Seasonal Unit Roots and Structural Breaks in agricultural time series: Monthly exports and domestic supply in Argentina

Mendez Parra, Maximiliano

University of Sussex

March 2015

SEASONAL UNIT ROOTS AND STRUCTURAL BREAKS IN AGRICULTURAL TIME SERIES: MONTHLY EXPORTS AND DOMESTIC SUPPLY IN ARGENTINA[1]

Maximiliano Mendez-Parra

2015

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.

---

[1] A longer version of this paper can be found in Mendez Parra, Maximiliano "*Futures prices, trade and domestic supply of agricultural commodities*." Doctoral Dissertation (2015), University of Sussex.

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 paper 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 paper, 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 paper, 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 paper, 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 paper 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.

This paper 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. 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.

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.

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

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

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

$$y_t = a_1y_{t-1} + a_{12}y_{t-12} + \epsilon_t + \beta_1\epsilon_{t-1} \qquad (2)$$

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

$$(1 - a_1L)(1 - a_{12}L^{12})y_t = (1 + \beta_1L)\epsilon_t \qquad (3)$$

Which can be rewritten as

$$y_t = a_1y_{t-1} + a_{12}y_{t-12} - a_1a_{12}y_{t-13} + \epsilon_t + \beta_1\epsilon_{t-1} \qquad (4)$$

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 (4) differs from the one described in (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 (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 (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

$$y_t = p_t + s_t + e_t \qquad (5)$$

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

$$p_t = \sum_{i=1}^{I} \alpha_i c_{it} \qquad (6)$$

$$s_t = \sum_{j=1}^{J} \delta_j d_{jt} \qquad (7)$$

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 (7) and (6) in (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 (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 paper.

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

$$s_t = \sum_{j=1}^{12} \beta_j d_{jt} = \beta_t \qquad (\beta_t = \beta_{t \pm 12k}, k = 1,2, \dots) \tag{8}$$

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.

$$y_t = \theta_0 + \alpha t + \sum_{i=1}^{12} \alpha_i D_i + e_t \tag{9}$$

Deterministic seasonality exists, following Pierce (1979) if all the α_i in (9) are not zero. Therefore, it is possible to test for the presence of deterministic seasonality with the hypothesis

$$H_0: \alpha_1 = \cdots = \alpha_{12} = 0 \tag{10}$$

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.

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 paper. 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. 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. 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 paper.

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.

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*[2], 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) 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.[3] Let y_t be the monthly series in question, generated by an autoregressive process of the form

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

[2] 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.
[3] A concise explanation of this procedure on quarterly data can be found in Ghysels and Osborn (2001), Harris and Sollis (2003) and Enders (2010).

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

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

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[4], 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

$$\Delta_{12}y_t = \mu + \beta t + \textstyle\sum_{k=2}^{12}\delta_k D_{kt} + \sum_{k=1}^{12}\pi_k Z_{k,t-1} + \sum_{i=1}^{\rho-1}\psi_i \Delta_{12}y_{t-i} + v_t \qquad (11)$$

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.

[4] 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 π/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).

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 paper 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 paper 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 I.1 (in Appendix I) 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.

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 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 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.

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 paper, 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 paper. 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 2, 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 3.2 Results of test for seasonal unit roots in monthly series

	0	π	$\pi/2$	$2\pi/3$	$\pi/3$	$5\pi/6$	$\pi/6$				
	π_1	π_2	$F_{3,4}$	$F_{5,6}$	$F_{7,8}$	$F_{9,10}$	$F_{11,12}$	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.

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

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

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

$$\hat{T}_{b,\pi i} = \underset{T_b}{argmin}\, t_{\pi i}(T_b) \qquad i = 1,2 \tag{13}$$

$$\hat{T}_{b,F_{o,e}} = \underset{T_b}{argmax}\, F_{odd,even}(T_b) \tag{14}$$

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

$$\hat{T}_b = argmax\, F_\theta(T_b) \tag{15}$$

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 (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

$$\hat{T}_b = 4 + argmax\ F_\theta(T_b) \quad (16)$$

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.

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 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[5] (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.

Table 3 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
$\pi/2$	$F_{3,4}$	25.26 *Oct-98*	27.98 *Oct-02*	16.48 *Jul-06*	14.88	20.11	23.59
$2\pi/3$	$F_{5,6}$	18.57 *Aug-98*	21.81 *Aug-97*	13.75 *Aug-06*	10.40	17.73	20.61
$\pi/3$	$F_{7,8}$	19.29 *Sep-02*	24.93 *Feb-04*	18.36 *Jul-06*	12.17	17.71	6.87
$5\pi/6$	$F_{9,10}$	20.40 *Aug-98*	14.36 *Apr-03*	14.58 *Aug-06*	11.88	11.44	10.17
$\pi/6$	$F_{11,12}$	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

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

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.

