Final version, July 2002

An Exact Multivariate Model-based Structural Decomposition

José Casals
Miguel Jerez†
Sonia Sotoca

Universidad Complutense de Madrid

Abstract: We propose a simple and structured procedure for decomposing a vector of time series into trend, cycle, seasonal and irregular components. Contrary to common practice, we do not assume these components to be orthogonal conditional on their past. However, the state-space representation employed assures that their estimates converge to values with null variances and covariances. Null variances are very important, as they assure that the components do not change when the sample increases. This lack of "revisions" is the most important feature of our method, in comparison with most alternative procedures. On the other hand, null covariances provide a solid statistical foundation to the decomposition, as it assures that a given component can be analyzed and interpreted independently of any other component(s). Other convenient properties of our method derive from the use of a state-space approach. First, defining the problem in state-space avoids dependence on particular model specifications, so the same procedure can be applied to a wide class of data representations including ARIMA, VARMAX, univariate transfer functions or structural time series models. Also, state-space methods deal easily with nonstandard situations, such as samples with missing values or constraints upon the structural components. Practical application of the procedure is illustrated both with simulated and real data. A MATLAB toolbox for time series modeling and decomposition is available via Internet.

Keywords: State-space models, Time-series decomposition, Latent factor, Seasonal adjustment

---

† José Casals is Senior Analyst, Caja de Madrid, Pº Castellana 189, 28046 Madrid, Spain (email: jcasalsc@cajamadrid.es). Miguel Jerez (email: mjerez@ccee.ucm.es) and Sonia Sotoca (email: sotoca@ccee.ucm.es) are Associate Professors, Departamento de Fundamentos del Análisis Económico II, Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Madrid, Spain. Sonia Sotoca and Miguel Jerez gratefully acknowledge financial support from CICYT, projects PB98-0789 and PB98-0810 respectively. Enrique Llopis provided the time series used in Section 6. Arthur B. Treadway, Jaime Terceiro, José Luis Gallego, Javier Pérez, Luis Puch and Marcos Bujosa read early drafts of this paper and made many useful suggestions. Valuable feedback from JASA editors and two anonymous referees is gratefully acknowledged.

1. INTRODUCTION.

Many works define the additive structural decomposition of a time series, z_t, as:

$$z_t = t_t + c_t + s_t + \varepsilon_t \tag{1}$$

where:

- t_t is the *trend component*, representing the long-term behavior of the series,
- c_t is the *cyclical component*, describing autocorrelated transitory fluctuations,
- s_t is the *seasonal component*, associated with persistent patterns over seasons, and
- ε_t is an *irregular component*.

The decomposition (1) has been of long-standing interest to time series analysts. For example, in an environmental study it may be important to measure the levels of a pollutant without "contamination" from short-term and/or seasonal fluctuations. In a different framework, a marketeer may want to estimate seasonal fluctuations of sales, to allocate efficiently his promotion budget. Last, statistical bureaus often publish quarterly or monthly economic indicators in seasonally adjusted form, to achieve comparability in the series values corresponding to different periods.

There are two basic approaches to estimate the components on the right-hand-side of (1): *ad-hoc methods* and *model-based methods*.

Ad-hoc methods consist of filtering the series by means of a differential equation, designed to extract the terms generating peaks of spectral power at previously chosen frequency ranges. Probably, the methods in the Census X-11 saga are most successful application of this approach (Shiskin, Young and Musgrave 1967; Findley, Monsell, Bell, Otto and Chen 1998). These methods have a clear advantage in usability, as they do not require any previous analysis. However, their mechanical application may yield a spurious decomposition, as these procedures implicitly assume that different time series follow the same stochastic process. Also, when the components are extracted using a symmetric filter, employing past and future information, the estimates change when the sample increases. This feature - often termed "revision" - is clearly a major problem when the decomposition is applied to generate public statistical information.

Potential inadequacy of *ad-hoc* methods inspired the *model-based methods*, which emphasize the coherence between the properties of the observed series and those of the structural components. This principle is at the basis of three methodologies, the *ARIMA-model-based*, the *Structural Time Series Models* (hereafter, "structural models") approach and the *Forecast Decomposition* method.

ARIMA-model-based techniques (Box, Hillmer and Tiao 1978; Hillmer and Tiao 1982) start from an ARIMA representation for z_t, and afterwards obtain a structural decomposition defined by individual

ARIMA processes for each component, constrained that their sum is observationally equivalent to the model for the time series. The TRAMO-SEATS decomposition (Gómez and Maravall 1996) is a well-known implementation of this approach. Similarly, structural models are directly set up in terms of the components in (1), which are represented by state-space models specified according to the properties of the time series (Harvey 1989; Pole, West and Harrison 1994; Young, Pedregal and Tych 1999) and constrained that their sum yields the time series. Both approaches have many common elements, as they rely on similar hypotheses and use equivalent signal extraction techniques. These coincidences are also in the origin of three shortcomings: first, they assume arbitrary independence restrictions between the components; second, theoretical and empirical components have different properties; and third, the components are affected by revisions. We will detail and discuss these shortcomings in Section 6, when comparing standard procedures with our proposal.

