Central de la Republica Argentina's nominal US Dollar/peso price, INDEC – Consumer Price Index Base 1999 (Índice de Precios al Consumidor), and US Bureau of Labor Statistics – Consumer Price Index.

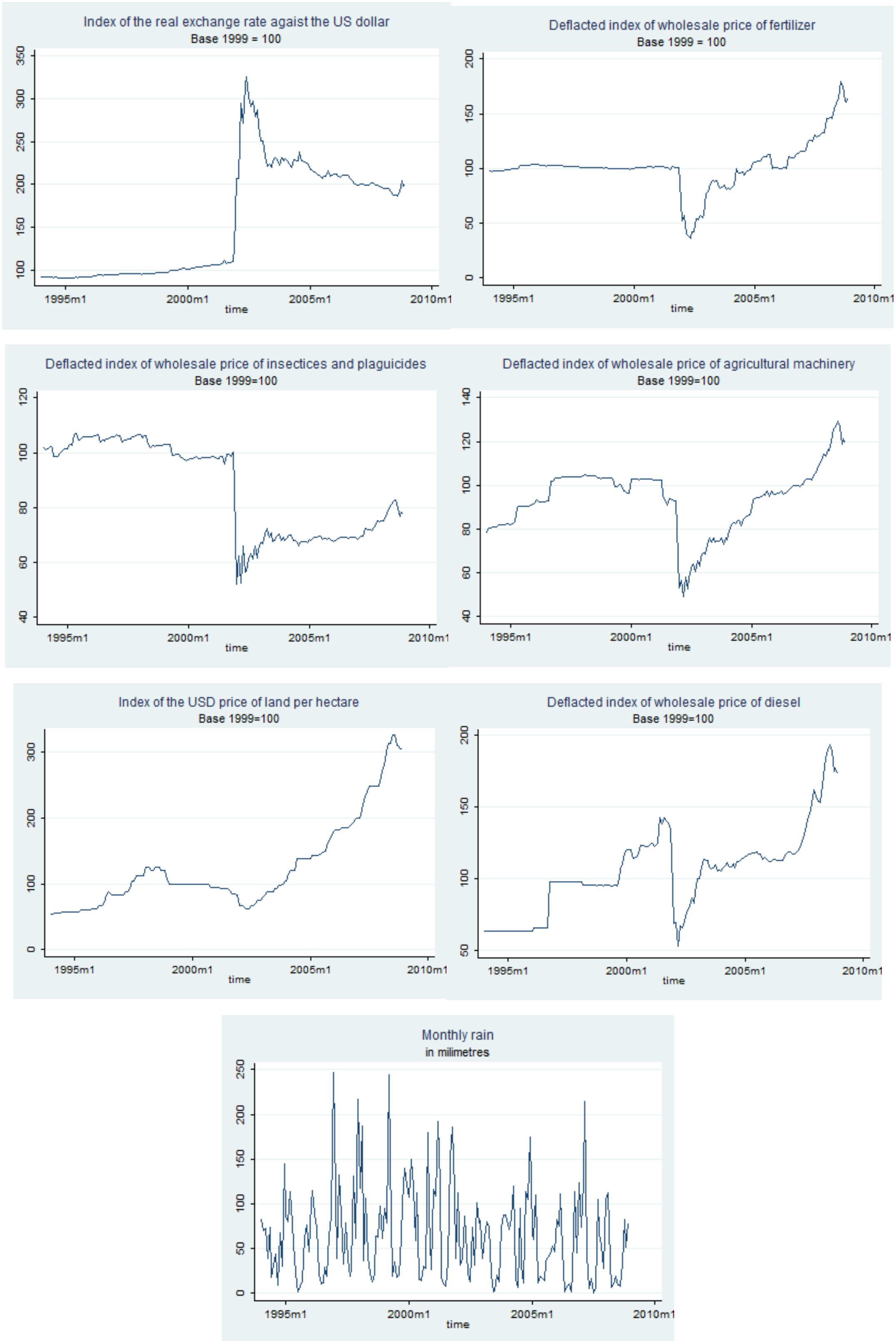
Additionally, variables associated to climate can be used to identify the functions. Rain, temperature and heliophany will notably affect production; however, only rain can introduce some variation in production as the other variables tend to be more associated to long-term cycles, and they tend to be more associated to the deterministic pattern of seasonality. In fact, either temperature or heliophany can replace seasonal deterministic dummy variables given their strong seasonal nature. Therefore, only rain has been considered as a viable variable to perform the identification of the supply functions. Average monthly rain has been calculated through the monthly sum of the average daily rain observations in multiple pluviometres located in La Pampa, Santa Fe, Buenos Aires and Cordoba provinces. The Instituto Nacional de Tecnología Agropecuaria (INTA) has collected these series. Figure 4.5 presents the plots of the seven variables.

The definition of rain as an identification variable presents a particular problem in a country the size of Argentina as rain patterns tend to be very different from one region to another. Whilst some areas could be suffering from a draught, other areas could have excessive rain that could complicate the activities. The variable used in the identification presents an average of rain over many regions, reducing the seasonal and stochastic effects. This means that it is unlikely that this variable might help to identify the supply functions.

Rain patterns, on the other hand, suggest the presence of a deterministic pattern associated to the rain regimes. However, as this series has been constructed using averages of multiple regions, these patterns are smoothed as the rain regimes differ between regions. Although there are climate cycles that might affect the series, they escape from the strict definition of phenomena that could generate stochastic seasonality as they are not innovations or breaks *per se*, but phenomena that could be modelled. Consequently, rain patterns are affected only by the existence of deterministic seasons.

It is difficult to characterise or identify deterministic seasonality in any of the economic or cost time series. However, they are generally characterised by the existence of distinctive periods: during and after the currency board (Plan de Convertibilidad). Prices are substantially more stable until December 2001/January 2002 when the devaluation occurred (further references can be found in Chapter 2). After the collapse of the currency board, prices observed a distinctive fall as a result of the increase in the nominal exchange rate, followed by a recovery as a result of the typical overshooting in the exchange rate (Dornbusch, 1976), and because of the adjustment of the prices in pesos or devaluation pass-through. However, this recovery has also been supplemented by an increase in the US dollar price, explained by the rise in the demand for these products. The real exchange rate index suggests the counterpart of this story, with a dramatic increase and at the time of the devaluation, and a sustained real appreciation because of the increase in the domestic peso denominated prices and the overshooting in the nominal exchange rate.

Figure 4.5: Additional variables used for identification of Error Correction Models



Augmented Dickey-Fuller unit root tests have been performed on the seven series: Table 4.7 presents the results. The third column presents the ADF statistic and suggests that unit roots that can be rejected are the rain, the real exchange rate and the index price of insecticides. In the rest of the variables, it is not possible to reject the null that innovations can have permanent effects on the level of the series. Additionally, the Portmanteau-Q test for white noise at one and the twelfth lag as well as the Breusch-Godfrey test, indicate that autocorrelation in the residuals of the unit root test regressions can be rejected in all cases.

Table 4.7: Augmented Dickey-Fuller unit root test on identification variables

Name of variable	Description	Model Definition	ADF-t	Q1	Q12	B-G
def_frtlzr_99	USD adjusted Wholesale Fertilizer Price Index	convert	0.97	0.025	1.2528	75.508
def_insect_99	USD adjusted Wholesale Insecticides and Pesticides Price Index	T,C, convert	-4.61	11.005	58.8826	129.26
def_mach_99	USD adjusted Wholesale Agricultural Machinery Price Index	convert	0.92	0.0136	0.4816	83.97
def_diesel_99	USD adjusted Diesel Fuel Price Index	convert	2.6	0.0014	0.3373	22.853
land_p_99	USD price of land per hectare	convert	3.29	0.0075	0.2526	20.178
Rer	Index of real exchange rate	T,C, convert	-5.63	13.71	139.62	124.81
Rain	Rain in millimetres	T,C, SD	-10.38	0.0006	0.569	6.449

Note: Convert identifies the period when the currency board was in place. T stands for trend, C for constant and SD for seasonal dummies.

Source: Own estimations

The above analysis suggests that it would only be possible to find a cointegration relationship between the dependent variable and the identification variable in three of the cases, as they are the only cases where there is a match between the orders of integration of the variables. This implies that, in a standard analysis, only these three variables should be included. However, as we are using these variables just for the identification of the supply relationships, and the analysis of the long-term relationships between these variables is not of interest here, we will include all of them in the cointegration relationships regardless of the matching of the integration order. Additionally, in order to capture the deterministic elements that might explain the export and the domestic supply of the commodities considered, deterministic seasonal dummies have been added in all the specifications.

Table 4.8 presents the results of the ECM for the quantity of exported maize (qemaz). The first two columns present the model selected considering the cointegration relationships using the four prices outlined, according to the result of the test that better rejected the null of no cointegration (with and without constant term). The third and fourth columns present the intermediate cases presented in Table 4.4. The autoregressive terms of qemaz are significant and its importance (measured by the size of its coefficients) is decreasing over time. More specifically, the effect of the immediate lag is lower (but significant) than the following and from this every period becomes less relevant. This result will also hold for prices, implying that in the supply decision, immediate history is substantially more important.

The domestic price of maize (pdmaz) is significant and negative; suggesting that an increase in the price paid domestically should decrease the export supply. Moreover, the lag selection process also suggests that the previous period price is also important in determining the export supply and operates in the same way (higher domestic prices imply lower export supply). The producer price (pqmaz) is significant and positive, suggesting that higher prices imply higher export supply. This might be explained because of additional output released by farmers that increase the export supply. As prices are higher, stocked production might be released increasing both the domestic and the export supply. This explanation, although possible and likely, cannot be validated as we have seen by the theoretical model developed in Chapter 1 that it was impossible to determine their sign analytically. Additionally, the export price of maize seems to play a non-significant role in the determination of the quantity exported, suggesting that in this case, the own direct price exerts no influence on the quantity supplied.

On the other hand, the future price effect on the quantity of maize exported is positive and significant. It remains complicated to conceptualise this result as future price might be reflecting hedging as well as speculation on the spot prices. Moreover, as seen in Chapter 1, there is also an element associated with the cross hedging that is impossible to capture in a single coefficient; however, it seems that the significant positive sign associated to the future price can help to explain the export supply.

Only the real exchange rate (rer) seems to play a significant role in the identification of the supply functions (only in the cases when the four price variables were used). The sign suggests that an increase in the real exchange rate would increase the export of maize: This is an expected result. In the rest of the specifications, although with the right sign, the coefficients are insignificant.

The models in columns three and four (with and without constant term) reflect the intermediate cases presented in the cointegration analysis. In this case, once the effect of the non-significant export prices and future prices are removed, the signs seem to reverse with a positive reaction to the domestic price and a negative for the producer price, in both cases with three lags. The signs are not in line to what the intuition would expect for them. Moreover, the fact that the identification variables are not playing a significant role casts some doubts on this specification.

The error correction term, represented by $\tilde{\omega}_{t-1}^{(1)}$, is not statistically significant in any of the cases analysed. This suggests that cointegration cannot characterise the relationship between these variables. The error correction term cannot guarantee the return to the long-term equilibrium because of a deviation generated by a stochastic event. The fact that the coefficient is not significant suggests that it is possible, through the Granger representation theorem, to reject the cointegration relationship. These results confirm that the result found in the previous section that rejected the possibility of cointegration.

Table 4.8: Error correction representation of the exports of maize

Variable	Model 1	Model 2	Model 3	Model 4
$L\Delta_{12}$.qemaz	-2.955***	-2.916***	-2.996***	-2.954***
	0.063	0.068	0.069	0.068
$L2\Delta_{12}$.qemaz	-3.727***	-3.610***	-3.764***	-3.667***
	0.158	0.172	0.174	0.174
$L3\Delta_{12}$.qemaz	-2.402***	-2.271***	-2.376***	-2.287***
	0.161	0.173	0.175	0.175
$L4\Delta_{12}$.qemaz	-0.666***	-0.618***	-0.638***	-0.611***
	0.067	0.071	0.071	0.072
Δ_{12}.pemaz	-0.247	-0.167		
	0.852	0.875		
Δ_{12} pdmaz	-54.048***	-9.544	-7.606	-4.270
	14.438	40.949	40.373	41.218
$L\Delta_{12}$.pdmaz	-59.022***	65.922	66.711	76.873
	14.623	95.443	93.771	95.698
$L2\Delta_{12}$.pdmaz		146.950	149.056	160.316*
		96.020	94.219	96.137
$L3\Delta_{12}$.pdmaz		72.895*	81.580**	87.246**
		41.014	40.005	40.801
Δ_{12}.pqmaz	57.567***	14.509	10.322	8.154
	15.414	41.522	40.875	41.743
$L\Delta_{12}$.pqmaz	71.968***	-50.149	-53.732	-61.660
	17.564	96.297	94.409	96.383
$L\Delta_{12}$.pqmaz	15.239*	-129.774	-133.954	-143.535
	7.820	96.729	94.623	96.579
$L3\Delta_{12}$.pqmaz	4.733	-68.202	-76.959*	-82.566**
	3.387	41.329	40.175	40.977
Δ_{12}.pfmaz_3	5.632**	5.697**		
	2.718	2.801		
$L\Delta_{12}$.pfmaz_3	11.913*	12.987**		
	6.383	6.530		
$L2\Delta_{12}$.pfmaz_3	10.569*	11.988*		
	6.356	6.475		
$L3\Delta_{12}$.pfmaz_3	5.363*	5.870**		

