Seasonally specific model analysis of UK cereals prices

Philip Kostov* and John Lingard**

* Department of Agricultural and Food Economics, Queen's University Belfast, Newforge Lane, Belfast, BT9 5PX, Northern Ireland, United Kingdom

** School of Agriculture, Food and Rural Development, University of Newcastle, Newcastle upon Tyne, NE1 7RU, United Kingdom

1. Introduction

Many economic time series contain regular patterns associated with seasonal causes (biological, meteorological etc.). The presence of such periodic features in economic datasets, particularly agricultural ones is well recognised. Nevertheless one rarely sees attempts to explicitly model them. Different automatic and semi-automatic methods of seasonal adjustment are sometimes employed to get rid of the seasonal features present in the data. The justification of these practices is that the longer-term behaviour of the time series is the main interest. In practice, we can handle only one time series at a time with seasonal adjustment procedures and the presence of non-linear computations and identifying assumptions in these procedures, may alter the relationships that exist between the variables we wish to consider simultaneously in a multivariate model (Sims, 1974; Wallis, 1974). Recent research points out that seasonal adjustment may significantly alter relevant time series properties such as invertibility (Maravall, 1995), linearity (Ghysels et al., 1996), cointegration (Granger and Siklos, 1995) and short-run co-movements (Cubadda, 1999).

Since there is convincing evidence of seasonal unit roots in common economic time series (Hylleberg et al. 1993), it is important to model them properly. The common practice of adding seasonal dummies to the set of regressors leads to misspecified models when seasonal unit roots are present (Abeysinghe, 1994). The analysis of seasonal cointegration, first proposed by Hylleberg et al. (1990), has gained interest from practitioners, (e.g. surveys by Franses and McAleer (1998) and Brendstrup et al. (2004)). It is preferable to use non-seasonally adjusted data and apply appropriate seasonal filters within the multivariate models we build for these series. Even with single time series, the seasonal adjustment procedure imposes a parametric model for the seasonal component that is independent from the model applied later to the non-seasonal component and can introduce spurious dynamics.

Two main approaches are used in jointly modelling seasonal and non-seasonal components. On the one hand, a deterministic modelling of the intra-annual movements, on the other, the introduction of unit-roots at seasonal frequencies related to the observation periodicity. Specification tests have been designed to distinguish situations in which one approach is more in concordance with the data (Canova and Hansen (1995), Caner (1998), Hasza and Fuller (1982), Hylleberg, Engle, Granger and Yoo (1990) *inter alia*).

When working simultaneously with several seasonally integrated processes in a multivariate set-up, the researcher is confronted with the problem of seasonal cointegration. One must determine the presence of seasonal cointegration at each possible and reasonable frequency and the dimension of each cointegration space to be able to specify and estimate VECM with seasonal error correction terms, possibly polynomial. Maximum likelihood procedures for a seasonally cointegrated process have been developed (Lee, 1992, Johansen and Schaumburg, 1998) Procedures developed for cointegrated processes at frequency of zero can be extended to the case of seasonal cointegration (Cubbada, 2001, Ahn et al., 2004), which involves the use of a complex number framework and complex Brownian processes.

Some have raised questions about the economic meaningfulness of the presence of seasonal unit roots (Sanjuán and Dawson, 2003 referring to Hatanaka (1996)). The main concern is that if two or more series are cointegrated, (otherwise they are not of much interest) then they exhibit a common trend. The presence of complex seasonal unit roots implies non-stationary seasonal patterns. In the presence of a common long-term trend however, the latter bounds the deviations of the seasonal patterns from a deterministic seasonal. In simple words, a non-stationary seasonal should not exist because it may in principle deviate too much from the common trend. Such an argument is however flawed. First, the point Hatanaka (1996) makes, refers to the way in which evolving seasonality is modeled through complex unit roots, not to the plausibility of the non-stationary seasonality itself. The alternative of deterministic seasonality is a dubious assumption:

Incidentally, a common statistical assumption, namely, "that the seasonal influence itself does not change so that a figure which measures the average seasonal movement for a considerable period may be regarded as the normal seasonal variation for any year in that period," is questioned.
(Green, 1935)

To clarify we reconsider the mainstream econometrics literature within the unobserved components framework adopted in this paper.