Last, forecast decomposition (Box, Pierce and Newbold 1987; Espasa and Peña 1995) consists of decomposing the h-steps-ahead forecast function of an ARIMA model for z_t into persistent, seasonal and transitory components. This method does not need arbitrary assumptions on the components and is not affected by revisions, as the filter employed depends only on past information. It also has weak points, as the variances of the components are larger than those obtained by symmetric filters and the components are heterogeneous, because the values computed in t are based in a different information set than those computed in any other moment. Also, there is some ambiguity about the choice of the lead-time for the forecast function.

In this paper we use state-space techniques to obtain the structural decomposition of a vector of time series generated by a linear stochastic process, addressing all the shortcomings mentioned above. This is achieved combining different theoretical results - some of them well known, others not so much - from the time series and control literature. Section 2 starts by describing how to obtain a convenient representation of the data generating process, which is a state-space model in a particular block-diagonal innovations form. Building on this model, Section 3 discusses how to characterize the structural components. To do this, we state a one-to-one correspondence between the eigenvalues of the block-diagonal transition matrix and the corresponding frequencies displaying peaks of spectral power. Using consensus ideas about the properties of the components, the state variables with unit eigenvalues are associated with the trend, those with peaks at seasonal frequencies are assigned to the seasonal component and the rest of the states are assigned to the cycle term. The structural components are then defined by unique linear combinations of the state variables, characterized by the coefficients in the observation equation. Section 4 discusses the estimation of the components. As the representation employed is detectable, the symmetric and asymmetric filtered estimates of the components converge to conditionally orthogonal values with null variances. In this sense, the decomposition obtained is exact. Section 5 organizes previous results into a six-stage methodology, which practical application is illustrated by the examples in Sections 6 and 7. Finally, Section 8 discusses previous results in the light of existing methods and suggests some additional applications. It also indicates how to obtain, via Internet, a MATLAB toolbox for time series modeling which includes an implementation of our procedure.

2. STATE-SPACE REPRESENTATION OF THE MODEL.

Assume that a structural decomposition analogous to (1) is to be computed for the $m \times 1$ random vector $\boldsymbol{z}_t$. For reasons that will be discussed in Section 4, our method requires $\boldsymbol{z}_t$ to be the output of a *steady-state innovations* state-space model (hereafter, innovations model) defined by:

$$\boldsymbol{x}_{t+1} = \boldsymbol{\Phi} \boldsymbol{x}_t + \boldsymbol{\Gamma} \boldsymbol{u}_t + \boldsymbol{E} \boldsymbol{a}_t \tag{2}$$

$$\boldsymbol{z}_t = \boldsymbol{H} \boldsymbol{x}_t + \boldsymbol{D} \boldsymbol{u}_t + \boldsymbol{a}_t \tag{3}$$

where:

- $\boldsymbol{x}_t$ is an $n \times 1$ vector of *state variables*,
- $\boldsymbol{u}_t$ is an $r \times 1$ vector of exogenous variables,
- $\boldsymbol{a}_t$ is an $m \times 1$ vector of errors, such that $\boldsymbol{a}_t \sim \text{iid}(\boldsymbol{0}, \boldsymbol{B})$.

The *transition equation* (2) characterizes all the dynamic structure of $\boldsymbol{z}_t$. On the other hand, the *observation equation* (3) describes how $\boldsymbol{z}_t$ is generated by the sum of: a) a linear combination of the dynamic components, given by $\boldsymbol{H}\boldsymbol{x}_t$, b) the instantaneous effect of the exogenous variables, given by $\boldsymbol{D}\boldsymbol{u}_t$, and c) the error $\boldsymbol{a}_t$. Starting from a general time-invariant state-space model, the representation (2)-(3) can be obtained in the following way:

Result 1. If the model for $\boldsymbol{z}_t$ is (or can be expressed as):

$$\boldsymbol{x}^*_{t+1} = \boldsymbol{\Phi} \boldsymbol{x}^*_t + \boldsymbol{\Gamma} \boldsymbol{u}_t + \boldsymbol{E}^* \boldsymbol{w}_t \tag{4}$$

$$\boldsymbol{z}_t = \boldsymbol{H} \boldsymbol{x}^*_t + \boldsymbol{D} \boldsymbol{u}_t + \boldsymbol{C} \boldsymbol{v}_t \tag{5}$$