Table 4.8: Error correction representation of the exports of maize

Variable	Model 1	Model 2	Model 3	Model 4
	2.726	2.785		
$\tilde{\omega}_{t-1}^{(1)}$	-0.273	-0.244	0.105	0.099
	1.302	1.318	1.309	1.337
Constant	20,390.969**		21,473.806**	
	8,385.973		8,361.088	
dm1	-26,931.761**	-1,760.752	-27,773.123**	-2,666.854
	12,933.768	8,386.519	12,824.129	8,478.688
	-	-	-	-
dm2	46,601.951***	28,388.440***	50,214.816***	30,014.382***
	11,371.896	8,313.405	11,360.645	8,373.580
dm3	416.417	23,105.062***	2,240.936	24,880.884***
	11,752.618	8,156.297	11,941.355	8,228.260
dm4	-15,733.852	2,425.670	-18,789.113	1,293.142
	11,590.894	8,505.546	11,431.712	8,518.211
dm5	-20,010.536*	1,133.314	-21,098.188*	1,307.236
	12,000.921	8,436.679	11,938.655	8,325.145
	-	-	-	-
dm6	45,521.488***	26,116.686***	44,179.285***	23,912.538***
	11,319.102	8,416.436	11,297.071	8,257.689
dm7	-905.841	22,905.961***	-7,230.039	16,597.940**
	12,253.002	8,238.067	12,090.252	7,918.672
dm8	-11,101.498	4,471.754	-7,073.532	11,705.007
	11,036.365	8,302.946	10,727.826	8,018.690
dm9	-26,141.985**	-3,387.845	-29,381.791**	-6,322.522
	11,725.848	8,126.583	11,859.910	7,915.209
	-	-	-	-
dm10	46,728.571***	27,546.266***	47,539.545***	28,123.298***
	11,170.618	8,070.379	10,834.121	7,927.039
dm11	-2,215.348	22,776.061***	-4,234.206	22,230.097***
	13,101.280	8,031.353	13,000.401	8,096.897
Δ_{12}.rain	-0.024	-0.073	-0.050	-0.075
	0.069	0.074	0.072	0.073
L6Δ_{12}.rain	0.059	-0.641	-0.604	-1.004*
	0.549	0.578	0.589	0.580
Δ_{12}.rer	2.181*	2.614**	0.302	0.832
	1.218	1.235	1.059	1.061
L6Δ_{12}.rer	-0.307	-0.296	0.098	0.123
	0.273	0.276	0.233	0.237
Δ_{12}.insect	1.223	2.692	-1.258	-0.717
	3.329	3.445	3.069	3.128
Δ_{12}.land	-0.808	-1.012	0.313	0.271
	1.070	1.085	0.995	1.016
Δ_{12}. Diesel	1.216	1.296	0.678	1.000
	0.880	0.891	0.885	0.895
Δ_{12}. Mach	1.915	1.492	0.817	1.396
	1.831	1.909	1.911	1.938
Δ_{12}. Frtlzr	1.723	1.954	1.183	1.362
	1.167	1.195	1.090	1.111
	162	162	162	162
R-squared	0.996	0.996	0.996	0.996
AIC	3784	3789	3787	3793
BJstat	1.504	2.645	0.297	1.099
B_godfrey	141.1	137.8	141.9	140.9
DW	2.880	2.825	2.864	2.868

Note: Standard errors under estimates

Source: Own estimation

Table 4.9 presents the results of the error correction representation for the quantity of exported wheat. In the specification that includes the four price variables, it can be seen that the domestic price of wheat (pdwht) is significant and with the expected sign (any increase in this price reduces the export supply) based on the economic intuition. Additionally, the export price (pewht) is positive and significant, suggesting that an increase in the export price would increase the export supply. The producer price, positive and significant, suggests that, as in the export of maize, an increase in this price increases the supply of both export and domestic products. The same role seems to play the future price that is positive and significant. The positive sign in the producer price might result difficult to conceptualise in a context of processors. However, if these processors are traders, they will act as intermediaries between farmers and the international or domestic demand. This would make the producer price to move in the same direction than the supply.

On the other hand, many of the identification variables proposed seem to locate the supply function. In general, production cost variables (insecticides, machinery and fertilizer) are negative and significant; suggesting that increases in these prices might reduce export supply. In terms of rain, the positive coefficient at the time of supply is hard to conceptualise as it is expected that rain at harvest time should reduce supply (domestic and export) because agricultural machinery cannot operate. Finally, the immediate coefficient of the real exchange rate is negative and significant but the lagged value is positive and significant, suggesting that an increase at the time of planting increases the supply at the time of harvest.

Table 4.9: Error correction representation for the quantity exported of wheat

	Model 1	Model 2	Model 3	Model 4
LΔ_{12}.qewht	-2.851*** 0.053	-2.850*** 0.055	-2.913*** 0.054	-2.912*** 0.055
L2Δ_{12}.qewht	-3.688*** 0.127	-3.680*** 0.132	-3.840*** 0.131	-3.826*** 0.135
L3Δ_{12}.qewht	-2.628*** 0.130	-2.621*** 0.135	-2.756*** 0.137	-2.732*** 0.141
L4Δ_{12}.qewht	-0.856*** 0.057	-0.865*** 0.059	-0.905*** 0.061	-0.897*** 0.062
Δ_{12}.pewht	1.427*** 0.362	1.222*** 0.373		
Δ_{12}.pdwht	-20.761*** 5.572	-22.953*** 5.773	-4.044 10.411	-1.434 10.593
LΔ_{12}.pdwht			22.689** 10.769	23.719** 11.094
Δ_{12}.pqwht	18.088*** 5.586	20.295*** 5.788	1.169 10.460	0.234 10.763
LΔ_{12}.pqwht	-2.345*** 0.692	-2.030*** 0.715	-27.079** 11.082	-24.513** 11.239
LΔ_{12}.pqwht			-2.226* 1.330	
Δ_{12}.pfwht_2	0.969** 0.440	1.205*** 0.453		

Table 4.9: Error correction representation for the quantity exported of wheat

	Model 1	Model 2	Model 3	Model 4
$L\Delta_{12}$.pfwht_2	1.242***	1.348***		
	0.418	0.435		
$\tilde{\omega}_{t-1}^{(1)}$	2.775	2.658	2.717	3.133
	4.954	5.166	5.182	5.339
Constant	20,659.057***		16,414.724***	
	5,857.131		6,040.125	
dm1	-22,245.609**	2,940.955	-17,452.671*	3,163.331
	9,276.796	6,175.067	9,337.030	6,336.904
dm2	-27,377.993***	-10,014.024	-25,063.395***	-7,914.687
	7,841.989	6,365.202	8,418.454	6,534.002
dm3	-8,048.501	14,878.273**	-3,899.717	11,441.800*
	9,183.886	6,765.290	9,301.397	6,797.576
dm4	-55,928.811***	-36,812.759***	-56,442.448***	-35,778.360***
	8,562.725	6,912.713	9,457.440	7,125.168
dm5	31,074.047***	52,946.783***	38,470.955***	54,626.805***
	8,745.486	6,430.358	9,306.553	6,987.048
dm6	-46,515.477***	-26,851.188***	-45,012.583***	-28,665.468***
	8,193.805	6,261.445	8,724.472	6,805.046
dm7	-32,899.135***	-11,706.775*	-29,726.996***	-10,827.864*
	8,333.714	6,021.913	8,688.148	6,271.397
dm8	-9,124.201	11,593.260*	-9,765.067	8,560.301
	8,315.794	6,138.285	8,763.864	6,246.494
dm9	-15,138.897*	6,063.113	-4,638.145	10,930.767*
	8,624.569	6,449.097	8,542.866	6,456.002
dm10	-9,584.802	8,395.802	-12,773.027	4,905.184
	8,162.944	6,648.040	8,681.685	6,704.198
dm11	-54,857.624***	-29,429.070***	-48,145.842***	-27,637.064***
	9,490.408	6,435.835	9,526.370	6,568.767
Δ_{12}.rain	0.119**	0.120**	0.173***	0.145**
	0.056	0.058	0.058	0.058
$L6\Delta_{12}$.rain	0.849**	0.662	0.900**	0.529
	0.394	0.407	0.437	0.429
Δ_{12}.rer	-2.937***	-2.685***	-3.380***	-2.568***
	0.839	0.872	0.989	0.932
$L6\Delta_{12}$.rer	0.465***	0.539***	0.397**	0.511***
	0.171	0.177	0.180	0.177
Δ_{12}. Insect	-5.507**	-4.827**	-6.279**	-4.469*
	2.260	2.348	2.539	2.442
Δ_{12}.land	-0.202	-0.216	-0.243	-0.338
	0.774	0.807	0.742	0.764
Δ_{12}.diesel	2.806***	2.913***	1.255*	1.533**
	0.672	0.700	0.659	0.673
Δ_{12}.mach	-6.265***	-5.596***	-3.972***	-3.547**
	1.556	1.611	1.456	1.493
Δ_{12}.frtlzr	-1.482*	-1.195	-1.899**	-1.624*
	0.852	0.884	0.914	0.932
R-squared	0.996	0.996	0.996	0.995
AIC	3679	3692	3699	3707
BJstat	0.119	0.318	0.604	0.355
B_godfrey	130.1	127.9	132	130.7
DW	2.798	2.801	2.801	2.826

Note: Standard errors under estimates

Source: Own elaboration

The error correction term, on the other hand, suggests that there is no cointegration relationship between the quantity of exported wheat and the variables proposed. Consequently, although the variables might have some power in explaining the short-term movement of the dependent variables, they do not represent a stable long-term relationship. Additionally, an intermediate specification of the export supply is presented in the third and fourth columns without improvements in terms of results.

The results of the error correction representation for the quantity of exported soybeans can be seen in Table 4.10. The export price of soybeans (pesoy) is the variable that presents the best results in terms of significance; however, the sign of the coefficients is against the theoretical model and any economic intuition. Neither the domestic price nor the producer price seems to play a significant role in the explanation of the export supply of soybeans. Only the future price seems to play some role in the determination of the quantity exported. However, its sign results difficult to conceptualise from the economic intuition or from the theoretical model.

On the other hand, the lack of significance of the identification variables suggests that the supply relationship could not be located in the price-quantity space. Only the price of land is significant but its sign is against the economic intuition. On the other hand, the error correction term is not significant suggesting that it was not possible to identify the cointegration relationship.

Table 4.10: Error correction representation of the quantity exported of soybeans

	Model 1	Model 2
$L\Delta_{12}$.qesoy	-2.770***	-2.747***
	0.053	0.053
$L2\Delta_{12}$.qesoy	-3.512***	-3.479***
	0.124	0.123
$L3\Delta_{12}$.qesoy	-2.396***	-2.387***
	0.118	0.118
$L4\Delta_{12}$.qesoy	-0.756***	-0.758***
	0.048	0.048
Δ_{12}.pesoy	-6.736***	-6.000***
	1.779	1.683
$L\Delta_{12}$.pesoy	-16.031***	-13.846***
	3.989	3.672
$L2\Delta_{12}$.pesoy	-13.986***	-11.413***
	4.044	3.599
$L3\Delta_{12}$.pesoy	-5.622***	-4.576***
	1.769	1.571
Δ_{12}.pdsoy	14.633	14.409
	11.593	11.513
$L\Delta_{12}$.pdsoy	28.142**	27.415**
	11.776	11.772
Δ_{12}.pqsoy	-9.362	-11.149
	11.516	11.521
$L\Delta_{12}$.pqsoy	-17.027	-21.893*
	12.110	11.897
$L2\Delta_{12}$.pqsoy	9.647**	3.262***
	4.518	1.220
$L3\Delta_{12}$.pqsoy	2.925	
	1.950	

Table 4.10: Error correction representation of the quantity exported of soybeans

	Model 1	Model 2
Δ_{12}.pfsoy_3	-1.161*	
	0.655	
LΔ_{12}.pfsoy_3	-2.709**	
	1.218	
L2Δ_{12}.pfsoy_3	-1.525**	
	0.717	
$\widetilde{\omega}_{t-1}^{(1)}$	0.142	0.149
	1.215	1.224
Constant	10,373.247	
	6,718.788	
dm1	2,605.340	15,174.268**
	10,550.699	6,653.445
dm2	-23,621.752**	-11,908.363*
	9,791.296	6,718.750
dm3	-16,722.605*	-9,234.794
	8,645.423	6,818.749
dm4	-3,519.701	9,244.785
	10,410.248	6,941.774
dm5	-5,751.378	3,313.170
	9,208.311	6,722.461
dm6	-4,347.521	6,754.213
	9,139.289	6,119.330
dm7	-36,363.226***	-27,070.547***
	8,525.175	5,655.766
dm8	3,333.364	15,636.530**
	9,663.770	6,091.610
dm9	14,789.853*	23,395.270***
	8,361.409	5,971.397
dm10	-42,299.025***	-32,489.687***
	8,940.294	6,239.740
dm11	-10,218.307	4,072.176
	11,013.289	6,688.779
Δ_{12}.rain	0.070	0.052
	0.053	0.052
Δ_{12}.rain_ma6	-0.483	-0.642
	0.463	0.445
Δ_{12}.rer	-0.017	0.101
	0.781	0.766
L6Δ_{12}.rer	-0.002	0.100
	0.186	0.175
Δ_{12}.def_insect_99	-3.879	-3.588
	2.647	2.639
Δ_{12}.land_p_99	1.633*	1.900**
	0.857	0.846
Δ_{12}.def_diesel_99	1.079	1.089
	0.826	0.832
Δ_{12}.Def_mach_99	2.378	2.769*
	1.605	1.605
Δ_{12}.def_frtlzr_99	-0.583	-0.762
	0.908	0.891
R-squared	0.996	0.995
AIC	3688	3689
BJstat	0.0862	0.0430
B_godfrey	147	146
DW	3.003	3.061

Note: Standard errors under estimates

Source: Own elaboration

The error correction representation for the quantity of domestically supplied maize is presented in Table 4.11. In the standard specification and assuming cointegration with the four prices suggested by the theoretical model, the domestic price (pdmaz) seems to be significant and presents the right sign. The other variable that results significant in this specification is the future price one lag ahead, presenting a negative sign that is still hard to conceptualise. However, as we have seen from the discussion in chapter 1, depending on the size of the speculation component as well as the cross hedging strategy followed, this coefficient might present different signs.

In terms of the identification of the supply equation, only the rain variable seems to play a significant role and presents a correct sign, suggesting that current and lagged increases in rain can reduce the domestic supply of maize. The rest of the variables are not significant. On the other hand, the error correction term represented by $\tilde{\omega}_{t-1}^{(1)}$, is not significant, which suggests that there is no cointegration relationship between the variables presented. This confirms the results found in the previous section with respect to the lack of cointegration relationship between the variables.