2. Methodology

We deal with a class of models for seasonal time series in an unobserved components framework, according to which each season follows specific dynamics but is also tied to the remaining seasons by a common disturbance. We adopt the seasonal specific model (Proietti, 2002) which is formulated in the time domain, as opposed to the frequency domain representation of conventional seasonal cointegration models.

It can be shown that this class of models nests the standard nonperiodic difference stationary time series case which admits the traditional decomposition into trends and a seasonal. In the more general case this does not need to be the case and the adopted model is particularly well suited for situations in which one or a group of seasons behave differently. In the latter case the constraint imposed by the trend-seasonal decomposition, namely that the latter component has a mean of zero over a number of consecutive observations equal to the seasonal period is too binding.

To illustrate the latter point take as an example maincrop potatoes prices. These are typically not traded when early potatoes appear in the market. Therefore the price dynamics in the latter period (typically one to two months) is radically different from the rest of the year. If the trough in this period is particularly deep it will drag down the trend and therefore affect the underlying growth in the series. The main idea is that when the information content of the seasons differs with respect to the long run behaviour of the series and if a subgroup is more variable (i.e. they behave more idiosyncratically), they should be appropriately discounted in extracting a nonperiodic signal that expresses the overall tendency of the series.
To clarify this, consider the extreme situation when the value of the series in a particular season can be equal or around some fixed value (e.g. a structural zero), as in the production of some strongly seasonal product. In this case even if some events are observed, they hardly tell us the general dynamics of the series. The zeros can be interpreted as missing values and this is equivalent to setting the variance of the season to infinity. The multivariate extension can deal with peculiar forms of seasonal or periodic cointegration that characterise only a subset of seasons. Since the model is defined in the time domain we move away from the usual notion of seasonal cointegration, which is defined in the frequency domain, and consequently avoid interpretation difficulties related to the complex seasonal unit roots (Tanaka, 1996).

In the seasonal specific model, each season evolves as a random walk, similarly to the evolving seasonal models (Hylleberg and Pagan, 1997). Seasonal specific models introduce periodic features without affecting the possibility of extracting a non-periodic signal, that provides an indication of the long run dynamics in the series. To illustrate the main idea behind the seasonal specific models let us introduce the simplest model of this kind, namely the seasonal specific local level model.
It can be presented as follows:

$$\begin{aligned} y_t &= \mu_{jt} + \xi_{jt} \\ \mu_{j,t+1} &= \mu_{jt} + \eta_{jt} \\ \eta_{jt} &= \eta_t + \eta^*_{jt} \end{aligned} \qquad (1)$$

where T time series y_t (t= 1,2,…,T) are observed with periodicity s. The model states that the series are characterized by seasonally specific levels μ_{jt} (j=1,2,…s).
The latter evolve as random walks. They are however driven by both seasonally specific (η^*_{jt}) and common (η_t) disturbances. The common disturbances bound the seasonal

levels together. Such a representation is not new, since factor models commonly employ it.
The level is observed with superimposed noise ξ_{jt}, which can also be represented as having an error component structure consisting of idiosyncratic and common to all seasons noise, similarly to the level disturbances.
We can think of the modeling procedure in the following way. Let us systematically sample the series so as to build 12 'yearly' time series, one for each month. Then we can model each individual time series as a local level (plus noise) model, which is the simplest unobserved components model (Harvey, 1989). All these models are however linked due to a common disturbance source, which will make them vary together. The common variation source induces common trends which can be represented via a linear combination of the unobserved trend components in these periodic models. Then obviously the remainder (i.e. the total minus the common trend weight) will form a weighted combination of these seasonal components that is devoid of long run dynamics. In this way we can obtain a decomposition into periodic (i.e. seasonal) and non-periodic (i.e. long-term) components.