where the errors $\boldsymbol{w}_t$, $\boldsymbol{v}_t$ are independent of the initial state, $\boldsymbol{x}^*_1$, and such that: a) $\boldsymbol{w}_t \sim \text{iid}(\boldsymbol{0}, \boldsymbol{Q})$, $\boldsymbol{v}_t \sim \text{iid}(\boldsymbol{0}, \boldsymbol{R})$, $\text{cov}(\boldsymbol{w}_t \boldsymbol{v}_t) = \boldsymbol{S}$ for all $t = 1, 2, \ldots, N$ and b) $\boldsymbol{x}^*_1 \sim (\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$. Then under weak assumptions (Casals, Sotoca and Jerez 1999, Theorem 1) $\boldsymbol{z}_t$ is also the output of (2)-(3) with:

$$\boldsymbol{x}_1 \sim (\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1 - \boldsymbol{P}) \tag{6}$$

$$\boldsymbol{P} = \boldsymbol{\Phi} \boldsymbol{P} \boldsymbol{\Phi}^T + \boldsymbol{E}^* \boldsymbol{Q} \boldsymbol{E}^{*T} - \boldsymbol{E} \boldsymbol{B} \boldsymbol{E}^T \tag{7}$$

$$\boldsymbol{E} = (\boldsymbol{\Phi} \boldsymbol{P} \boldsymbol{H}^T + \boldsymbol{E}^* \boldsymbol{S} \boldsymbol{C}^T) \boldsymbol{B}^{-1} \tag{8}$$

$$\boldsymbol{B} = \boldsymbol{H} \boldsymbol{P} \boldsymbol{H}^T + \boldsymbol{C} \boldsymbol{R} \boldsymbol{C}^T \tag{9}$$

Note that model (4)-(5) includes two error terms, $\boldsymbol{w}_t$ and $\boldsymbol{v}_t$, whereas the innovations model (2)-

(3) has only one, $\boldsymbol{a}_t$. This means that both are observationally equivalent but some information that is explicit in the covariance matrices of general state-space model ($\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{S}$) is lost when they are combined in the single noise covariance of the innovations model, $\boldsymbol{B}$.

Result 1 has the advantage of generality, as it supports any model that can be written as a linear fixed-coefficients state-space model, but requires to solve the nonlinear equations (7)-(9). Ionescu, Oara and Weiss (1997) describe efficient and stable procedures to do it, but this can be an undesirable complication for many users. This inconvenience can be avoided when the model for the time series is (or can be written as) a VARMAX process. In this case the matrices in (2)-(3) can be obtained directly from the VARMAX parameters, see the Appendix.

The states in expression (2) may combine the dynamic features of different structural components. Result 2 shows how to restructure the model dynamics in an equivalent and more meaningful representation:

Result 2. If z_t is realized by (2)-(3), then it is also realized by the model:

$$\bar{\boldsymbol{x}}_{t+1} = \bar{\boldsymbol{\Phi}}\bar{\boldsymbol{x}}_t + \bar{\boldsymbol{\Gamma}}\boldsymbol{u}_t + \bar{\boldsymbol{E}}\boldsymbol{a}_t \tag{10}$$

$$z_t = \bar{\boldsymbol{H}}\bar{\boldsymbol{x}}_t + \boldsymbol{D}\boldsymbol{u}_t + \boldsymbol{a}_t \tag{11}$$

where the states in (10)-(11) are related to those in (2)-(3) by a linear transformation $\bar{\boldsymbol{x}}_t = \boldsymbol{U}\boldsymbol{x}_t$, such that $|\boldsymbol{U}| \neq 0$. Accordingly, the matrices in (10)-(11) are related to those in (2)-(3) by the expressions: $\bar{\boldsymbol{\Phi}} = \boldsymbol{U}\boldsymbol{\Phi}\boldsymbol{U}^{-1}$, $\bar{\boldsymbol{\Gamma}} = \boldsymbol{U}\boldsymbol{\Gamma}$, $\bar{\boldsymbol{E}} = \boldsymbol{U}\boldsymbol{E}$, $\bar{\boldsymbol{H}} = \boldsymbol{H}\boldsymbol{U}^{-1}$. Finally, the transformation matrix $\boldsymbol{U}$ should be chosen so that the transformed transition matrix $\bar{\boldsymbol{\Phi}}$ is block-diagonal.

The block-diagonal representation (10)-(11) clarifies the model dynamics, as it decomposes the response of the system to shocks in the inputs, $\boldsymbol{a}_t$ and $\boldsymbol{u}_t$, into several basic movements corresponding to different eigenvalues of the transition matrix. These basic movements represent independent reactions of the state variables.

Probably the Jordan decomposition is the most famous block-diagonalizing algorithm, even if it is very unstable (Golub and Van Loan 1996). We will use a simpler and more stable method, which consists of applying the Schur decomposition to $\boldsymbol{\Phi}$ and diagonalizing the resulting real Schur matrix by solving a system of Sylvester equations (Petkov, Christov and Konstantinov 1991, pp. 103-106).