The additional intermediate specifications of the error correction model given in columns three to six did not reveal improvement in the identification of any long-term relationship between the variables postulated. Moreover, the lack of significance of the error correction term indicates that there is no cointegration relationship between these variables.

Table 4.11: Error correction representation of the quantity domestically supplied of maize

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$L\Delta_{12}$.qdmaz	-3.133***	-3.133***	-3.101***	-3.084***	-3.097***	-3.071***
	0.059	0.059	0.057	0.057	0.055	0.056
$L2\Delta_{12}$.qdmaz	-4.373***	-4.374***	-4.263***	-4.228***	-4.256***	-4.197***
	0.148	0.148	0.142	0.142	0.140	0.141
$L3\Delta_{12}$.qdmaz	-3.139***	-3.141***	-3.001***	-2.975***	-2.999***	-2.941***
	0.154	0.153	0.143	0.144	0.142	0.144
$L4\Delta_{12}$.qdmaz	-0.996***	-0.997***	-0.922***	-0.920***	-0.926***	-0.907***
	0.066	0.064	0.058	0.059	0.058	0.059
Δ_{12}..pemaz	0.043	0.046				
	0.190	0.185				
$L\Delta_{12}$.pemaz	-0.249	-0.242				
	0.458	0.449				
$L2\Delta_{12}$.pemaz	-0.670	-0.666				
	0.462	0.457				
$L3\Delta_{12}$.pemaz	-0.385*	-0.383*				
	0.198	0.196				
Δ_{12}..pdmaz	2.305*	2.315*	0.545	0.520	-0.082	-0.152
	1.198	1.186	0.407	0.406	0.196	0.199
$L\Delta_{12}$.pdmaz	4.726*	4.746*	0.800*	0.869**		

Table 4.11: Error correction representation of the quantity domestically supplied of maize

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	2.713	2.689	0.422	0.419		
L2Δ_{12}.pdmaz	4.618*	4.641*				
	2.752	2.723				
L3Δ_{12}.pdmaz	2.411*	2.425**				
	1.223	1.203				
Δ_{12}..pqmaz	-2.205*	-2.216*	-0.496	-0.481	0.094	0.164
	1.225	1.211	0.417	0.414	0.199	0.203
LΔ_{12}.pqmaz	-4.282	-4.304	-0.731*	-0.788*		
	2.747	2.721	0.426	0.421		
LΔ_{12}.pqmaz	-3.984	-4.009				
	2.771	2.739				
L3Δ_{12}.pqmaz	-2.162*	-2.177*				
	1.230	1.207				
Δ_{12}..pfmaz_1	-0.166***	-0.167***				
	0.050	0.048				
LΔ_{12}..pfmaz_1	-0.171***	-0.172***				
	0.051	0.050				
Δ_{12}..pfmaz_2			-0.016	0.188*		
			0.033	0.110		
LΔ_{12}..pfmaz_2			-0.057*	0.415		
			0.032	0.255		
L2Δ_{12}..pfmaz_2				0.441*		
				0.251		
L3Δ_{12}..pfmaz_2				0.148		
				0.103		
Δ_{12}..pfmaz_3					0.076**	0.220***
					0.030	0.081
LΔ_{12}..pfmaz_3					0.055*	0.420**
					0.030	0.188
L2Δ_{12}..pfmaz_3						0.376**
						0.186
L3Δ_{12}..pfmaz_3						0.158**
						0.080
$\widetilde{\omega}_{t-1}^{(1)}$	-0.054	-0.052	1.294	1.155	0.827	0.880
	2.972	2.959	2.927	2.911	2.921	2.910
Constant	21.508		304.599		403.221	
	290.103		268.019		261.178	
dm1	14.482	44.636	-572.181	-107.976	-682.635	-185.234
	487.364	267.409	424.527	263.933	413.524	260.138
dm2	-1,066.332***	-1,049.990***	-1,216.930***	-1,019.448***	-1,336.816***	-896.924***
	350.379	271.228	350.338	278.459	346.878	274.114
dm3	1,259.061***	1,282.372***	1,086.124***	1,468.066***	931.730**	1,281.861***
	404.086	252.755	394.310	264.972	379.953	258.042
dm4	-553.298	-532.365**	-1,073.328***	-818.826***	-963.082***	-555.002**
	385.465	261.316	353.714	263.976	356.082	255.136
dm5	290.450	313.378	132.131	466.961*	-275.262	156.543
	409.495	267.301	398.986	268.776	372.158	254.524
dm6	-1,095.837***	-1,077.014***	-1,306.118***	-1,062.492***	-1,150.048***	-808.172***
	370.634	268.894	359.735	265.308	356.533	252.673
dm7	1,409.523***	1,433.963***	890.870**	1,305.443***	784.620**	1,280.291***
	421.123	260.962	366.881	248.550	361.403	244.877
dm8	-893.405**	-874.417***	-921.840**	-693.919***	-1,192.227***	-845.795***
	365.190	259.242	354.402	247.592	345.259	245.878
dm9	627.934	651.980**	154.029	438.079*	223.773	608.003**
	411.794	252.662	377.922	246.154	383.602	248.582
dm10	-1,101.368***	-1,085.072***	-1,186.775***	-857.812***	-1,316.095***	-959.219***
	335.682	252.649	329.651	257.424	326.870	260.256

Table 4.11: Error correction representation of the quantity domestically supplied of maize

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
dm11	878.647*	908.220***	387.591	779.044***	186.316	769.934***
	473.405	253.897	433.418	255.039	415.121	256.143
Δ_{12}..rain	-0.006**	-0.006**	-0.002	-0.001	-0.003	-0.002
	0.002	0.002	0.002	0.002	0.002	0.002
L6Δ_{12}..rain	-0.084***	-0.084***	-0.050***	-0.055***	-0.039**	-0.039**
	0.021	0.019	0.017	0.017	0.016	0.016
Δ_{12}.rer	0.011	0.011	-0.049	-0.072**	0.020	0.029
	0.034	0.033	0.034	0.036	0.032	0.033
L6Δ_{12}..rer	0.002	0.002	0.001	0.003	-0.006	-0.008
	0.008	0.008	0.007	0.007	0.007	0.008
Δ_{12}.insect	-0.055	-0.055	-0.182*	-0.246**	-0.086	-0.062
	0.100	0.099	0.094	0.101	0.091	0.096
Δ_{12}..land	0.005	0.005	0.018	0.034	-0.018	-0.020
	0.036	0.035	0.034	0.034	0.030	0.031
Δ_{12}..diesel	-0.019	-0.019	0.007	0.012	0.016	0.020
	0.029	0.029	0.027	0.027	0.026	0.026
Δ_{12}..mach	0.106*	0.106*	0.067	0.076	0.072	0.070
	0.060	0.059	0.056	0.056	0.055	0.055
Δ_{12}..frtlzr	0.044	0.044	0.006	-0.004	0.069**	0.071**
	0.036	0.036	0.033	0.033	0.034	0.034
R-squared	0.997	0.997	0.997	0.997	0.997	0.997
AIC	2599	2597	2603	2602	2602	2602
BJstat	0.0525	0.0438	3.555	2.151	1.989	0.478
B_godfrey	148.7	148	145.6	145.8	147.5	148.5
DW	2.892	2.893	2.832	2.855	2.862	2.894

Note: Standard errors under estimates

Source: Own estimation

Table 4.12 presents the error correction representation of the domestic supply of wheat. The domestic price of wheat (pdwht) is positive and significant, which is in line with the theory developed in Chapter 1 and with the economic intuition. On the other hand, the price of futures with three lags is also significant and positive, suggesting that increases in this price, increases the domestic supply via the increase in the supply of the input. However, as we have discussed, it is difficult to conceptualise the coefficient of this price given the different elements that can affect the behaviour of agents with respect to the futures price. The other price variables do not play a significant role in the determination of the domestic supply.

In terms of the identification of the supply function, the variables associated with the rain are significant and present the expected signs. The cost variables, on the other hand, present mixed results with the price of machinery being negative (as expected) and the price of insecticides positive.

The error correction term, on the other hand, cannot help to correct the deviations in the long run relationship because of not being significant. This suggests that this specification cannot accurately represent the long-term relationship between the quantity of wheat supplied domestically and the prices specified in the theoretical model. The intermediate specification presented in columns three and four seem to reinforce the role of the domestic price in the determination of the domestic supply. However, the relationship cannot be characterised as stable in the long run because of the lack of significance in the error correction term.

Table 4.12: Error correction representation of the domestic supply of wheat

	Model 1	Model 2	Model 3	Model 4
$L\Delta_{12}$.qdwht	-2.957***	-2.958***	-2.910***	-2.918***
	0.060	0.062	0.060	0.062
$L2\Delta_{12}$.qdwht	-4.022***	-3.982***	-3.939***	-3.918***
	0.137	0.140	0.136	0.140
$L3\Delta_{12}$.qdwht	-2.863***	-2.809***	-2.788***	-2.759***
	0.136	0.138	0.134	0.138
$L4\Delta_{12}$.qdwht	-0.928***	-0.881***	-0.894***	-0.858***
	0.059	0.059	0.058	0.058
Δ_{12}.pewht	0.171	-0.021	0.198	0.211
	0.231	0.151	0.241	0.249
$L\Delta_{12}$.pewht	0.264	-0.351	0.353	0.319
	0.526	0.263	0.550	0.567
$L2\Delta_{12}$.pewht	0.419	-0.325**	0.617	0.554
	0.518	0.147	0.542	0.559
$L3\Delta_{12}$.pewht	0.351		0.502**	0.451*
	0.219		0.230	0.236
Δ_{12}.pdwht	2.566*	6.861**	2.870*	3.317**
	1.478	2.662	1.479	1.557
$L\Delta_{12}$.pdwht	4.480*	15.179**	3.789	5.278**
	2.494	5.967	2.497	2.664
$L2\Delta_{12}$.pdwht	3.253**	13.809**	2.805*	4.226***
	1.427	5.998	1.445	1.608
$L3\Delta_{12}$.pdwht		4.362*		0.555***
		2.561		0.199
Δ_{12}.pqwht	-2.202	-6.670**	-2.440	-3.012*
	1.511	2.690	1.484	1.548
$L\Delta_{12}$.pqwht	-3.462	-14.548**	-2.353	-4.123
	2.578	6.022	2.513	2.612
$L2\Delta_{12}$.pqwht	-2.223	-13.072**	-1.155	-2.809*
	1.522	6.028	1.484	1.511
$L3\Delta_{12}$.pqwht	0.354*	-4.089	0.639***	
	0.193	2.570	0.195	
Δ_{12}.pfwht_3	0.034	0.111		
	0.117	0.118		
$L\Delta_{12}$.pfwht_3	0.362	0.596**		
	0.289	0.290		
$L2\Delta_{12}$.pfwht_3	0.616**	0.856***		
	0.284	0.285		
$L3\Delta_{12}$.pfwht_3	0.307***	0.399***		
	0.111	0.113		
$\tilde{\omega}^{(1)}_{t-1}$	8.547	8.480	4.569	4.068
	6.799	6.912	6.794	7.011
Constant	-2,360.885***		-2,521.213***	
	800.331		834.608	
dm1	5,197.345***	1,298.257**	5,135.938***	1,000.540
	1,478.257	629.992	1,522.082	672.704
dm2	-1,869.228**	-2,902.929***	-1,399.996*	-2,551.616***
	725.251	662.388	775.234	700.937
dm3	6,876.252***	3,506.492***	6,761.885***	3,217.541***

Table 4.12: Error correction representation of the domestic supply of wheat

	Model 1	Model 2	Model 3	Model 4
	1,326.145	657.712	1,351.680	697.102
dm4	-1,270.422	-2,908.647***	-817.531	-2,579.848***
	844.233	662.519	904.349	708.576
dm5	5,847.043***	2,751.005***	5,687.595***	2,445.108***
	1,211.968	656.795	1,249.238	698.621
dm6	-1,854.963**	-3,369.493***	-1,455.073*	-3,194.535***
	821.501	614.516	866.910	654.465
dm7	6,684.444***	3,710.606***	6,856.054***	3,636.441***
	1,131.459	541.745	1,195.899	582.217
dm8	-1,026.931	-3,167.201***	-1,298.596	-3,260.353***
	847.053	574.394	852.773	603.484
dm9	5,375.611***	2,699.804***	6,391.860***	3,225.893***
	1,183.528	653.414	1,247.078	663.588
dm10	-610.573	-1,956.115***	-1,370.896*	-2,763.751***
	778.675	645.432	790.266	665.230
dm11	5,327.103***	1,378.377**	6,129.985***	2,076.091***
	1,408.373	605.841	1,467.861	650.060
Δ_{12}.rain	-0.015***	-0.014**	-0.022***	-0.020***
	0.005	0.006	0.005	0.006
L6Δ_{12}.rain	-0.098***	-0.086**	-0.118***	-0.092**
	0.036	0.039	0.038	0.041
Δ_{12}.rer	0.150*	0.091	0.158*	0.108
	0.086	0.085	0.086	0.089
L6Δ_{12}.rer	-0.035*	-0.024	-0.046**	-0.028
	0.018	0.018	0.019	0.018
Δ_{12}.insect	0.646**	0.427*	0.652***	0.459*
	0.249	0.243	0.242	0.251
Δ_{12}.land	-0.108	-0.094	-0.143*	-0.124
	0.075	0.075	0.075	0.078
Δ_{12}.diesel	0.086	0.111	0.072	0.098
	0.071	0.071	0.067	0.069
Δ_{12}.mach	-0.337**	-0.325**	-0.381**	-0.404**
	0.137	0.138	0.147	0.155
Δ_{12}.frtlzr	-0.106	-0.078	-0.027	0.033
	0.092	0.093	0.081	0.084
R-squared	0.998	0.998	0.998	0.998
AIC	2833	2841	2855	2864
BJstat	7.650	7.712	8.875	14.65
B_godfrey	119.6	120.3	132.6	131.8
DW	2.871	2.758	3.005	2.885

Note: Standard errors under estimates

Source: Own elaboration

The error correction representation for the domestic supply of soybeans is presented in Table 4.13. In this case, the export price is significant and negative, suggesting that an increase in this price would generate a decrease in the domestic supply of soybeans, an expected result from the economic intuition point of view. At the same time, the producer and the price for the futures with two lags are significant and positive, indicating that an increase of these prices should increase the quantities available for the domestic market (and the export market as well). However, from the discussion in chapter 1, it is very little what can be added from the analytical point of view. The domestic price is not significant.