The above is illustrated by rewriting (1) in state space form, which is given below[1]:

$$\begin{aligned} y_t &= x_t'\mu_t + \xi_t \\ \mu_{t+1} &= \mu_t + i\eta_t + \eta_t^* \\ Var\left(\eta_t^*\right) &= \mathrm{N} \end{aligned} \qquad (2)$$

where the vector $x_t' = [0,...,0,1,0,...0]$ selects the season. Evidently $x_t = x_{t-s}$. Moreover the seasonally specific levels are stacked together in the sx1 vector $\mu_t = [\mu_{1t},....,\mu_{1t}]'$, i is an sx1 vector of ones, and the seasonally specific disturbances, as well as the noise component are similarly stacked in row vectors. Typically the covariance matrix N will be diagonal, although specifications allowing for correlated idiosyncratic disturbances are possible. We use the state space form representation since it allows for convenient estimation via the Kalman filter. The Kalman filter is a recursive algorithm for Gaussian and conditionally Gaussian state space form models which allows for maximum likelihood estimation since it can be used to construct the log-likelihood via the prediction error decomposition (Harvey, 1989).

Since the seasonally specific models can be viewed as structural time series models (Harvey, 1989) fitted to the periodic time series, it is straightforward to generalize the simple seasonally specific local level model in (2) above to include other unobserved components. Initial fitting of a structural time series model with a seasonal component can be used as a starting point in specifying the appropriate seasonally specific model. Similar to the unobserved components literature, it often will appear that a reasonable choice is the seasonally specific counterpart of the basic structural model (Harvey, 1989),

[1] The state space form representation is not unique (Harvey, 1989). We present the form used in the estimation process. Furthermore the state space form representation used follows Koopman et al. (1999) rather than Harvey(1989).

namely the seasonally specific local trend model. It is a modest extension to the local level model in allowing for (a possibly stochastic) slope component. Assuming that the noise does not have any idiosyncratic component[2], its state space form is:

$$\begin{aligned} y_t &= x_t'\mu_t + \xi_t & \xi_t &\sim NID\left(0, \sigma_\xi^2\right) \\ \mu_{t+1} &= \mu_t + \beta_t + i\eta_t + \eta_t^* & \eta_t^* &\sim NID\left(0, \mathrm{N}_\eta\right) \\ \beta_{t+1} &= \beta_t + i\nu_t + \nu_t^* & \nu_t^* &\sim NID\left(0, \mathrm{N}_\nu\right) \end{aligned} \tag{3}$$

The model presented in (3) has too many sources of variation and for identification reasons needs some restriction to be imposed. It has been shown (Proietti, 2004) that homogeneity (i.e $\mathrm{N}_\nu = q\mathrm{N}_\eta$ for some positive scalar q) is sufficient to provide unique decomposition into periodic and non-periodic components. Furthermore, one may notice that a positive (i.e. nonzero) variance for the slope component implies an I(2) process. When both q and the common slope variance are zero, the model reduces to a deterministic slope model, which implies an I(1) process with deterministic drift. The multivariate extensions are straightforward within the line of the seemingly unrelated time series equations for each of the univariate time series. The only additional parameters necessary for multivariate seasonally specific models are correlation coefficients for the contemporaneous correlation between the disturbances of the individual time series.

$$\begin{aligned} y_{kt} &= x_{kt}'\mu_{kt} + \xi_{kt} \\ \mu_{kt+1} &= \mu_{kt} + \beta_{kt} + i\eta_{kt} + \eta_{kt}^* \\ \beta_{kt+1} &= \beta_{kt} + i\nu_{kt} + \nu_{kt}^* \end{aligned} \tag{4}$$

with

$$\begin{aligned} \xi_{kt} &\sim NID\left(0, \sigma_{k\xi}^2\right) \\ \eta_{kt} &\sim NID\left(0, \sigma_{k\eta}^2\right) \\ \eta_{kt}^* &\sim NID\left(0, \mathrm{N}_{k\eta}\right) \\ \nu_{kt} &\sim NID\left(0, \sigma_{k\nu}^2\right) \\ \nu_{kt}^* &\sim NID\left(0, \mathrm{N}_{k\nu}\right) \end{aligned}$$

and the symmetric *k*x*k* matrices[3] Θ_ξ, Θ_η, Θ_ν collecting the contemporaneous correlations between the corresponding non-periodic disturbances. Similarly the correlations between the different seasonally specific disturbances will be collected in the s symmetric *k*x*k* matrices $\Theta_{\eta^* j}$ and $\Theta_{\nu^* j}$ (*j=, 1,…s*). If homogeneity is imposed across the individual time series the matrices $\Theta_{\nu^* j}$ become superficial.