3. CHARACTERIZATION OF THE STRUCTURAL COMPONENTS.

The literature often characterizes the properties of structural components in the frequency domain. Following Burman (1980), the trend is represented by a peak at low frequencies, the irregular component should display no significant peak at any frequency, the seasonal component has peaks at the basic seasonal frequency and its multiples. By elimination, any other component should be assigned to the cyclic term. The implementation of these ideas in the block-diagonal model (10)-(11) is easy, as the eigenvalues of the transition matrix characterize unambiguously the properties of the states in both time and frequency domains, see e.g., West (1997).

Assume that $\lambda_{j,k} = a_j \pm b_j i$ is a pair of conjugate eigenvalues of $\bar{\mathbf{\Phi}}$ in (10), associated with the states *j* and *k*. Under these conditions, it is easy to show that both states generate a peak in the pseudo-spectrum in the frequency $f_{j,k} = (2\pi)^{-1} \arctan(b_j/a_j)$, where $f_{j,k}$ is in cycles per unit time. There are several particular cases worth considering:

1) If the real part of $\lambda_{j,k}$ is zero, then $b_j/a_j = \infty$, implying that $f_{j,k} = 1/4$ if $b_j > 0$ or $f_{j,k} = 3/4$ if $b_j < 0$.

2) If an eigenvalue is a real number, then $b_j/a_j = 0$, implying that $f_{j,k} = 0$ if $a_j > 0$ or $f_{j,k} = 1/2$ if $a_j < 0$.

3) If $b_j = 0$ and a_j tends to one, the spectral power tends to infinity at $f_{j,k} = 0$.

According to these results, the states in the block-diagonal model can be naturally assigned to the structural components with the rules summarized in Table 1.

[Insert Table 1]

where $\boldsymbol{F_S}$ is the set of seasonal frequencies defined as: $\boldsymbol{F_S} = \{ f_j = k_j/s \ ; \ k_j = 1, 2, \ldots, [s/2] \}$, being s the seasonal period and $[s/2] = s/2$ if s is even, or $[s/2] = (s-1)/2$ if s is odd.

Building on this analysis, the characterization of the structural components can be done as follows. Assume that $\boldsymbol{z}_t$ is generated by the block-diagonal innovations model (10)-(11). Consider also the multivariate extension of (1):

$$\boldsymbol{z}_t = \boldsymbol{t}_t + \boldsymbol{c}_t + \boldsymbol{s}_t + \boldsymbol{d}_t + \boldsymbol{\varepsilon}_t \qquad (12)$$

where $\boldsymbol{t}_t$, $\boldsymbol{c}_t$, $\boldsymbol{s}_t$ and $\boldsymbol{\varepsilon}_t$ are $m \times 1$ vectors of trend, cycle, seasonal and irregular components, respectively. A new term, $\boldsymbol{d}_t$, represents the instantaneous effects of exogenous variables on $\boldsymbol{z}_t$. In the structural decomposition framework, this term is often used to model calendar effects or outliers.

Under such conditions, the components in (12) can be characterized by restructuring the block-diagonal model (10)-(11) as:

$$\begin{bmatrix} \bar{x}^t_{t+1} \\ \bar{x}^c_{t+1} \\ \bar{x}^s_{t+1} \end{bmatrix} = \begin{bmatrix} \bar{\Phi}^t & 0 & 0 \\ 0 & \bar{\Phi}^c & 0 \\ 0 & 0 & \bar{\Phi}^s \end{bmatrix} \begin{bmatrix} \bar{x}^t_t \\ \bar{x}^c_t \\ \bar{x}^s_t \end{bmatrix} + \begin{bmatrix} \bar{\Gamma}^t \\ \bar{\Gamma}^c \\ \bar{\Gamma}^s \end{bmatrix} u_t + \begin{bmatrix} \bar{E}^t \\ \bar{E}^c \\ \bar{E}^s \end{bmatrix} a_t \tag{13}$$

$$z_t = \begin{bmatrix} \bar{H}^t & \bar{H}^c & \bar{H}^s \end{bmatrix} \begin{bmatrix} \bar{x}^t_t \\ \bar{x}^c_t \\ \bar{x}^s_t \end{bmatrix} + D\,u_t + a_t \tag{14}$$

where $\bar{x}^t_t$ is the vector of nonstationary states, $\bar{x}^c_t$ is the vector of stationary (nonseasonal) states and $\bar{x}^s_t$ is the vector of seasonal states. Accordingly, the structural components are defined as:

$$t_t = \bar{H}^t \bar{x}^t_t \tag{15}$$

$$c_t = \bar{H}^c \bar{x}^c_t \tag{16}$$

$$s_t = \bar{H}^s \bar{x}^s_t \tag{17}$$

$$d_t = D\,u_t \tag{18}$$

$$g_t = a_t \tag{19}$$

The components given by (15)-(19) are unique although there are infinite block-diagonalizing matrices. This happens because block-diagonalization is a similar transformation of the original transition matrix.