In terms of representation and identification of the supply function, only the price of the fertilizer is significant and presents an expected sign. The rest of the cost variables, as well as the climate ones, are not significant. The same applies to the real exchange rate.

Finally, the fact that the error correction term is not significant, casts some doubts on the capability of this model to represent the long run relationship between the quantity of soybeans supplied and the prices outlined in the theoretical model. The other specifications presented in columns three and four suggest a more important role for the producer price of soybeans. Nevertheless, being the error correction term insignificant is very little what can be concluded.

Table 4.13: Error correction representation for the domestic supply of soybeans

	Model 1	Model 2	Model 3	Model 4
LΔ_{12}.qdsoy	-2.923***	-2.928***	-3.042***	-3.045***
	0.060	0.060	0.060	0.060
L2Δ_{12}.qdsoy	-3.840***	-3.872***	-4.197***	-4.215***
	0.147	0.146	0.153	0.152
L3Δ_{12}.qdsoy	-2.614***	-2.670***	-3.028***	-3.060***
	0.158	0.156	0.169	0.164
L4Δ_{12}.qdsoy	-0.780***	-0.812***	-0.975***	-0.994***
	0.072	0.070	0.077	0.074
Δ_{12}.pesoy	-5.155***	-5.240***		
	1.530	1.539		
LΔ_{12}.pesoy	-15.041***	-15.186***		
	3.403	3.423		
L2Δ_{12}.pesoy	-13.661***	-13.786***		
	3.426	3.447		
L3Δ_{12}.pesoy	-4.348***	-4.447***		
	1.498	1.506		
Δ_{12}.pdsoy	14.105	13.288	-7.340	-7.503
	10.041	10.091	5.592	5.584
LΔ_{12}.pdsoy	18.274*	17.924*		
	9.950	10.011		
Δ_{12}.pqsoy	-10.963	-9.985	10.134*	10.402*
	10.139	10.184	5.810	5.798
LΔ_{12}.pqsoy	-10.514	-9.776	5.355***	5.625***
	10.239	10.293	1.944	1.918
LΔ_{12}.pqsoy	3.940***	4.173***	1.930	2.111*
	1.127	1.124	1.194	1.176
Δ_{12}.pfsoy_2	0.406***	0.384***		
	0.125	0.125		
$\tilde{\omega}_{t-1}^{(1)}$	0.218	0.259	0.404	0.428
	1.196	1.204	1.268	1.267
Constant	9,316.151		6,067.569	
	5,770.603		6,866.210	
dm1	6,641.813	17,906.089***	16,318.721	23,517.302***
	9,006.419	5,730.879	10,544.921	6,690.494
	-	-	-	-
dm2	26,361.458***	17,355.702***	25,045.537***	19,101.307***
	7,792.533	5,475.415	9,337.056	6,470.232
dm3	-9,022.361	-228.852	-6,179.575	-468.121
	7,957.675	5,837.971	9,379.785	6,791.982
dm4	-13,698.555	-4,133.299	-11,139.204	-4,926.875
	8,467.831	6,087.906	9,898.020	6,962.011
dm5	8,965.181	18,430.349***	12,679.986	18,897.505***
	8,405.092	6,060.584	9,835.609	6,867.136

Table 4.13: Error correction representation for the domestic supply of soybeans

	Model 1	Model 2	Model 3	Model 4
dm6	-16,565.095**	-7,551.113	-12,209.751	-6,413.607
	8,025.674	5,801.474	9,360.738	6,672.982
	-	-		
dm7	25,721.453***	15,873.170***	-19,509.937**	-13,118.577**
	8,073.743	5,322.264	9,591.596	6,294.646
dm8	-3,971.712	4,930.371	-5,852.370	31.512
	7,678.513	5,377.285	9,178.981	6,313.033
dm9	17,000.985**	26,223.917***	22,098.198**	28,016.915***
	7,819.519	5,372.957	9,209.924	6,316.391
	-	-	-	-
dm10	31,789.028***	22,500.779***	25,021.909***	18,936.990***
	8,022.120	5,625.746	9,550.096	6,611.911
dm11	-19,747.206**	-8,935.141	-19,230.287*	-12,226.370*
	8,858.248	5,834.488	10,476.626	6,845.729
Δ_{12}.rain	-0.041	-0.048	-0.004	-0.010
	0.049	0.049	0.057	0.056
$L6\Delta_{12}$.rain	-0.300	-0.437	-0.141	-0.249
	0.431	0.426	0.484	0.468
Δ_{12}.rer	-0.586	-0.456	-1.048	-0.956
	0.723	0.723	0.858	0.851
L6Δ_{12}.rer	-0.102	-0.101	-0.112	-0.109
	0.174	0.175	0.180	0.180
Δ_{12}.insect	3.763*	4.113*	0.892	1.179
	2.142	2.145	2.447	2.424
Δ_{12}.land	1.203	1.124	-0.948	-0.964
	0.774	0.777	0.833	0.832
Δ_{12}.diesel	-0.249	-0.159	0.324	0.356
	0.707	0.709	0.702	0.701
Δ_{12}.mach	-2.381	-2.283	-2.064	-2.081
	1.492	1.500	1.581	1.579
Δ_{12}.frtlzr	-2.165**	-2.109**	-1.041	-0.983
	0.937	0.942	0.978	0.975
R-squared	0.995	0.995	0.992	0.992
AIC	3667	3668	3723	3722
BJstat	0.350	0.471	17.27	16.64
B_godfrey	131.7	130.9	147.8	145.3
DW	2.888	2.908	2.967	2.978

Note: Standard errors under estimates

Source: Own elaboration

The inspection of the results of the ECMs reveals that none of the cases analysed allow us to make conclusions about the long-run relationships between the variables. In none of the cases was it possible to find a significant error correction terms. This suggests that the variables are not effectively cointegrated and confirms some of the results found in the previous section. Consequently, the results seem to be in line with the conclusions extracted when the McKinnon critical values are used to perform the unit root tests on the residuals for the cointegration test.

Although some of the price variables seem to present some association or relationship with the quantity variables, it is difficult to characterise it as long term and stable because of the poor cointegration results. Moreover, although some of the signs have been identified, and they are in line with the theoretical model and economic intuition, there is no consistency across the models and products analysed. Whilst in some cases, for example, it was possible to identify some form of cross price effect between export prices and domestic supply (and *vice versa*); this relationship has failed to appear in others. At the same time, contrary to economic intuition and the theoretical model, it was not possible in all cases to verify a direct price effect on the quantities supplied.

Moreover, as discussed in chapter 1, from the theoretical point of view, it is not possible solely with the information available to conceptualise many of the results found. Particularly in the case of the coefficient of the futures price, depending on the size of the speculation component, the sign of the coefficient in the supply function could be either positive or negative. Moreover, the fact that the futures markets can be used to hedge against fluctuations in the producer price but also as a substitute of the cross-hedging strategy between domestic and export price, introduces another source of uncertainty. The information available cannot help to identify any of these possibilities. This means that additional analysis and estimation is necessary to characterise the type of behaviour of agents with respect to the futures market; and eventually, help to characterise the sign of this coefficient.

Additional attempts should be considered in order to identify adequately the supply functions. Although costs, profits and climate are clearly the right type of variables to identify the functions, the variables chosen have generally failed to identify properly the functions. In many cases, these variables have resulted not significant and, in others, the signs associated with them are hard to associate to economic intuition. On the other hand, although some of the variables chosen are of relevance to agricultural production, they might not necessarily be the adequate variables to identify supply functions of economic process where processors or traders are involved. Although the real exchange rate can affect both processors and suppliers alike, costs of production and climate variables may have different effects on each type of agent. The finding of a proper identification variable or set of variables that can characterise this type of complex process is of paramount importance and needs to continue being pursued.

4.8. ALTERNATIVE APPROACHES

When looking for explanations for these results, several causes may be behind them. One possibility is the data and the length on the period under use. With respect to the latter, as was discussed, it may be counterproductive to go further into the past. The alternative may be to use newer data. At the time this exercise was performed, the data used was the latest available. It is possible now to add three or four years of latest data to increase the length of the sample. However, whilst additional data may improve the power in the inference, it is unlikely that it would substantially change the results obtained.

There is, however, also the problem that some of the variables used have been constructed and some measurement errors might have affected their behaviour. The data on quantities are unlikely to suffer from these problems, but some of the prices might have been affected. If the definition of these prices cannot be improved, the alternative might be trying in a different country where better data can be found. However, in any case, the choice of using monthly data could be implicated and it is recommended to continue with it. In fact, the estimation of this model using data for a different country should necessarily be carried out for validation purposes, even if the results presented here were more conclusive.

A second group of explanations may lie on the election of the cointegration method chosen. The E-G methodology was selected based on its simplicity and clarity of exposition and by the ease in interpreting their results. The application of this methodology seems to develop more naturally making it more amenable to the application. Nevertheless, the fact that it involves different steps in its application makes it prone to the commission of mistakes. These mistakes can be transmitted to the following stages without warning, and this can completely invalidate the process. On the other hand, it should be considered that the large-sample properties in their estimation and inference might present problems when shorter samples are used. This suggests that even the cointegration conclusions we have found may not be valid.

This suggests that the application of a Johansen-type of cointegration approach may be advisable. Although this approach is harder to implement, less clear and present some interpretational problems, as we have seen, it is a very reliable method that has been extensively applied and, particularly for monthly data, has received more attention than the EG approach. However, it is unlikely that a different cointegration method may substantially alter the conclusions, although we have seen that literature has identified some cases. It may improve some of the results, but it would be extremely suspicious if, because of a change in the method, diametrically opposed conclusions were reached. However, this methodology also requires larger sample sizes, which means that its application will depend on the improvement in the quantity of data available.

Alternatively, there is the possibility that we have not selected the most suitable model between all the options in terms of the lags considered for each variable in the ECM, simply because they were not in the sample. The solution to this may be to extend the number of lags considered and select from a bigger sample. However, there is the possibility that a very long lag structure may be selected, with implications in terms of the power of the test given by the additional number of variables to be estimated. Whilst by adding additional lags we may improve the model, it is extremely unlikely that this will be the case.

A third attempt at explanation goes deeper into the cointegration concept and casts some doubt about its relevance. It may be the case that cointegration is not a suitable methodology for the estimation of this model. Cointegration requirements, that the variables involved should be integrated of the same order, may be seen to be a bit rigid and an alternative estimation approach may be recommended.

One possibility is the estimation of a more general dynamic model through a transfer function. Given that the variables involved contain a unit root, an ADL model may provide a good approach for the estimation of the model. Their requirements are not as stringent as the cointegration approach and, if they are met, consistent estimators of the true parameters may be obtained.

However, in this case the issue of the exogeneity of the independent variables should be addressed. This implies that before the estimation, an analysis of the cross-correlograms should be performed in order to detect the type of model to estimate (in terms of their autoregressive and moving average components), as well as the issue of the exogeneity of the independent variables. In case the exogeneity cannot be rejected, an instrumental variables approach should be followed to address this issue.

Nevertheless, before completely discarding the possibility of cointegration, some alternative and less restrictive cointegration approaches may be attempted. Fractional cointegration, for example, is seen generally as less restrictive than standard cointegration. Fractional integrated processes were identified originally by Granger (1986) and have received important applications by Cheung and Lai (1993) and Baillie and Bollerslev (1994). In fractional cointegration, residuals of the cointegration equation are integrated of a real number between 0 and 1. This is because fractional integration processes also present mean-reversion in the errors, even when they are not processes integrated of order zero.

Seasonality, on the other hand, can be treated in an analogous way. Gil-Alana and Robinson (2001) have developed tests for seasonal fractional integration. This is very relevant since, given that they are testing a less stringent hypothesis, the application of these tests may suggest the presence of additional unit roots to those already found in this chapter, which might change the posterior analysis.

The final explanation may be that, effectively, the theory developed may not be an accurate representation of the reality. It may happen that elements not considered in the original theoretical development may be very relevant and their exclusion can make the theory invalid. The results obtained here may be reflecting that fact. Nevertheless, it may be too soon to discard completely a theoretical model based only on just one estimation exercise, particularly when doubts exist about the data, approach and methodology used to perform the empirical validation.

4.9. CONCLUSIONS

This chapter had two defined purposes. It tried, first, to address the issue of seasonal cointegration on monthly data using the Engle-Granger approach; and second, it tried to determine if the variables defined in the theoretical model could be characterised through a long-term relationship. In particular, monthly data on three agricultural commodities produced and commercialised in Argentina were used to estimate the model. In order to perform this analysis, different definitions of prices for commodities in Argentina in recent history were tested for the presence of seasonal unit roots with and without structural breaks.

Besides the zero frequency, no matching seasonal unit roots could be found between the prices and the quantities analysed in the second chapter, which eliminates the possibility of seasonal cointegration. Consequently, only the standard cointegration analysis could be performed on the proposed model.

The results we have encountered so far reveal that cointegration cannot be sustained by the data. The parameters of the error correction model reveal that, in any of the cases analysed, a cointegration relationship exists between the variables. However, if standard DF critical values are used in the cointegration tests, in contrast to those proposed by MacKinnon, cointegration cannot be rejected. This contradiction, as well as the violation of some of the assumptions of the cointegration analysis, particularly normality, impedes making strong conclusions about the existence of cointegration relationships.

Error correction models were attempted as well. In addition to the variables under consideration, cost and climate variables were included for identification purposes, as well as seasonal deterministic dummy variables. Although the error correction models might present some estimation problems, the conclusion of no cointegration is still sustained. This means that it was not possible to prove the existence of a long term and stable relationship between the variables outlined in the theoretical model.

The length of the data used may present problems. Extending the sample further into history may be problematic since grains markets in Argentina were regulated before the sample, implying that historical series may respond to different dynamics. In recent years, on the other hand, newer data, has been available that can increase the length of the series to perform the estimation. However, it is also true that the length of data used in this analysis were not particularly shorter than similar cointegration exercises. Nevertheless, monthly data should continue to be used.