[2] This assumption is employed in the empirical application.
[3] These are defined simply as storage devices. Upper or lower triangular matrices, stacked vectors or a list of correlation coefficients can be used instead.

One may note that the correlation matrices $\Theta_{\eta^* j}$ have the same rank as the covariance matrices corresponding to the seasonal components. Therefore inference about common seasonal features can be based directly on the rank of the seasonal correlation matrices $\Theta_{\eta^* j}$. Rank restrictions in a state space form model are in principle testable using standard LM tests. Such an inference is simple in bivariate systems, since seasonal cointegration reduces to the simple requirement that for some j the idiosyncratic disturbances are perfectly correlated (i.e. the correlation coefficients are ± 1). Then there is a common single source of seasonal specific disturbances in season j and there exists a linear combination that contains no idiosyncratic features corresponding to those seasons. In this case the non-periodic component that can be extracted from that linear combination will depend solely on that season. In the homoscedastic case (when $\mathrm{N}_{k\eta} = \sigma^2_{\eta^*} I_s$ where I_s is an *sxs* unit matrix, i.e. when the disturbance variance is common across the seasons) perfect correlation amongst the seasonally specific disturbances of the k series implies that the series are seasonally cointegrated (i.e. there exist a linear combination that only displays deterministic seasonality). Similarly perfect correlation amongst some seasonal disturbances, but not amongst others means that there is seasonal cointegration only at some frequencies.

In addition to the convenience of specifying the seasonal component in the time domain, another advantage of the class of models is that the unit root framework, which is superimposed in conventional modeling (after unit root testing, which may not be consistent with the later modeling) is part of the model itself. The presence of unit roots in an unobserved components framework is equivalent to positive variance components, which is a testable assumption within the modeling procedure. In general a bootstrap score test that is equivalent to the locally best invariant test can be constructed (Koopman et al., 1999: 46). The latter is known to be asymptotically pivotal (Tanaka, 1996, Chapter 10.7).

Common trends and common features in unobserved components frameworks simply mean that the disturbance matrices driving these trends or other features (e.g. seasonality) have less than full rank. Such restrictions are easily modeled and tested. Different non-stationary components (and thus multiple sources of variability) can be accommodated in this framework without the need for complex units roots used in the conventional seasonality literature.

3. Data

UK wheat and barley monthly prices obtained from the Eurostat NewCronos database for the period January 1969 to March 2004 are used to demonstrate the methodological approach[4]. It is well known that these two price series follow similar dynamics. It is thus expected that they will have common trends (i.e. will be co-integrated). It should also be reasonable to expect that the cointegration property will hold for all seasonal frequencies. A visual inspection of the plots of the wheat and barley prices confirms these expectations. The other important property of these series is the presence of unit roots at the seasonal frequencies. Since the issue of seasonal unit roots is not essential in this

[4] DEFRA updates its price databases in April each year. Eurostat receives this data and makes it available.

framework[5], we will omit it here. Nevertheless the results from an extensive set of such tests is available upon request. Given the seasonal unit roots, a question arises whether these two series are fully co-integrated (i.e. co-integrated at all seasonal frequencies).

4. Results

We estimate the homogeneous version of model (4). Although this is a restricted form of the more general model presented earlier, it is easy to interpret in the conventional sense of co-integration. The Ox (Doornik, 2001) version of the SsfPack routines (Koopman *et al*., 1999) for state space form models manipulation and estimation is used.
The variance estimates are presented in table 1.

Table 1 Variance estimates (*10^6) with their standard errors

	Wheat		Barley	
	Variance	SE	Variance	SE
Level	106,960	(53,043.97)	91,607	(43,781.07)
slope	1.5E-199	(0.00000)	0	(0.00000)
Irregular	4,790	(1,757.97)	1,943.9	(921.79)
Seasonal	655.94	(384.21)	748.34	(412.30)

The main issue of interest in this model are the correlation coefficients of the disturbances which are presented in table 2.