Last, expressions (12) and (15)-(19) can be interpreted as a decomposition of the one-step-ahead forecast function. To see this, note that (19) defines the irregular component to be the error term of the model. Then the components in (15)-(18) add up to the corresponding fitted value, when computed within the sample, whereas the component in (19) coincides with the residual. When computed out of the sample, (15)-(18) add up to the corresponding forecast.

4. ESTIMATION OF THE STRUCTURAL COMPONENTS.

Given the characterization of the structural components made in Section 3, the problem reduces to obtaining estimates of the state variables and combining them according to (15)-(19). The literature provides two basic algorithms to do this: the Kalman filter and the fixed-interval smoother (Anderson and Moore 1979). These methods differ mainly in the information considered. While the former is a one-sided asymmetric filter, providing estimates of the first and second-order moments of the states conditional on past information, the latter is a two-sided symmetric filter, which exploits all the information in the sample. Therefore, smoothed estimates are more efficient in general. In this paper we use a smoothing algorithm (Casals, Jerez and Sotoca 2000) which has two advantages for this specific application. First, it allows for both, stationary and nonstationary roots. Second, it uses the fixed-coefficients innovations structure of the model to reduce computational costs, see Casals et al. (2000, sec. 3.2).

Denoting the information set up to time *i* as $\boldsymbol{\Omega}^i = \{z_1, z_2, \ldots, z_i, \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_i\}$, and the smoothed estimates by $\boldsymbol{x}_{t|N} = E(\boldsymbol{x}_t | \boldsymbol{\Omega}^N)$ and $\boldsymbol{P}_{t|N} = E[(\boldsymbol{x}_t - \boldsymbol{x}_{t|N})(\boldsymbol{x}_t - \boldsymbol{x}_{t|N})^T | \boldsymbol{\Omega}^N]$, *N* being the sample size, the following result holds:

Result 3. In an innovations model, if the eigenvalues of $\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{H}$ are in or inside the unit circle, the smoothing covariance, $\boldsymbol{P}_{t|N}$, converges to zero as *t* increases. If all the eigenvalues of $\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{H}$ are inside the unit circle, this convergence is exponentially fast.

In summary, Result 3 implies that the states in an innovations model are observable, so their smoothed estimates converge to exact values. The practical interest of this property arises from two facts. First, the estimates for the states and the structural components, which are linear combinations of these states, can be treated as actual data and interpreted separately, as they are mutually independent. Second, the empirical structural components will not be revised, as revisions are proportional to the uncertainty affecting the estimates.

This important property can be derived as a Corollary of Theorem 4.2. in De Souza, Gevers and Goodwin (1986), where detectability is shown to be a necessary and sufficient condition for convergence of the difference Riccati equation to the unique strong solution of the algebraic Riccati equation. For our purposes, the following proof is simpler and more direct.

Proof. Equation (3) implies that $\boldsymbol{a}_t = z_t - \boldsymbol{H}\boldsymbol{x}_t - \boldsymbol{D}\boldsymbol{u}_t$. Substituting this expression in (2) yields:

$$\boldsymbol{x}_{t+1} = (\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{H})\boldsymbol{x}_t + (\boldsymbol{\Gamma} - \boldsymbol{E}\boldsymbol{D})\boldsymbol{u}_t + \boldsymbol{E}z_t \tag{20}$$

If the initial states $\boldsymbol{x}_1$ were known, then all the sequence of states would be exactly determined by (20) because the inputs in the right-hand-side of this expression, $\boldsymbol{u}_t$ and z_t, belong to the information set $\boldsymbol{\Omega}^N$. Usually $\boldsymbol{x}_1$ is unknown and treated as a random variable. Then the covariance of smoothed estimates, conditional on all the sample, would be:

$$P_{t|N} = (\mathbf{\Phi} - \boldsymbol{EH})^{t-1} P_{1|N} [(\mathbf{\Phi} - \boldsymbol{EH})^{t-1}]^T \tag{21}$$

where the eigenvalues of $\mathbf{\Phi} - \boldsymbol{EH}$ coincide with the reciprocal roots of the MA terms implied by (2)-(3). If the model is strictly invertible, then the eigenvalues of $\mathbf{\Phi} - \boldsymbol{EH}$ are inside the unit circle and $\boldsymbol{P}_{t|N}$ converges to zero at an exponential rate as t increases. On the other hand, if some eigenvalues of $\mathbf{\Phi} - \boldsymbol{EH}$ are in the unit circle, this convergence occurs as the eigenvalues of $\boldsymbol{P}_{1|N}$ associated with the noninvertible MA roots converge to zero, but it is not exponentially fast. $\square$

The convergence to zero implied by (21) can happen in three different ways.

1) When the model has a pure AR structure of p-th order, the trace becomes zero after processing exactly p observations. This property holds even when the process has unit AR roots.