The data used is also suspect. Some of the series used are not observed values and have been constructed instead. Improved definitions of the prices that are not available should be considered and used. Alternatively, the estimation of the model using another set of data for other products and/or country where good quality series are available, will help and could serve as a comparison for future research.

The estimation methodology is also under scrutiny. The Engle-Granger approach for cointegration was preferred given its simplicity, clarity and ease of interpretation of its estimates. However, its error prone procedure that can carry mistakes along the whole process without warning, and some estimations problems with respect to its large-sample properties are not applicable to the small sample we have used, could also be blamed. Consequently, a Johansen cointegration approach, or less restrictive frameworks such as fractional cointegration, are recommended.

However, the possibility that the cointegration analysis may not be the most appropriate approach to estimation should be considered. Alternative, more general, dynamic models, such as Autoregressive Distributed Lag models, should be considered. This would require analysing and addressing the issue of exogeneity of the independent variables that the cointegration approach does not require.

OVERALL CONCLUSIONS

This research has addressed aspects of the determination of the export and domestic supply of storable agricultural commodities when futures markets are available and output is subject to technological risk. Much of the literature has focused on these issues separately, and this is an attempt to reconcile all those elements in one single framework. The analytical framework is validated empirically using the case of three agricultural commodities in Argentina. Moreover, it also presents a technical discussion about the issue of seasonal unit roots in agricultural commodities by presenting general insights and contributions for research beyond the topic of discussion.

Chapter 1 developed a model whereby a trader transforms a single storable input produced under technological risk into both an exported and a domestically supplied product. Futures markets are available to hedge against fluctuations in the spot price of the input. This treatment unites the contributions on the second stages of production made by Hirshleifer (1988) and cross-hedging by Anderson and Danthine (1980, 1981) in the determination of the supply of exports and domestic agricultural commodities. However, the model is flexible enough to consider general situations where processors transform inputs into several outputs.

The well-known separation result between output and the hedging decision does not hold in this framework, and both elements must be determined jointly in the trader's problem. In this model, traders tend to buy futures, providing a natural counterpart to producers and other agents, who tend to sell them. This has been verified even under less restrictive conditions than the previous findings. However, this still cannot produce an unbiased futures price, and only under very special conditions will the bias disappear. Even when traders are assumed to be neutral to risk, in contrast to the other agents under the same assumption, a bias is still necessary to attract some agents to cover the short position taken by the rest of the agents. On the other hand, it is verified that, under special circumstances, the model collapses into those already analysed by the literature, suggesting coherence in the treatment between this approach and previous ones.

However, the parametric approach followed has generated very complex mathematical formulations that complicate its analytical study. The inclusion of a second stage of production with two markets to supply implies the interaction of several parameters and variables, particularly the expected variances and covariances between the prices that intervene in the model. The parameters and coefficients of the supply equations are mainly affected by the expectations that agents form about the volatility of prices. However, the supply equations found are linear and, under the assumption of the stability of coefficients, they can be estimated econometrically.

In contrast to the case where no second stages in production are considered, the long-run equilibrium price of both processed products depends not only on the structural parameters of the model but also on the price of the input. It was also found that the bias in the future price remains in the long run, and only when processors are risk neutral does this bias disappear, in contrast to the solution in the short run.

The decision in terms of which market to supply is explained not only by differences in prices but also by differences in the expected variances and covariances of the prices, since the trader is also interested in reducing the volatility of his/her profits. This suggests that, *ceteris paribus*, the supply of a particular market may increase if the volatilily of the price in the other market is high. This verifies the cross-hedging strategy followed by taking complementary positions in the spot and futures markets for all products available.

Additionally, the futures market can offset the volatility in the price of the input, and can reduce part of the effect that the volatility in the price of both processed products has on the supply, through the way in which input and processed product prices move together.

Empirical validation of the export and domestic supply equations developed in Chapter 1 has been attempted using cointegration analysis. Since the data to be used present important seasonal patterns, a seasonal cointegration approach was carried out. Nevertheless, before the estimation it was necessary to test the series under study for the presence of seasonal unit roots using the HEGY test.

A time series based on agricultural processes presents very acute forms of seasonality. Seasons where no exports or domestic supply are recorded are frequent in monthy data, are part of the domain of the series and are the result neither of missing values nor of lack of registration. The HEGY test has never been applied before in data with these characteristics; this raises questions about the validity of the procedure and addresses the critical values used to perform the inference. On the other hand, temporal aggregation into data of higher frequencies in order to remove this issue from the data may present estimation and inference problems, as well as hide relevant information about the processes.

A Monte Carlo simulation exercise was performed in order to verify that the critical values, for monthly data, are not affected by the presence of zero values. It was found that the critical values tabulated using data affected by zero values are no different to those already found by the literature. Given that additional model specifications of the testing equation have been considered, the critical values obtained here supplement those already found for monthly data.

When the HEGY test indicates the presence of a unit root, it is also necessary to verify that the test has not been affected by the presence of an unknown structural break. No applications have been found for the HEGY test under the presence of unknown structural breaks on any kind of monthly data. Therefore, critical values to perform the HEGY test in this context with monthly data were not available. Monte Carlo simulations to obtain critical values when series present zero values have been carried out. It was found that zero values do not seem to affect the critical values. Therefore, those obtained here could eventually be used on any application of monthly data with any type of data where there is suspicion of structural breaks. Nevertheless, as occurs when quarterly data are considered, the critical values of the HEGY test in monthly data when structural breaks are considered tend to be different to those when no breaks are present in the data.

In the context of the HEGY test under unknown structural breaks, the misplacement in the selection of the break date found in quarterly data by Harvey, Leybourne and Newbold (2002), when the break is selected by maximising the significance of the break dummy, could not be verified using monthly data in this exercise. Differences in the way in which the DGP in the Monte Carlo simulations in both exercises could be responsible, as well as differences in the size of the breaks, were considered.

Given that the HEGY tests under the presence of unknown structural breaks require the estimation of a testing equation with several parameters, a problem with the power of the test may appear if the data sample is not particularly long. As a consequence, the modelling of the testing equation of the HEGY test under an unknown structural break, using just one dichotomic variable to capture the break, was attempted and critical values were obtained using Monte Carlo simulations. It was found again in this case that the distribution of the test statistic is not affected by the presence of zero values in the series. Additionally, the misplacement in the selection rule of the break using non-seasonal data in Harvey, Leybourne and Newbold (2001) has been verified in our seasonal context. This suggests that this misplacement is the result of the way in which the structural break is modelled and not of the seasonality in data or the presence of zero values.

An empirical application was performed using monthly exports and the domestic supply of soybeans, maize and wheat in Argentina between 1994 and 2008, revealing that unit roots at the zero frequency could not be rejected in any of the cases; this indicated the presence of a stochastic trend in the series. On the other hand, although it was not possible to reject seasonal unit roots in some cycles in some of the series, the characterisation of the series, as presenting stochastic seasonality, cannot be sustained. This suggests that the case for stochastic seasonality in the supply of these commodities is not very strong. Additionally, a deterministic approach to seasonality has shown that it can explain a relatively high share of the monthly variation in the series analysed. This indicates that deterministic seasonality would be an adequate or sufficient approach for the seasonality observed in the series. These results have been backed by the graphical inspection of the series in question and their comparison with the theoretical ACFs and PACFs for pure seasonal periods.

Although the tests indicate that stochastic seasonality would not be sustained easily and, consequently, the seasonal cointegration technique would be impossible to apply, the theoretical analysis of this technique was performed in Chapter 4. This approach has been chosen given its less restrictive assumptions about the exogeneity of the explanatory variables in the model. Seasonal cointegration through the EGHL approach, using monthly data, has received very little attention, and this approach is chosen because of its transparency and the ease of interpretation of the coefficients.

Seasonal unit root tests on prices of exports, domestic supply, futures and input price of soybeans, maize and wheat in Argentina between 1994 and 2008 were carried out using the same methodology as in Chapter 3. Long-run unit roots cannot be rejected in any of the cases, and there is some weak evidence of additional seasonal unit roots in some of the frequencies. However, given that the seasonal unit roots on prices found do not match the frequencies of the unit roots found in the quantities in Chapter 3, and that the evidence for stochastic seasonality is weak as well, the possibility of seasonal cointegration is rejected. Additionally, as already mentioned, the evidence found of stochastic seasonality in the quantities is not strong. Therefore, the standard cointegration approach is followed, given that all series involved present unit roots at the zero frequency.

In general, the results when cointegration tests were applied on the residuals of the cointegration equations using the ADF test were ambiguous. Different conclusions were obtained depending on the critical values used: rejecting the null of no cointegration when Dickey and Fuller (1979) critical values were used, whilst favouring it when MacKinnon (2010) critical values were considered. Additionally, the violation of the normality assumption in the residuals of the testing equations cannot help to support any of the conclusions. However, attempts made with different types of specification, including intermediate cases, suggest that the case for cointegration is very weak.

Despite the results of the cointegration tests and the ambiguity of some of the other tests, an error correction representation was attempted in order to obtain estimates of the parameters involved in the export and the domestic supply equations developed in the first model. Although the normality assumption has been violated in some of the cases, inference about the parameters of the error correction term in the ECM representation reveals that any of the relationships analysed could be characterised by cointegration. Although some of the signs of the estimates are in line with the theory outlined in Chapter 1 and the economic intuition, the fact that the error correction terms which resulted were not significant (as well as some of the parameters) suggests that it is difficult to sustain a long-term relationship characterised by cointegration between the variables of the model.

Identification of the supply functions to be estimated in the ECM has been attempted using the different climate and cost variables associated with the agricultural activity. These variables were analysed graphically and in terms of the presence of the unit roots. However, the capability of these variables to correctly identify the supply functions has not been, in general, very satisfactory. Apart from the real exchange rate that seems to present an association with the supply functions, the rest of the variables have not provided sufficient explanatory power for the variation in the quantity supplied. These variables have also been accompanied by seasonal deterministic dummies with the objective of capturing the seasonal movements of the series.

The short length of the data considered may be part of the explanation for these results. However, extending the data into the past may not be convenient given important institutional changes in Argentina that may complicate and distort the problem under study; only adding newer data may help. Moreover, it is possible that the most appropriate specification has not been selected among all the specifications considered in terms of the autoregressive lag structure of the dependent and independent variables, used to obtain properly behaving residuals. However, it is highlighted during the analysis that in any case the choice of working with monthly data can be responsible.

On the other hand, given the lack of observable prices, some of those used had to be constructed using different assumptions about their composition that may have introduced some measurement errors affecting the results. If better prices cannot be found, particularly observed prices, the only alternative to this problem lies in attempting to estimate the model using data from a different country and/or different products. In fact, this suggestion is not only based on the difficulties found, but also on the continuous empirical validation that any theory or model should face.

The selection of the EG cointegration approach may also be responsible. Given its error-prone procedure, mistakes that might have been made in the early stages can be carried along later steps without warning. The single-step Johansen-type of approach may address these issues, paying the price of less insight based on its minor interpretational capabilities. Alternatively, the less restrictive fractional cointegration approach may be another way of addressing these estimation issues.

Nevertheless, it is also considered that cointegration may not be the most appropriate approach to estimate this model, given its particularly rigid requirements. Alternative and extensively used dynamic models, such as the Autoregressive Distributed Lag Model, are suggested, given their less stringent requirements. However, the issue of the exogeneity of the explanatory variables, key when supply equations are estimated, should be addressed first in order to apply these techniques.

On the other hand, whilst the possibility that the theoretical model developed in Chapter 1 may not be a good description of how domestic and export supplies behave cannot be ruled out. The fact that this is the first attempt to empirically validate this model suggests that it is too soon to discard it. This also suggests that further empirical validation must be attempted using newer and improved data for the cases under study, a different set of commodities and/or country, an alternative cointegration technique such as Johansen or fractional cointegration, and/or more general dynamic models. These constitute some of the ways, particularly with respect to the validation of this model, which could be followed in the research avenue opened by this thesis.

The parametric approach followed in the development of the model in Chapter 1 has been very useful for the inference and econometric estimation. However, a more general approach in their formulation might have avoided some of the analytical difficulties encountered. This constitutes another line of future research that might shed additional light on the behaviour of the export and domestic supply of agricultural commodities. However, it may require alternative empirical validation efforts outside those presented here. Simulation exercises may be attempted in this case.

The theoretical model can accommodate the reality of the export and domestic supply of agricultural annual crops. Further lines of research should address the situation of non-storable agricultural commodities or of continuous production, as well as considering other commodities, such as oil and minerals, for which futures markets operate. It will first require an analysis of the type of technological risk they face (if any) and of how these markets are structured. The characterisation of traders between supply and demand undertaken in this thesis may not be appropriate in a context where production and commercialisation are vertically integrated; this is typical of oil and mining companies.

However, as the nature of the commodities under study is different, the nature of the time series that might be used for the empirical validation may change as well. This suggests that these series may have different characteristics from those observed in the series used in this thesis. Cycles associated with the life of oil and mining reservoirs may have important effects on the nature of the time series used, and which may require alternative testing and estimation techniques beyond those presented here. The possibility of attempting the estimation with data with more transversal elements, such as cross panels, should not be excluded.

BIBLIOGRAPHY

Amemiya, T., & Wu, R. (1972). 'The Effect of Aggregation on Prediction in the Autorregresive Model'. *Journal of the American Statistical Association, 67*(339), 628-632.

Anderson, R. W., & Danthine, J. P. (1980). 'Hedging and joint production: theory and illustrations'. *The Journal of Finance, 35*(2), 487-498.

Anderson, R. W., & Danthine, J. P. (1981). 'Cross Hedging'. *The Journal of Political Economy, 89*(6), 1182-1196.

Anderson, R. W., & Danthine, J.-P. (1983). 'Hedger diversity in future markets'. *The Economic Journal, 93*, 370-389.

Andrade, I., Clare, A., O'Brien, R., & Thomas, S. (1999). 'Tests for stochastic seasonality applied to daily financial time series'. *The Manchester School, 67*(1), 39-59.

Ardeni, P. (1989). 'Does the Law of One Price really hold for commodity prices?' *American Journal of Agricultural Economics, 71*(3), 661-669.