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.

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.

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.

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.

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.

Beaulieu, J., & Miron, J. A. (1993). 'Seasonal unit roots in aggregate U.S. data'. *Journal of Econometrics, 55*, 305-328.

Bell, W., & Hilmer, S. (1984). 'Issues Involved with the Seasonal Adjustment of Economic Time Series'. *Journal of Business and Economic Statistics, 2*, 291-230.

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.

Chow, G. C. (1960). 'Tests of equality between sets of coefficients in two linear regressions'. *Econometrica, 28*(3), 591-605.

Cox, C. (1976).' Futures, trading and market information'. *Journal of Political Economy, 84*, 1215-1237.

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.

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.

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.

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.

HoFranses, 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.

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.

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.

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.

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.

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.

Hylleberg, S., Engle, R. F., Granger, C. W., & Yoo, B. S. (1990). 'Seasonal Integration and Cointegration'. *Journal of Econometrics, 44*, 215-238.

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.

Mugambe, G. K., & Reilly, B. (2007). 'Seasonality and Industrial Production in Uganda'. *African Development Review, 19*(3), 501-518.

Osborn, D. (2002). 'Cointegration for Seasonal Time Series Processes'.

Osborn, D. R. (1993). 'Seasonal cointegration'. *Journal of Econometrics, 55*, 299-303.

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.

Rodrigues, P. M., & Osborn, D. R. (1999). 'Performance of seasonal unit root tests for monthly data'. *Journal of Applied Statistics, 26*(8), 985-1004.

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.

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.

Wilcox, D. (1992). 'The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work'. *American Economic Review, 82*, 922-941.

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)

$$Z_{1,t} = (1 + B + B^2 + B^3 + B^4 + B^5 + B^6 + B^7 + B^8 + B^9 + B^{10} + B^{11})y_t$$

$$Z_{2,t} = -(1 - B + B^2 - B^3 + B^4 - B^5 + B^6 - B^7 + B^8 - B^9 + B^{10} - B^{11})y_t$$

$$Z_{3,t} = -(1 - B + B^5 - B^7 + B^9 - B^{11})y_t$$

$$Z_{4,t} = -(1 - B^2 + B^4 - B^6 + B^8 - B^{10})y_t$$

$$Z_{5,t} = -\frac{1}{2}(1 + B - 2B^2 + B^3 + B^4 - 2B^5 + B^6 + B^7 - 2B^8 + B^9 + B^{10} - 2B^{11})y_t$$

$$Z_{6,t} = \frac{\sqrt{3}}{2}(1 - B + B^3 - B^4 + B^6 - B^7 + B^9 - B^{10})y_t$$

$$Z_{7,t} = \frac{1}{2}(1 - B - 2B^2 - B^3 + B^4 + 2B^5 + B^6 - B^7 - 2B^8 - B^9 + B^{10} + 2B^{11})y_t$$

$$Z_{8,t} = -\frac{\sqrt{3}}{2}(1 + B - B^3 - B^4 + B^6 + B^7 - B^9 - B^{10})y_t$$

$$Z_{9,t} = -\frac{1}{2}\left(\sqrt{3} - B + B^3 - \sqrt{3}B^4 + 2B^5 - \sqrt{3}B^6 + B^7 - B^9 + \sqrt{3}B^{10} - 2B^{11}\right)y_t$$

$$Z_{10,t} = \frac{1}{2}\left(1 - \sqrt{3}B + 2B^2 - \sqrt{3}B^3 + B^4 - B^6 + \sqrt{3}B^7 - 2B^8 + \sqrt{3}B^9 - B^{10}\right)y_t$$

$$Z_{11,t} = \frac{1}{2}\left(\sqrt{3} + B - B^3 - \sqrt{3}B^4 - 2B^5 - \sqrt{3}B^6 - B^7 + B^9 + \sqrt{3}B^{10} + 2B^{11}\right)y_t$$

$$Z_{12,t} = \frac{1}{2}\left(1 + B + 2B^2 + \sqrt{3}B^3 + B^4 - B^6 - \sqrt{3}B^7 - 2B^8 - \sqrt{3}B^9 - B^{10}\right)y_t$$

APPENDIX I

STATISTICAL TABLES OF THE HEGY TEST WITH MONTHLY DATA

Table I. 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 I. 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

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