2) If the model has all the MA roots inside the unit circle, then $\mathrm{tr}(\boldsymbol{P}_{t|N})$ converges exponentially to zero as $t \to N$. The decay of the coefficients in the corresponding autoregressive representation governs the rate of convergence. Therefore, a MA root close to the unit circle slows convergence and vice versa. Note that this type of convergence also holds for MA roots outside the unit circle, as any such model can be written in a strictly invertible form.

3) When the MA term has unit roots, convergence is asymptotic. In this case the trace of $\boldsymbol{P}_{t|N}$ converges to a small value for each sample size, and this value tends to zero as the sample increases. Then, convergence of $\mathrm{tr}(\boldsymbol{P}_{t|N})$ to zero is assured when $t \to \infty$.

The following corollaries provide further insight into the implications of Result 3.

Corollary 3.1. The Kalman filter covariance of an innovations model, $\boldsymbol{P}_{t|t-1}$, converges to zero.

Proof. The proof is identical to that of Result 3, replacing the subindex N in (21) by t-1. $\square$

Corollary 3.2. In an innovations model the smoothed and Kalman filter estimates of the states converge to the same values, so that $\|\boldsymbol{x}_{t|N} - \boldsymbol{x}_{t|t-1}\| \underset{t \to \infty}{\to} 0$, being $\|\ \|$ a vector norm.

Proof. Immediate from Result 3 and Corollary 3.1. $\square$

Therefore, one-sided Kalman filter estimates also converge to exact values (Corollary 3.1) and, consequently, to the smoothed estimates of the states (Corollary 3.2). For an on-line application the recursivity of Kalman filtering would be a clear computational advantage, specially if the recursion is simplified in the light of Corollary 3.1. On the other hand the use of filtering instead of smoothing implies some inefficiency, as $\boldsymbol{P}_{t|N}$ converges to zero faster than $\boldsymbol{P}_{t|t-1}$.

5. STRUCTURE OF THE METHOD.

Building on the ideas discussed in Sections 2, 3 and 4, the structural decomposition of a vector of time series, z_t, can be organized in the following steps:

Step 1) Obtain an adequate representation for z_t and the equivalent innovations model, using Result 1 or the simplified procedure described in the Appendix.

Step 2) According to Result 2, obtain the equivalent block-diagonal innovations model.

Step 3) Compute smoothed estimates of the states, $\bar{\boldsymbol{x}}_{t|N}$, and the corresponding covariances, $\boldsymbol{P}_{t|N}$, which by Result 3 converge to zero. For on-line applications, the Kalman filter estimates, $\bar{\boldsymbol{x}}_{t|t-1}$, may be an adequate alternative.

Step 4) Classify the different states according to their eigenvalues. Compute estimates of the trend, cycle and seasonal components by combining the estimates of the states with the corresponding coefficients in the observation equation, see (15)-(17).

Step 5) Compute the instantaneous effect of the exogenous variables as indicated in (18).

Step 6) Compute estimates of the irregular component as $\boldsymbol{a}_{t|N} = z_t - \bar{\boldsymbol{H}}\bar{\boldsymbol{x}}_{t|N} - \boldsymbol{D}\boldsymbol{u}_t$, see (11) and (19).

To illustrate the practical application of this methodology, Sections 6 and 7 present two examples.

The first example consists of a worst-case comparison of our method with standard procedures using simulated data. To this end, we simulate an stochastic process which can be interpreted both, as a structural model and an ARIMA-based decomposition. The components are then extracted using the direct state-space representation and the equivalent innovations model. In this exercise our method is in deliberate disadvantage, as the data generating process makes explicit certain variance proportionality constraints which get lost in the innovations representation. Despite this fact, the results obtained by both methodologies are very similar, but our components show clear advantages due to their convergence to exact values.

In the second example we illustrate the flexibility and multivariate capacities of our method in a multivariate framework. First, an empirical analysis concludes that the best representation for these series is a VARMAX model, with common trend and common pseudo-cycle constraints. Using our procedure the extraction and separate analysis of common and specific components is straightforward. Also, the presence of some missing values in the sample is easily treated.

6. AN EXAMPLE WITH SIMULATED DATA: COMPARING DIFFERENT METHODS FOR TIME SERIES DECOMPOSITION.

Consider the data generating process defined in Table 2. Following the terminology of Harvey (1989), the trend component follows a particular case of the *stochastic trend model*, implying that the trend in t+1 is equal to the trend in t plus a random walk derivative of the trend, Δ_t. The seasonal component follows a quarterly *dummy variable seasonality model*, where the sum of the seasonal components over a year is a random disturbance. The error terms η_t, ω_t and e_t are gaussian white noise processes, with an instantaneous covariance matrix:

$$COV\begin{bmatrix}\eta_t\\ \omega_t\\ e_t\end{bmatrix} = \begin{bmatrix}1/1600 & 0 & 0\\ 0 & .1 & 0\\ 0 & 0 & 1\end{bmatrix} \quad (22)$$

Note that, according to the hypotheses of most decomposition methods, the shocks affecting different components are drawn from independent distributions. Also, the variance-ratio, $\sigma^2_\eta/\sigma^2_e = 1/1600$, coincides with the value assumed in the HP filter for quarterly data (Hodrick and Prescott 1997).