Arslan, R., Rapetti, M. & Skott, P. (2012). 'The real exchange rate and economic development'. *Structrual Change and Economic Dynamics*, 23 (2), 151-169

Armington, P. S. (1969). 'A theory of demand for products distinguished by place of production'. *International Monetary Fund Staff Papers, 16*(1), 159-178.

Baillie, R., & Bollerslev, T. (1994). 'Cointegration, fractional cointegration and exchange rate'. *The Journal of Finance, 49*(2), 737-745.

Beaulieu, J., & Miron, J. A. (1993). 'Seasonal unit roots in aggregate U.S. data'. *Journal of Econometrics, 55*, 305-328.

Beenstock, M., Goldin, E., & Nabot, D. (1999). 'The demand for electricity in Israel'. *Energy Economics, 21*, 168-183.

Bell, W., & Hilmer, S. (1984). 'Issues Involved with the Seasonal Adjustment of Economic Time Series'. *Journal of Business and Economic Statistics, 2*, 291-230.

Binet, M.-E., & Zaied, Y. B. (2011, November). 'A Seasonal Integration and Cointegration Analysis of Residential Water Demand in Tunisia'. *Working Paper 22, 22*. Rennes, France: University of Rennes 1.

Bohl, M. (2000). 'Nonstationary stochastic seasonality and the German M2 money demand function'. *European Economic Review, 44*, 61-70.

Bohl, M. T., & Sell, F. L. (1998). 'Demand for cash balances in Germany: theoretical underpinnings and empirical evidence'. *Applied Economics, 30*, 1017-1026.

Bolt, J., & van Zanden, J. L., (2013) 'The First Update of the Maddison Project. Re-Estimating Growth before 1820', *Maddison-Project Working Paper* WP-4

Braun, O., & Joy, L. (1968). 'A model of economic stagnation - A case study of the Argentine economy'. *The Economic Journal, 78*(312), 868-887.

Britto, R. (1984). 'The simultaneous determination of spot and futures prices in a simple model with production risk'. *The Quarterly Journal of Economics, 99*(2), 351-365.

Cachanosky, R. (2013, October 2). 'El fracaso del blanqueo'. *La Nacion*.

Calvo, G. A., Izquierdo, A., & Talvi, E. (2003). 'Sudden Stops, the Real Exchange Rate, and Fiscal Sustainability: Argentina's Lessons'. *Working Paper 9828*.

Caminero, E., & Diaz-Emperanza, I. (1997). 'Estimation and testing of cointegration relationships with strongly seasonal monthly data'. *Kybernetika, 33*(6), 607-631.

Caporale, G. M., Cunado, J., & Gil-Alana, L. A. (2012). 'Deterministic versus stochastic seasonal fractional integration and structural breaks'. *Statistics and Computing, 22*(2), 349-358.

Carlton, D. W. (1984). 'Futures Markets: Their purpose, their history, their growth, their success and failures'. *The Journal of Futures Markets, 4*(3).

Carter, C. A. (1999). 'Commodity futures market: A survey'. *The Australian Journal of Agricultural and Resource Economics, 43*(2), 209-247.

Cavallo, A. (2013). 'Online and official price indexes: Measuring Argentina's inflation'. *Journal of Monetary Economics, 60*, 152-165.

Cellini, R., & Cuccia, T. (2009). '*Museum and monument attendace and toursim flow: A time series analysis approach'.* Munich: Munich Personal RePEc Archive.

Chari, V. V., Jogannathan, R., & Jones, L. (1990). 'Price Stability and Futures Trading in Commodities'. *The Quarterly Journal of Economics, 105*(2), 527-534.

Cheung, Y.-W., & Lai, K. (1993). 'A fractional cointegration analysis of purchasing power parity'. *Journal of Business & Economic Statistics, 11*(1), 103-111.

Chow, G. C. (1960). 'Tests of equality between sets of coefficients in two linear regressions'. *Econometrica, 28*(3), 591-605.

Clements, M. P., & Madlener, R. (1999 йил May). 'Seasonality, cointegration, and forescastin UK residential energy demand'. *Scottish Journal of Political Economy, 46*(2), 185-206.

Corden, W. (1984). 'Booming sector and Dutch Disease Economics: Survey and consolidation'. *Oxford Economic Papers, 36*(3), 359-380.

Cox, C. (1976).' Futures, trading and market information'. *Journal of Political Economy, 84*, 1215-1237.

Danthine, J. P. (1978). 'Information, future prices and stabilizing speculation'. *Journal of Economic Theory, 17*, 79-98.

Darne, O. (2004). 'Seasonal cointegration for monthly data'. *Economic Letters, 82*, 349-356.

De la Torre, A., Levy Yeyati, E., Schmukler, S. L., Ades, A., & Kaminsky, G. (2003).' Living and Dying with Hard Pegs: The Rise and Fall of Argentina's Currency Board'. *Economia, 3*(2), 43-107.

de Pablo Valenciano, J., Perez Mesa, J., & Levy Mangin, J. (2008). 'The Spanish Tomato Export Sector in the Almeria Region: An Econometric Approach'. *International Advances in Economic Research, 14*, 316-328.

Deaton, A., & Laroque, G. (1992).' On the Behaviour of Commodity Prices'. *The Review of Economic Studies, 59*(1), 1992.

Diaz Alejandro, C. (1970). '*Essays on the Economic History of the Argentine Republic'.* New Haven: Yale University Press.

Dickey, D., & Fuller, W. (1979).' Distribution of the estimatiors for autorregresive time series with a unit root'. *Journal of the American Statistical Association, 84*, 427-431.

Dornbush, R. (1076). 'Expectations and Exchange Rate Dynamics'. *Journal of Political Economy*, 84 (6)

Eastwood, R., Lipton, M., & Newell, A. T. (2010). 'Farm Size'. In R. Evenson, & P. Pingal (Eds.), *Handbook of Agricultural Economics* (Vol. 4, pp. 6-65). Elsevier.

Enders, W. (1995). '*Applied Econometric Time Series'* (1st Edition ed.). John Wiley & Sons.

Enders, W. (2010). '*Applied Econometric Time Series'* (3rd Edition ed.). Hoboken,, NJ, USA: Jonh Wiley & Sons, Inc.

Engle, R. F., & Granger, C. W. (1987). 'Cointegration and error correction: Representation, estimation and testing'. *Econometrica, 55*, 151-165.

Engle, R. F., Granger, C. W., Hylleberg, S., & Lee, H. S. (1993). 'Seasonal cointegration'. *Journal of Econometrics, 55*, 275-298.

European Commision. (2013). '*Agriculture and Rural Development - Frequently asked questions'.* Retrieved from http://ec.europa.eu/agriculture/faq/

Facciano, G. (2011). '*Los contratos de futuro de soja en el mundo'.* Buenos Aires: Mercado a Termino de Buenos Aires.

Fanelli, J., Frenkel, R., & Rozenwurcel, G. (1990). '*Growth and structural reform in Latin America'. Where we stand.* UNCTAD. Buenos Aires: CEDES.

Franses, P. (1996). 'Recent Advances in Modelling Seasonality'. *Journal of Economic Surveys, 10*(3), 299-345.

Franses, P. H. (1991). '*Model selection and seasonality in time series'.* Amsterdam: Thesis Publishers.

Franses, P. H. (1991). 'Seasonality, non-stationarity and the forescasting of monthly time series'. *International Journal of Forecasting, 7*, 199-208.

Franses, P. H. (1993). 'A method to select between periodic cointegration and seasonal cointegration'. *Economic Letters, 41*, 7-10.

Franses, P. H., & McAleer, M. (1998). 'Cointegration analysis of seasonal time series'. *Journal of Econometric Surveys, 12*(5), 651-678.

Franses, P. H., & Volgelsang, T. J. (1998). 'On seasonal cycles, unit roots and mean shifts'. *The Review of Economics and Statistics, 80*, 231-240.

Franses, P., & Hobijn, B. (1997). 'Critcal values for unit root tests in seasonal time series'. *Journal of Applied Statistics, 24*(1), 25-47.

Franses, P., & Volgesang, T. (1998). 'On seasonal cycles, unit roots and mean shifts'. *The Review of Economics and Statistics, 80*, 231-240.

Frenkel, R., & Rapetti, M. (2008). 'Five years of competitive and stable real exchange rate in Argentina, 2002-2007'. *International review of applied economics*, 22 (2), 215-226

Garbade, K. D., & Silber, W. L. (1983). 'Price Movements and Price Discovery in Futures and Cash Markets'. *The Review of Economics and Statistics, 65*(2), 289-297.

Gemmill, G. (1985). 'Forward contracts or international buffer stocks? A study of their relative efficency in stabilising commodity export earnings'. *The Economic Journal, 378*, 400-417.

Gerchunoff, P., & Llach, L. (1998). '*El ciclo de la ilusion y el desencanto: un siglo de politicas economicas argentinas'.* Buenos Aires: Ariel Sociedad Economica.

Ghysels, E. (1990). 'Unit-Root Tests and the Statistical Pitfalls of Seasonal Adjustment: The Case of U.S'. Postwar Real Gross National Product. *Journal of Business & Economics Statistics, 8*(2), 145-152.

Ghysels, E., & Osborn, D. R. (2001). '*The econometric analysis of seasonal time series'.* Cambridge: Cambridge University Press.

Ghysels, E., Lee, H. S., & Noh, J. (1994). 'Testing for unit roots in seasonal time series'. *Journal of Econometrics, 62*, 415-442.

Giberti, H. C. (1986). '*Historia economica de la ganaderia argentina'.* Buenos Aires: Hyspamerica.

Gil-Alana, L., & Robinson, P. (2001). 'Testing of seasonal fractional integration in UK and Japanese consumption and income'. *Journal of Applied Econometrics, 16*, 95-114.

Granger, C. (1981). 'Some properties of time series data and their use in econometric model specification'. *Journal of Econometrics, 16*, 121-130.

Granger, C. (1986). 'Developments in the study of cointegrated economic variables'. *Oxford Bulletin of Economics and Statistics, 48*(3), 213-228.

Gray, R., & Rutledge, D. (1971). 'The economics of commodity futures markets: a survey'. *Review of Marketing and Agricultural Economics, 39*(4), 57-108.

Grilli, E., & Yang, M. (1988). 'Primary commodity prices, manufactures good prices, and the terms of trade of Developing Countries: What the long run shows'. *World Bank Economic Review, 2*(1), 1-47.

Hall, A. (1994). 'Testing for a unit root in time series with pretest data based model selection'. *Journal of Business and Economic Studies, 12*, 461-470.

Hamori, S., & Tokihisa, A. (2001). 'Seasonal cointegration and the money demand function: some evidence from Japan. '*Applied Economics Letters*(8)', 305-310.

Harris, R., & Sollis, R. (2003). '*Applied Time Series Modelling and Forescasting.* Chichester': John Wiley & Sons Ltd.

Harvey, D. I., Leybourne, S. J., & Newbold, P. (2002). 'Seasonal unit root test with seasonal mean shifts'. *Economic Letters, 76*, 295-302.

Harvey, D., Leybourne, S., & Newbold, P. (2001). 'Innovational outlier unit root tests with an endogenously determined break in level'. *Oxford Bulletin of Economics and Statistics, 63*(5), 559-575.

Hasan, M. S. (2011). 'Seasonal Cointegration and Long-Run Neutrality of Money in the USA'. *Economic Notes by Banca Monte dei Paschi di Siena, 40*(3), 93-105.

Haug, A. (2002). 'Temporal Aggregation and the Power of Cointegration Tests: a Monte Carlo Study'. *Oxford Bulletin of Economics and Statistics, 64*(4), 399-412.

Herwartz, H., & Reimers, H.-E. (2003). 'Seasonal cointegration analysis for German M3 money demand'. *13*, 71-18.

Hirshleifer, D. (1988). 'Risk, futures pricing, and the organization of production in commodity markets'. *The Journal of Political Economy, 96*(6), 1206-1220.

Hirshleifer, D. (1989, September). 'Futures trading, storage and the division of risk: A multiperiod analysis'. *The Economic Journal*, 700-719.

Hirshleifer, D., & Subrahmanyam, A. (1993). 'Futures versus share contracting as means of diversifying output risk'. *The Economic Journal, 103*, 620-638.

Holthausen, D. M. (1979). 'Hedging and the competitive firm under price uncertainty'. *The American Economic Review, 69*(5), 989-995.

Huang, T.-H., & Shen, C.-H. (1999). 'Applying the seasonal error correction model to the demand for international reserves in Taiwan'. *Journal of International Money and Finance, 18*, 107-131.

Hueth, D., & Schmitz, A. (1972). 'International trade in intermediate and final goods: Some welfare implications of destabilized prices'. *The Quarterly Journal of Economics, 86*(3), 352-365.

Hylleberg, S., Engle, R. F., Granger, C. W., & Yoo, B. S. (1990). 'Seasonal Integration and Cointegration'. *Journal of Econometrics, 44*, 215-238.

Jacks, D. S. (2007). 'Populists versus theorists: Futures markets and the volatility of prices'. *Explorations in Economic History, 44*, 342-362.

Johansen, S. (1988). 'Statistical Analysis of Cointegration Vectors'. *Journal of Economic Dynamics and Control, 12*, 231-54.

Johansen, S., & Schaumburg, E. (1999). 'Likelihood analysis of seasonal cointegration'. *Journal of Econometrics, 88*, 301-339.

Jorgenson, D. (1966). 'Rational distributed lag functions'. *Econometrica, 34*(1), 135-149.

Kamara, A. (1982). 'Issues in futures markets: A survey'. *The Journal of Futures Markets, 2*(2), 261-294.

Kawai, M. (1983). 'Price volatility of storable commodities under rational expectations in spot and futures markets'. *International Economic Review, 24*(2), 435-459.

Kawai, M. (1983b). 'Spot and futures prices of nonstorables commodities under rational expectations'. *The Quarterly Journal of Economics, 98*(2), 235-254.

Kawai, M., & Zilcha, I. (1986). 'International trade with forward-futures markets under exchange rate and price uncertainty'. *Journal of International Economics, 20*, 83-98.