Table 2 Correlation coefficients of model disturbances

	Correlation coefficient	SE
Level	0.7503777	(0.04524)
Slope	0.9999993	(0.00000)
Irregular	-0.9999976	(0.26594)
N1	0.9999991	(0.15506)
N2	0.9999993	(0.15236)
N3	0.9999983	(0.15266)
N4	0.9999994	(0.09967)
N5	0.9999991	(0.13535)
N6	0.9999984	(0.19873)
N7	-0.5418890	(0.54275)
N8	-0.9999990	(0.15124)
N9	-0.9999990	(0.12286)
N10	0.9999989	(0.13968)
N11	0.9999994	(0.11148)
N12	0.9999989	(0.15415)

[5] Canova and Hansen's (1995) extension of the KPSS test to seasonal frequencies and the generalisation due to Caner (1998) are explicitly formulated as unobserved components models. Nevertheless, the latter are not fully compatible with the models considered here.

Most of the disturbance components are perfectly correlated, as is expected in the case of full cointegration. Since the slope variance is virtually zero (see table 1) thus ruling out the I(2) possibility, the slopes correlation of one basically reflect the stationary slopes. The perfect correlation of the error terms suggests long-term co-movement (conventional econometric model only contain error term disturbances).
The striking feature of the results is the possible lack of full seasonal co-integration. What is even more surprising however is that only a single month, July, is responsible for this lack of cointegration. In the conventional seasonal cointegration models, such a result is infeasible, since they are defined in the frequency, rather than in the time domain, as in the current model. Strictly speaking the correlation coefficient for July has a very large standard error and is not statistically different from –1. Its unusually large standard error however suggest a possibility for prevailing co-movement of the July wheat and barley prices with occasional deviations from this pattern. The latter large error term causes a relatively big standard error in the error term correlation, and also lack of co-movement in the levels.
It is interesting to enquire what causes this lack of co-integration, and to what extent this is due to the properties of the model specification. The current specification imposes the strong restriction of homogeneity of the seasonal disturbance, which is unrealistic for price time series. Nevertheless, relaxing this assumption (results available upon request) does not reduce the July standard error or improve the residuals diagnostics.
Furthermore, it is clear that in the current model setting the lack of levels co-movement is due to the unusually variable July component. It is possible to restrict the slope correlation coefficient to 1, which results in an estimate of -0.99992 (0.87292) for the July disturbances correlation (full results available from the authors upon request). This confirms that the July seasonal disturbances cause the lack of full co-integration between wheat and barley prices.

The residuals of the current model (as well as its heteroscedastic version) exhibit strong serial correlation. It is possible to incorporate correlated residuals in the modeling context (see Proietti, 2004). For price data which is characterized by strong volatility (aka stochastic volatility) however explicitly modeling the latter may a better alternative.

5. Conclusions

Seasonal features in econometric models are often ignored or treated simplistically. This paper applies a unifying unobserved components approach to modeling seasonality to wheat and barley prices in the UK. Our results reveal that although, broadly speaking, these two time series are co-integrated, they may not necessarily move together in July. Such a result is consistent with the main uses of wheat (food) and barley (feed). The quality of the wheat harvest defines what proportion of wheat cannot be used for its main purpose and is thus reallocated to feed use. The latter is directly attributable to the specific harvest and makes its way through prices immediately before the harvest when its quality can be actually ascertained. In case of weather-related deviations from the expected harvest quality, food wheat and feed barley prices may be subjected to different disturbances. Since the latter are relatively rare, using aggregated data or models that disallow such a possibility (such as the conventional seasonal cointegration models) would belittle the influence of such shocks. They would normally appear as outliers in the modelling context.

References

Abeysinghe, T. (1994) Deterministic Seasonal Models and Spurious Regressions, *Journal of Econometrics*, 61, 259-72.

Ahn S.K., S. Cho and B. Chan Seong (2004) Inference of Seasonal Cointegration: Gaussian Reduced Rank Estimation and Tests for Various Types of Cointegration, *Oxford Bulletin of Economics and Statistics*, 66 (2), 261-284.