[Insert Table 2]

The second column of Table 2 shows how the model can be interpreted as an ARIMA-model-based decomposition, where the trend follows an integrated random-walk and the seasonal component follows a nonstationary AR(3). Also, the observable time series follows an $IMA(1,5)\times(1,0)_4$ process, which can be approximately factored to the very common $IMA(1,1)\times(1,1)_4$ airline model.

Using Result 1, this model can be written in an equivalent innovations form, and then block-diagonalized using Result 2. The representation obtained is shown in Table 3.

[Insert Table 3]

Note that: a) the states are now clearly separated in different (1×1) and (2×2) blocks, and b) the eigenvalues of the transition matrix are identical to the inverse AR roots of the stochastic processes in Table 2. According to their frequencies, the first and second states correspond to the trend and the third to fifth states correspond to the seasonal component. Taking into account this classification and the coefficients in the last row of Table 3, the structural components are given by:

$$t_t = \bar{x}_{1t} \quad (23)$$

$$s_t = .619\bar{x}_{3t} - .342\bar{x}_{4t} - .577\bar{x}_{5t} \tag{24}$$

$$g_t = z_t - t_t - s_t \tag{25}$$

Combining the state-space model in Table 3 with (23)-(25), we can derive ARIMA representations for these components. Table 4 compares them with the corresponding models of the data generating process.

[Insert Table 4]

Note that both sets of models have the same autoregressive structures, but different error terms. While the components of the data generating process follow pure AR processes, the components implied by the innovations model receive shocks from the delayed irregular component, affected by different moving average polynomials and scale parameters.

We now generate 200 random draws from the data generating process (the data can be obtained at *www.ucm.es/info/icae/e4/*) and compute the smoothed estimates of the components using both, the direct formulation in Table 2 and the equivalent innovations model in Table 3. Figure 1 compares both sets of empirical components.

[Insert Figure 1]

Despite their overall similarity, the uncertainty affecting both decompositions is very different. Figure 2 compares the trace of the covariance matrices corresponding to the structural and the innovations model components.

[Insert Figure 2]

Note that the uncertainty of the components derived from the structural model display the U-shape characteristic of symmetric filters, meaning that the estimates at the center of the sample are more precise than those at the extremes. When the decomposition is applied to obtain seasonally adjusted data this fact is very important, because an increase in the sample generates a revision proportional to the uncertainty of the components. On the other hand, the components implied by the innovations model converge to exact values, as was stated in Result 3, so they will not be affected by new data.

Finally, Figure 3 shows the sample autocorrelation functions of the second-order differences of the smoothed trends.

[Insert Figure 3]

Note that the differenced structural model trend displays a nonstantionary second-order autoregressive structure, far from the white noise behavior that could be expected from the theoretical model, see Table 4. Therefore, the level of this variable has at least a fourth-order AR structure, with at least three roots on the unit circle. This fact suggests some mutual contamination between the trend and quarterly components. In comparison, the autocorrelations of the trend implied by the innovations model are coherent with its IMA(2,1) model, see Table 4.

In fact, contamination of the empirical structural model trend is to be expected because, despite the independence constraint in (22), the smoothed estimates of the components in this model are correlated. For example, the instantaneous conditional correlations between the trend and seasonal states vary between ±.20 at the beginning and at the end of the sample, being almost null at the middle. In comparison, the empirical components resulting from innovations model converge to independent values and, therefore, its properties correspond neatly with those of the theoretical models.

This result provides useful insight about the implications of the widespread assumption that the shocks affecting components are conditionally uncorrelated. Many authors recognize that this restriction is not justified on theoretical grounds, but find two compelling practical motivations to adopt it.

First, structural components are often analyzed separately and, therefore, some kind of independence among them is desirable. This example shows clearly that this constraint is both unnecessary and harmful. It is unnecessary because one can compute independent estimates of the components, even if the shocks affecting them are mutually correlated. It is harmful because it generates an undetectable statistical structure for the components, even when the model for the data is intrinsically detectable.

Second, the independence constraint is often used to estimate the models for the components. For example, in the ARIMA-model-based framework this is a standard assumption to solve uniquely the equations relating the model for the time series and the models for the components. On the other hand the literature about structural models does not discuss the value of this assumption but, according to our experience, maximum likelihood algorithms show symptoms of extreme ill-conditioning when nonzero covariances are allowed. Therefore, we guess that in structural modeling this assumption is employed, at least partially, to improve the behavior of estimation procedures. In comparison, our approach adopts the more natural assumption that there is a single source of shocks affecting all the components. Besides being closer to most time series models, this example shows clearly that it is a more effective symplificatory hypothesis.