Keynes, J. M. (1930). '*A treatise of money* (Vol. II)'. London: Macmillan.

Kunst, R. M. (1993). 'Seasonal Cointegration, Common Seasonals, and Forecasting Seasonal Series'. *Empirical Economics, 18*, 761-776.

Kunst, R. M., & Franses, P. H. (2011). 'Testing for Seasonal Unit Roots in Monthly Panels of Times Series'. *Oxford Bulletin of Economics and Statistics, 73*(4), 469-488.

Leuthold, R. (1974). 'The price perfomance on the futures markets of a nonstorable commodity: Live beef catlle'. *American Journal of Agricultural Economics, 56*(2).

Ljung, G., & Box, G. (1978). 'On a Measure of Lack of Fit in Time Series Models'. *Biometrica, 65*, 297-303.

MacKinnon, J. (1991). 'Critical vallues for co-integration tests'. In R. F. Engle, & C. W. Granger, *Long-Run Economic Relationships* (pp. 267-276). Oxford, UK. : Oxford University Press.

MacKinnon, J. (2010). '*Critical Values for Cointegration Tests'.* Working Paper No. 1227, Queen's University, Department of Economics, Kingston.

Martin-Alvarez, F. J., Cano-Fernandez, V. J., & Caceres-Hernandez, J. J. (1999). 'The introduction of seasonal unit roots and cointegration to test index aggregation optimality: And application to a Spanish farm price index'. *Empirical Economics, 24*, 403-414.

Massell, B. F. (1969). 'Price Stabilization and Welfare'. *Quarterly Journal of Economics, 83*(2), 284-298.

McCallum, J. (1995). 'National borders matter: Canada-U.S. regional trade patterns'. *The American Economic Review, 85*(3), 615-623.

McErlean, S., Wu, Z., Moss, J., IJpellar, J., & Doherty, A. (2003). 'Do EU direct payments to beef producers belong in the 'blue box'?' *The Australian Journal of Agricultural and Resource Economics, 47*(1), 55-73.

Moffett, M. (2012, June 13). 'Dollars Become Scarce as Argentina Cries Peso'. The Wall Street Journal.

Mugambe, G. K., & Reilly, B. (2007). 'Seasonality and Industrial Production in Uganda'. *African Development Review, 19*(3), 501-518.

Nerlove, M. (1972). 'Lags in economic behavior'. *Econometrica, 40*(2), 221-251.

Netz, J. (1995). 'The effect of futures markets and corners on storage and spot price variability'. *American Journal of Agricultural Economics, 77*(1), 182-193.

Newbery, D. M. (1987). 'When do futures destabilize spot prices?' *International Economic Review, 28*(2), 291-297.

Newbery, D. M., & Stiglitz, J. E. (1981). '*The theory of commodity price stabilization: A study in the economics of risk'.* Oxford: Clarendon Press.

Newbery, D., & Stiglitz, J. (1979).' The theory of commodity price stabilisation rules: welfare impacts and supply responses'. *The Economic Journal, 356*, 799-817.

Obstfeld, M., & Rogoff, K. (2001). 'The six major puzzles in international macroeconomics: Is there a common cause?' In B. Bernanke, & K. Rogoff, *NBER Macroeconomics Annual 2000* (Vol. 15, pp. 339-412). MIT Press.

Olivo, S. L. (2010). 'Condiciones para el desarrollo de los mercados de futuro'. *Documento de Trabajo 420*.

Osborn, D. (2002). 'Cointegration for Seasonal Time Series Processes'.

Osborn, D. R. (1993). 'Seasonal cointegration'. *Journal of Econometrics, 55*, 299-303.

Ouerfelli, C. (2008). 'Co-integration analysis of quarterly European toursim demand in Tunisia'. *Tourism Management*(29), 127-137.

Peck, A. E. (1976). 'Futures markets, supply response, and price stability'. *The Quarterly Journal of Economics, 90*(3), 407-423.

Perez-Pascual, P., & Sanz-Carnero, B. (2009). 'Law of one price: evidence from the Spanish wheat market'. *The Annals of Regional Science, 47*, 329-351.

Perron, P. (1988). 'Trends and rando walks in macroeconomic time series: Further evidence from a new approach'. *Journal of Economic Dynamics and Control, 12*, 297-332.

Perron, P. (1989). 'The Great Crash, the Oil Shock and the unit root hypothesis'. *Econometrica, 57*, 1361-1402.

Perron, P. (1997). 'Further evidence on breaking trend functions in macroeconomic variables'. *Journal of Econometrics, 80*, 355-385.

Perron, P., & Vogelsang, T. (1992). 'Nonstationarity and level shifts with an application to purchasing power parity'. *Journal of Business & Economics Statistics, 10*(3), 301-320.

Pierce, D. A. (1979). 'Seasonal Adjustment when both Deterministic and Stochastic Seasonality are Present. In A. Zellner, *Seasonal Analysis of Economic Time Series* (pp. 242-280)'. National Bureau of Economic Research.

Prebisch, R. (1959). 'Commercial Policy in the Underdeveloped Countries'. *The American Economic Review, 49*(2), 251-273.

Rodrigues, P. M., & Osborn, D. R. (1999). 'Performance of seasonal unit root tests for monthly data'. *Journal of Applied Statistics, 26*(8), 985-1004.

Rodrik, D. (2006). 'Goodbye Washington Consensus, Hello Washington Confusion? A review of the World Bank's Economic Growth in the 1990s: Learning from a Decade of Reform'. *Journal of Economic Literature, 44*(4), 973-987.

Rossana, R., & Seater, J. (1995). 'Temporal Aggregation and Economic Time Series'. *Journal of Business and Economics Statistics, 13*(4), 441-451.

Shiller, R., & Perron, P. (1985). 'Testing the Random Walk Hypothesis: Power versus Frequency of Observations'. *Economic Letters, 18*, 381-186.

Silvestrini, A., & Veredas, D. (2008). 'Temporal Aggregation of Univariate and Multivariate Time Series Models: A Survey'. *Journal of Economic Surveys, 22*(8), 458-497.

Smith, J., & Otero, J. (1997). 'Structural breaks and seasonal integration'. *Economic Letters, 56*, 13-19.

Stein, J. L. (1979). 'Spot, forward and futures'. *Research in Finance, 1*, 225-310.

Street, J. H. (1974). 'The Ayres-Kuznets Framework and Argentine Dependency'. *Journal of Economic Issues, 3*(4), 707-728.

Tasseven, O. (2008). 'Modelling Seasonality - An Extension of the HEGY Approach in the Presence of Two Structural Breaks'. *Panoeconomicus*(4), 465-484.

Taylor, A. M. (1998). 'Testing for unit roots in monthly time series'. *Journal of Time Series Analysis, 19*(3), 349-368.

The Economist (2012, February 25) 'Don't lie to me, Argentina'.

The Economist (2013, July 4) ‘A fistful of financial instruments’.

The Guardian (2012, December 17) 'Argentina faces IMF penalties over failure to meet inflation deadline'

Tiffin, R., & Dawson, P. J. (2000). 'Structural breaks, cointegration and the farm-retail price for lamb'. *Applied Economics, 32*, 1281-1286.

Tokihisa, A., & Hamori, S. (2001). 'Seasonal Integration for daily data'. *Econometric Reviews, 20*(2), 187-200.

Turnovsky, S. (1979). 'Futures markets, private storage and price stabilization'. *Journal of Public Economics, 12*, 301-327.

Turnovsky, S. J. (1983). 'The Determination of spot and futures prices with storable commodities'. *Econometrica, 51*(5), 1363-1387.

Turnovsky, S. J., & Campbell, R. B. (1985). 'The stabilising and welfare properties of futures markets: A simulation approach'. *International Economic Review, 92*(2), 227-303.

Viaene, J. M., & Zilcha, I. (1998). 'The behavior of competitive exporting firms under multiple uncertainty'. *International Economic Review, 39*(3).

Wilcox, D. (1992). 'The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work'. *American Economic Review, 82*, 922-941.

Williamson, J. (1990). 'What Washington Means by Policy Reform. In J. Williamson (Ed.), *Latin American Adjustment: How Much Has Happened?'* Peterson Institute for International Economics.

Working, H. (1949). 'The Theory of the price of storage'. *American Economic Review, 39*, 591-609.

Wright, B. D., & Williams, J. C. (1982). 'The economic role of commodity storage'. *The Economic Journal, 92*(367), 596-614.

Zivot, E., & Andrews, D. (1992). 'Further evidence on Great Crash, the oil price shock and the unit root hypothesis'. *Journal of Business and Economic Statistics, 10*, 251-270.

APPENDICES

1.1. APPENDIX 1

DEFINITION OF THE TRANSFORMED VARIABLES IN THE HEGY TEST WITH MONTHLY DATA

The $Z_{k,t}$variables presented in expression are defined below. For further details, see Beaulieu and Miron (1993)

$$Z_{1,t} = (1 + B + B^2 + B^3 + B^4 + B^5 + B^6 + B^7 + B^8 + B^9 + B^{10} + B^{11})y_t$$

$$Z_{2,t} = -(1 - B + B^2 - B^3 + B^4 - B^5 + B^6 - B^7 + B^8 - B^9 + B^{10} - B^{11})y_t$$

$$Z_{3,t} = -(1 - B + B^5 - B^7 + B^9 - B^{11})y_t$$

$$Z_{4,t} = -(1 - B^2 + B^4 - B^6 + B^8 - B^{10})y_t$$

$$Z_{5,t} = -\frac{1}{2}(1 + B - 2B^2 + B^3 + B^4 - 2B^5 + B^6 + B^7 - 2B^8 + B^9 + B^{10} - 2B^{11})y_t$$

$$Z_{6,t} = \frac{\sqrt{3}}{2}(1 - B + B^3 - B^4 + B^6 - B^7 + B^9 - B^{10})y_t$$

$$Z_{7,t} = \frac{1}{2}(1 - B - 2B^2 - B^3 + B^4 + 2B^5 + B^6 - B^7 - 2B^8 - B^9 + B^{10} + 2B^{11})y_t$$

$$Z_{8,t} = -\frac{\sqrt{3}}{2}(1 + B - B^3 - B^4 + B^6 + B^7 - B^9 - B^{10})y_t$$

$$Z_{9,t} = -\frac{1}{2}\left(\sqrt{3} - B + B^3 - \sqrt{3}B^4 + 2B^5 - \sqrt{3}B^6 + B^7 - B^9 + \sqrt{3}B^{10} - 2B^{11}\right)y_t$$

$$Z_{10,t} = \frac{1}{2}\left(1 - \sqrt{3}B + 2B^2 - \sqrt{3}B^3 + B^4 - B^6 + \sqrt{3}B^7 - 2B^8 + \sqrt{3}B^9 - B^{10}\right)y_t$$

$$Z_{11,t} = \frac{1}{2}\left(\sqrt{3} + B - B^3 - \sqrt{3}B^4 - 2B^5 - \sqrt{3}B^6 - B^7 + B^9 + \sqrt{3}B^{10} + 2B^{11}\right)y_t$$

$$Z_{12,t} = \frac{1}{2}\left(1 + B + 2B^2 + \sqrt{3}B^3 + B^4 - B^6 - \sqrt{3}B^7 - 2B^8 - \sqrt{3}B^9 - B^{10}\right)y_t$$

1.2. APPENDIX II

STATISTICAL TABLES OF THE HEGY TEST WITH MONTHLY DATA

Table II. 1 Critical values from the distribution of test statistics for seasonal unit roots. DGP: $\Delta_{12}\tilde{y}_t = e_t$

Trend	Intercept	Seasonal Dummies	T	`t':π_1				`t':π_2				F:π_{odd},π_{even}			
				0.01	0.025	0.05	0.1	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
Y	N	Y	200	-3.81	-3.06	-2.79	-2.49	-3.40	-3.06	-2.79	-2.49	5.06	6.05	7.02	8.26
			400	-3.90	-3.63	-3.37	-3.10	-3.45	-3.15	-2.88	-2.58	5.33	6.34	7.31	8.55
			∞	-4.01	-3.72	-3.46	-3.16	-3.50	-3.20	-2.91	-2.61	5.47	6.51	7.48	8.74
N	N	Y	200	-3.31	-2.98	-2.72	-2.42	-3.32	-3.01	-2.77	-2.48	5.06	6.06	7.01	8.27
			400	-3.37	-3.08	-2.83	-2.52	-3.41	-3.12	-2.86	-2.57	5.33	6.36	7.31	8.57
			∞	-3.46	-3.15	-2.87	-2.57	-3.47	-3.15	-2.91	-2.61	5.46	6.49	7.48	8.74
N	N	N	200	-1.43	-1.10	-0.82	-0.49	-2.59	-2.26	-1.97	-1.67	2.35	3.06	3.77	4.71
			400	-1.15	-0.86	-0.61	-0.33	-2.66	-2.33	-2.06	-1.74	2.45	3.18	3.90	4.87
			∞	-0.98	-0.72	-0.49	-0.23	-2.71	-2.40	-2.13	-1.80	2.53	3.26	4.01	5.01
N	Y	N	200	-3.31	-2.99	-2.74	-2.45	-2.58	-2.26	-1.96	-1.65	2.34	3.03	3.73	4.70
			400	-3.39	-3.09	-2.82	-2.51	-2.67	-2.33	-2.04	-1.73	2.43	3.16	3.88	4.85
			∞	-3.52	-3.18	-2.91	-2.59	-2.71	-2.36	-2.11	-1.79	2.52	3.25	3.99	4.98
N	Y	Y	200	-3.28	-3.00	-2.75	-2.44	-3.34	-3.04	-2.78	-2.49	5.04	6.04	6.99	8.24
			400	-3.40	-3.08	-2.81	-2.51	-3.43	-3.11	-2.86	-2.56	5.33	6.35	7.32	8.57
			∞	-3.48	-3.17	-2.89	-2.58	-3.46	-3.15	-2.90	-2.61	5.47	6.51	7.48	8.73
Y	Y	N	200	-3.82	-3.53	-3.30	-3.01	-2.58	-2.25	-1.98	-1.66	2.32	3.01	3.72	4.66
			400	-3.95	-3.64	-3.40	-3.10	-2.68	-2.37	-2.08	-1.74	2.43	3.15	3.88	4.85
			∞	-4.00	-3.69	-3.45	-3.15	-2.71	-2.37	-2.11	-1.78	2.52	3.25	3.99	4.98
Y	Y	Y	200	-3.85	-3.54	-3.29	-3.01	-3.36	-3.06	-2.78	-2.49	5.05	6.06	7.00	8.25
			400	-3.96	-3.64	-3.39	-3.11	-3.45	-3.12	-2.87	-2.58	5.34	6.36	7.32	8.57
			∞	-4.01	-3.72	-3.46	-3.16	-3.50	-3.20	-2.91	-2.61	5.47	6.51	7.48	8.74