Ahn, S.K., and G.C. Reinsel (1994) Estimation of Partially Non-Stationary Vector Autoregressive Models with Seasonal Behavior, *Journal of Econometrics*, 62, 317-350.

Brendstrup, B., S. Hylleberg, M.Nielsen. L. Skipper, and L. Stentoft, (2004) Seasonality in Economic Models, *Macroeconomic Dynamics*, Vol. 8, pp. 362-394.

Caner, M. (1998), A locally optimal seasonal unit-root test., *Journal of Business and Economic Statistics* 16(3) 349-356.

Canova, F. and B.E. Hansen (1995) Are seasonal patterns constant over time? A test for seasonal stability,. Journal of Business and Economic Statistics, 13, 237-252.

Cubadda G. (2001) Complex Reduced Rank Models for Seasonally Cointegrated Time Series, *Oxford Bulletin of Economics and Statistics*, 63, 497-511.

Cubadda, G. (1999) Common Cycles in Seasonal Non-Stationary Time Series, *Journal of Applied Econometrics*, 14, 273-291.

Dickey, D., D. Hasza, and W. Fuller (1984) Testing for unit roots in seasonal time series., *Journal of the American Statistical Association* 79, 355.367.

Doornik, J.A. (2001) *Ox. An Object-Oriented Matrix Programming Language*, Timberlake Consultants Press, London.

Franses, P.H., and M. McAleer (1998), Cointegration Analysis of Seasonal Time Series, *Journal of Economic Surveys*, 12, 651-78.

Ghysels, E., Granger, C.W.J., and P.L. Siklos (1996), Is Seasonal Adjustment a Linear or Nonlinear Data-Filtering Process? (with comments), *Journal of Business and Economic Statistics*, 14, 374-97.

Granger, C.W.J., and P.L. Siklos (1995), Systematic Sampling, Temporal Aggregation, Seasonal Adjustment, and Cointegration: Theory and Evidence, *Journal of Econometrics*, 66, 357-69.

Green, R. M. (1935) Seasonal and Short-Time Fluctuations in Wheat Prices in Relation to the Wheat-Price Cycle, *Kansas Technical Bulletin 39*, September, Kansas Agricultural Experiment Station, Topeca: W.C. Austin [historical document]. Available online at http://www.oznet.ksu.edu/historicpublications/pubs/stb039.pdf

Harvey, A.C. (1989) *Forecasting, Structural Time Series and the Kalman Filter*, Cambridge University Press, Cambridge U.K.

Hylleberg S. and A. Pagan (1997) Seasonal Integration and the Evolving Seasonals Model, *International Journal of Forecasting*, 13 (3), 329-340.

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

Hylleberg, S., Jorgensen, C., and N. K. Sorensen (1993), Seasonality in Macroeconomic Time Series, *Empirical Economics*, Vol. 18, pp. 321-335.

Johansen, S., and E. Schaumburg (1998), Likelihood analysis of seasonal cointegration, *Journal of Econometrics*, 88, pp. 301-339.

Koopman S.J., N. Shepard, and J.A. Doornik (1999) Statistical algorithms for models in state space using SsfPack 2.2, *Econometrics Journal*, 2, 113-166.

Lee, H. S. (1992), Maximum Likelihood Inference on Cointegration and Seasonal Cointegration, *Journal of Econometrics*, 54, pp. 1-47.

Maravall, A. (1995), Unobserved Components in Economic Time Series, in Pesaran, M.H., and M.R. Wickens (eds.) *Handbook of Applied Econometrics. Volume 1. Macroeconomics*, Blackwell, 12-72.

Proietti, T. (2002) Seasonal specific structural time series models, Working paper ECO 2002/10, San Domenico: European University Institute.

Proietti, T. (2004) Seasonal specific structural time series, *Studies in Nonlinear Dynamics and Econometrics*, Berkeley Electronic Press, vol. 8(2), pages 16.

Sanjuán, A. I. and P. J. Dawson (2003) Price transmission, BSE and structural breaks in the UK meat sector**,** *European Review of Agricultural Economics* 30(2): 155-172.

Tanaka, K. (1996) *Time Series Analysis: Nonstationarity and Noninvertible Distribution Theory*, NewYork: Wiley.