Table II. 2 Critical values from the distribution of test statistics for seasonal unit roots with zero values under the presence of unknown structural break with a standard normal break size

$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b)$ $\quad i = 1,2$ and $\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b)$														
		Seasonal	\`t':π_1				\`t':π_2				F:π_{odd},π_{even}			
Trend	Intercept	Dummies	0.01	0.025	0.05	0.1	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
Y	N	Y	-5.54	-4.93	-4.56	-4.16	-4.99	-4.58	-4.27	-3.93	11.16	12.91	14.98	17.82
N	N	Y	-5.91	-5.18	-4.55	-3.95	-5.02	-4.61	-4.29	-3.95	11.29	13.13	15.21	18.11
N	N	N	-2.83	-2.41	-2.07	-1.73	-4.10	-3.70	-3.40	-3.07	6.61	8.02	9.44	11.26
N	Y	N	-6.15	-5.41	-4.69	-4.05	-4.08	-3.67	-3.37	-3.03	6.58	7.95	9.31	11.14
N	Y	Y	-5.91	-5.19	-4.54	-3.94	-5.00	-4.61	-4.28	-3.94	11.27	13.07	15.09	17.86
Y	Y	N	-5.76	-5.17	-4.64	-4.24	-4.02	-3.59	-3.34	-3.01	6.49	7.86	9.12	10.91
Y	Y	Y	-5.55	-4.93	-4.54	-4.13	-4.98	-4.61	-4.27	-3.92	11.17	12.92	14.92	17.74
$\hat{T}_b = argmax\, F_\theta(T_b)$														
			\`t':π_1				\`t':π_2				F:π_{odd},π_{even}			
Trend	Intercept	Seasonal Dummies	0.01	0.025	0.05	0.1	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
Y	N	Y	-5.04	-4.59	-4.14	-3.69	-4.50	-3.99	-3.66	-3.23	8.51	10.33	11.96	14.30
N	N	Y	-5.56	-4.75	-4.16	-3.47	-4.50	-4.03	-3.68	-3.30	8.73	10.57	12.26	14.80
N	N	N	-2.44	-2.03	-1.64	-1.23	-3.65	-3.27	-2.94	-2.55	4.84	6.20	7.57	9.35
N	Y	N	-6.14	-5.34	-4.62	-3.87	-3.56	-3.11	-2.80	-2.43	4.56	5.82	7.13	8.92
N	Y	Y	-5.58	-4.75	-4.15	-3.46	-4.51	-4.04	-3.69	-3.29	8.70	10.57	12.25	14.53
Y	Y	N	-5.72	-5.10	-4.53	-4.00	-3.39	-3.07	-2.70	-2.36	4.39	5.55	6.88	8.64
Y	Y	Y	-5.04	-4.58	-4.14	-3.69	-4.50	-3.99	-3.66	-3.23	8.51	10.36	11.95	14.32

Table II. 3 Critical values from the distribution of test statistics for seasonal unit roots with zero values under the presence of unknown structural break with a large break

$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b)$ $i = 1,2$ and $\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b)$														
		Seasonal	`t':π_1				`t':π_2				F:π_{odd},π_{even}			
Trend	Intercept	Dummies	0.01	0.025	0.05	0.1	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
Y	N	Y	-5.53	-4.94	-4.42	-3.80	-7.12	-6.63	-6.18	-5.75	38.41	43.34	48.19	54.80
N	N	Y	-5.55	-4.96	-4.41	-3.79	-7.40	-6.81	-6.40	-5.95	40.59	46.08	51.44	59.10
N	N	N	-5.60	-4.95	-4.39	-3.65	-6.93	-6.28	-5.78	-5.24	33.29	39.08	44.98	53.10
N	Y	N	-5.95	-5.28	-4.64	-3.92	-6.87	-6.23	-5.73	-5.19	32.45	38.02	43.71	51.14
N	Y	Y	-5.59	-4.99	-4.45	-3.80	-7.35	-6.83	-6.36	-5.91	40.36	45.80	51.45	59.18
Y	Y	N	-5.83	-5.19	-4.62	-3.93	-6.53	-5.92	-5.41	-4.92	29.42	34.30	38.72	45.03
Y	Y	Y	-5.56	-4.97	-4.45	-3.81	-7.14	-6.66	-6.20	-5.77	38.44	43.32	48.21	54.67
$\hat{T}_b = argmax\, F_\theta(T_b)$														
		Seasonal	`t':π_1				`t':π_2				F:π_{odd},π_{even}			
Trend	Intercept	Dummies	0.01	0.025	0.05	0.1	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
Y	N	Y	-4.37	-3.44	-2.76	-2.02	-5.95	-5.12	-4.43	-3.82	15.68	23.54	30.18	38.20
N	N	Y	-4.50	-3.59	-2.88	-2.13	-6.03	-5.30	-4.69	-4.05	21.26	28.06	34.36	41.43
N	N	N	-3.74	-2.61	-2.00	-1.42	-4.82	-3.93	-3.29	-2.74	7.43	15.38	21.75	28.95
N	Y	N	-4.06	-2.95	-2.30	-1.66	-4.82	-3.93	-3.33	-2.73	7.36	15.37	21.67	29.37
N	Y	Y	-4.66	-3.74	-2.93	-2.16	-5.99	-5.19	-4.58	-3.94	21.37	28.12	34.22	41.55
Y	Y	N	-3.55	-2.64	-2.07	-1.56	-4.52	-3.60	-3.05	-2.54	5.48	8.49	13.64	23.24
Y	Y	Y	-4.41	-3.44	-2.76	-2.03	-5.96	-5.11	-4.44	-3.82	15.69	23.58	30.09	37.86

Table II. 4 Percentage of times particular break dates are selected using $\widehat{T}_b = argmax\ F_\theta(T_b)$

Trend	Y	N	N	N	N	Y	Y
Intercept	N	N	N	Y	Y	Y	Y
Seasonal Dummies	Y	Y	N	N	Y	N	Y
<=T_b-13	11.46	19.84	8.86	8.70	19.47	3.44	11.36
T_b-12	4.12	4.94	1.91	1.65	5.02	1.35	4.19
T_b-11	7.05	8.42	3.06	3.48	7.76	2.55	7.14
T_b-10	18.76	20.71	12.80	12.53	21.06	8.25	18.79
T_b-9	4.69	4.51	3.25	2.87	4.87	2.09	4.66
T_b-8	2.62	2.27	1.24	1.32	2.40	1.03	2.68
T_b-7	2.23	1.77	0.87	1.06	1.85	0.83	2.22
T_b-6	2.37	1.65	0.92	0.95	1.85	0.79	2.25
T_b-5	2.24	1.85	1.05	1.00	1.53	0.96	2.12
T_b-4	2.57	1.91	1.32	1.30	2.04	1.38	2.53
T_b-3	2.80	2.46	2.01	2.18	2.38	2.00	2.71
T_b-2	3.53	2.64	3.09	3.34	2.62	3.39	3.48
T_b-1	3.99	3.40	4.34	4.92	3.30	4.99	4.00
T_b	5.23	4.07	6.97	7.18	4.16	7.18	5.37
>=T_b	26.34	19.56	48.31	47.52	19.69	59.77	26.50

Table II. 5 Critical values from the distribution of test statistics for seasonal unit roots under the presence of unknown structural break using single dichotomic variable for identification. Standard normal break size.

$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b)$ $i = 1,2$ and $\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b)$															
		Seasonal	`t':π_1				`t':π_2				F:π_{odd},π_{even}				
Trend	Intercept	Dummies	0.01	0.025	0.05	0.1	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
Y	N	Y	-5.80	-5.20	-4.75	-4.33	-3.41	-3.09	-2.79	-2.49	5.29	6.43	7.54	9.05
N	N	Y	-5.92	-5.21	-4.60	-4.02	-3.53	-3.15	-2.83	-2.52	5.36	6.52	7.65	9.21
N	N	N	-2.72	-2.40	-2.07	-1.74	-3.01	-2.61	-2.29	-1.92	3.73	4.92	6.17	7.88
N	Y	N	-6.13	-5.40	-4.80	-4.18	-3.02	-2.60	-2.29	-1.93	3.67	4.84	6.02	7.72
N	Y	Y	-5.77	-5.12	-4.60	-4.01	-3.48	-3.12	-2.83	-2.50	5.31	6.43	7.47	9.02
Y	Y	N	-5.90	-5.32	-4.90	-4.43	-3.08	-2.62	-2.29	-1.93	3.72	4.88	6.09	7.80
Y	Y	Y	-5.69	-5.11	-4.72	-4.30	-3.49	-3.11	-2.84	-2.51	5.27	6.39	7.41	8.93
$\hat{T}_b = argmax\, F_\theta(T_b)$															
		Seasonal	`t':π_1				`t':π_2				F:π_{odd},π_{even}				
Trend	Intercept	Dummies	0.01	0.025	0.05	0.1	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
Y	N	Y	-5.76	-5.15	-4.70	-4.29	-3.29	-2.99	-2.68	-2.39	4.87	5.94	6.97	8.45
N	N	Y	-5.90	-5.18	-4.57	-3.99	-3.43	-3.03	-2.73	-2.43	4.91	6.00	7.08	8.58
N	N	N	-2.59	-2.18	-1.80	-1.38	-2.97	-2.56	-2.25	-1.87	3.43	4.56	5.79	7.41
N	Y	N	-6.12	-5.36	-4.77	-4.13	-2.92	-2.49	-2.20	-1.84	3.33	4.44	5.56	7.17
N	Y	Y	-5.73	-5.08	-4.55	-3.98	-3.37	-3.03	-2.74	-2.41	4.88	5.93	6.94	8.34
Y	Y	N	-5.85	-5.25	-4.79	-4.36	-2.97	-2.51	-2.19	-1.83	3.34	4.45	5.57	7.23
Y	Y	Y	-5.62	-5.06	-4.67	-4.24	-3.36	-3.01	-2.74	-2.42	4.84	5.90	6.90	8.35

Table II. 6 Critical values from the distribution of test statistics for seasonal unit roots under the presence of unknown structural break using single dichotomic variable for identification. Large break size.

			$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b)$ $i = 1,2$ and $\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b)$											
		Seasonal	\`t':π_1				\`t':π_2				F:π_{odd},π_{even}			
Trend	Intercept	Dummies	0.01	0.025	0.05	0.1	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
Y	N	Y	-3.44	-3.14	-2.81	-2.41	-5.85	-5.31	-4.80	-4.21	19.17	24.67	30.42	38.28
N	N	Y	-2.98	-2.04	-1.39	-0.76	-6.51	-5.86	-5.30	-4.64	25.49	32.57	40.30	50.95
N	N	N	-2.59	-1.88	-1.39	-0.88	-6.54	-5.66	-4.86	-4.16	21.58	29.37	38.15	50.45
N	Y	N	-2.70	-2.11	-1.61	-1.16	-6.28	-5.34	-4.63	-3.89	18.16	24.77	32.27	42.98
N	Y	Y	-2.80	-2.02	-1.37	-0.73	-6.60	-5.92	-5.31	-4.66	25.48	32.58	40.02	50.32
Y	Y	N	-4.26	-4.01	-3.82	-3.57	-5.33	-4.56	-4.02	-3.37	12.27	16.90	22.17	29.06
Y	Y	Y	-3.39	-3.04	-2.75	-2.39	-6.01	-5.29	-4.79	-4.21	19.27	24.64	30.12	37.32
			$\hat{T}_b = argmax\, F_\theta(T_b)$											
		Seasonal	\`t':π_1				\`t':π_2				F:π_{odd},π_{even}			
Trend	Intercept	Dummies	0.01	0.025	0.05	0.1	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
Y	N	Y	-2.68	-1.62	-0.95	-0.38	-3.57	-3.16	-2.85	-2.49	4.15	5.42	6.93	9.24
N	N	Y	-2.55	-1.57	-0.94	-0.33	-3.55	-3.18	-2.87	-2.52	4.16	5.45	6.98	9.12
N	N	N	-2.10	-1.47	-1.04	-0.55	-2.79	-2.40	-2.06	-1.69	2.65	3.52	4.45	5.71
N	Y	N	-2.25	-1.60	-1.16	-0.68	-2.79	-2.40	-2.05	-1.69	2.62	3.47	4.38	5.65
N	Y	Y	-2.52	-1.52	-0.94	-0.31	-3.64	-3.22	-2.88	-2.52	4.16	5.44	6.99	9.05
Y	Y	N	-2.26	-1.69	-1.23	-0.68	-2.83	-2.42	-2.06	-1.69	2.64	3.52	4.37	5.60
Y	Y	Y	-2.37	-1.56	-0.93	-0.37	-3.56	-3.17	-2.86	-2.49	4.17	5.49	6.92	9.05

Table II. 7 Percentage of times particular break dates are selected using $\widehat{T}_b = argmax\ F_\theta(T_b)$ and using a single dichotomic variable for identification.

Trend	Y	N	N	N	N	Y	Y
Intercept	N	N	N	Y	Y	Y	Y
Seasonal Dummies	Y	Y	N	N	Y	N	Y
<=T_b-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
T_b	0.00	0.00	0.00	0.00	0.00	0.00	0.00
T_b+1	0.01	0.00	0.00	0.00	0.00	0.00	0.00
T_b+2	89.39	90.04	88.84	88.62	89.76	88.26	89.89
T_b+3	7.89	7.56	8.35	8.52	7.72	8.66	7.62
>=T_b+4	2.71	2.40	2.81	2.86	2.52	3.08	2.49