7. AN EXAMPLE WITH REAL DATA: COMMON FEATURES IN WHEAT PRICES.

The degree of market integration in different periods is an attractive and debated issue to quantitative historians. Many of their works share two common features: a) they focus in grain markets, due to their importance in the pre-industrial economies, and b) are based in the statistical analysis of time series of grain prices measured in different locations. In this example we analyze the yearly average prices of the wheat sold in the *Monasterio de Sandoval* (S_t) and in the *Fábrica de la iglesia de Alaraz* (A_t), between 1691 and 1788 (N=98), in reales per fanega. The second series has six missing values due to discontinuities in the source. Both series can be obtained at *www.ucm.es/info/icae/e4/*. The goals of the analysis are: a) testing if these series show common statistical features (Engle and Kozicki 1993) and, if found, b) apply our method to estimate them.

After a standard univariate analysis, the following models were fitted to the log-transformed series and to the difference of (log) prices:

$$(1 - \underset{(.098)}{.049}B + \underset{(.098)}{.289}B^2)\nabla \log S_t = \hat{a}_t^S \; ; \; \hat{\sigma}_S^2 = .077, \; Q(10) = 5.55, \; p = 4.12, \; d = .54 \tag{26}$$

$$(1 - \underset{(.100)}{.004}B + \underset{(.100)}{.343}B^2)\nabla \log A_t = \hat{a}_t^A \; ; \; \hat{\sigma}_A^2 = .113, \; Q(10) = 12.25, \; p = 3.99, \; d = .59 \tag{27}$$

$$(1 - \underset{(.102)}{.280}B)(\log A_t - \log S_t) = \underset{(.031)}{.107} + \hat{a}_t \; ; \; \hat{\sigma}^2 = .067, \; Q(10) = 6.81 \tag{28}$$

where the figures in parentheses are the standard deviations of the estimates; $Q(10)$ is the Ljung-Box Q statistic, computed with the first 10 lags of the sample autocorrelation function; and the values p, d are, respectively, the period and damping factor of the pseudo-cycle implied by the corresponding AR(2) factor. Due to the presence of missing values, the estimates in (27) and (28) were obtained using state-space techniques as described in Terceiro (1990, chap. 5).

Then, single log prices are adequately represented by ARIMA(2,1,0) processes, implying that both series are nonstationary and have harmonic pseudo-cycles with a period of roughly four years. Also, the model for the difference of log prices is a stationary AR(1) with a constant term, meaning that both series share a common trend and a common harmonic cycle. The presence of these cofeatures supports the idea that both markets were substantially integrated in this historical period.

Further evidence can be found by building a multivariate model, constrained to represent these cofeatures. A bivariate analysis (Jenkins and Alavi 1981) provided the following model:

$$\begin{bmatrix} 1 - \underset{(.089)}{.231} B & 0 \\ 0 & (1 - \underset{(.086)}{.026} B + \underset{(.082)}{.247} B^2)(1 - B) \end{bmatrix} \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} \underset{(.026)}{.110} \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ \underset{(.090)}{.619} B & 1 \end{bmatrix} \begin{bmatrix} \hat{a}_{1t} \\ \hat{a}_{2t} \end{bmatrix} \tag{29}$$

$$\hat{\boldsymbol{\Sigma}}_a = \begin{bmatrix} .068 & -.003 \\ = & .052 \end{bmatrix} \quad ; \quad \boldsymbol{Q}(10) = \begin{bmatrix} 6.99 & 11.06 \\ 9.18 & 9.93 \end{bmatrix} \tag{30}$$

where $\boldsymbol{z}_t$ is defined as:

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \log A_t \\ \log S_t \end{bmatrix} \tag{31}$$

and $\boldsymbol{Q}(10)$ is a matrix of Ljung-Box statistics for the residual sample autocorrelations and cross-correlations. Table 5 summarizes the structure of model (29)-(30) in block-diagonal innovations form. According to the eigenvalues of the transition matrix, the first state corresponds to the trend, the second state corresponds to a cyclical movement and the third and fourth states correspond to a harmonic pseudo-cycle.

[Insert Table 5]

As the sample has missing values (at 1727-29, 1731, 1759 and 1782) the trace of the smoothed covariance matrices displays a special behavior. It drops quickly to zero but shows transitory peaks of uncertainty in the years following a missing value. Due to this effect, the trace is nonzero (with four decimals of precision) in 1728-33, 1760-61 and 1783-84.

Obtaining the structural decomposition of $\boldsymbol{z}_t$ is now trivial, taking into account the observation equation in Table 5. However, it would be more meaningful to compute the decomposition in terms of the (log) level of the series. By (31) the variables in $\boldsymbol{z}_t$ are related with the log-prices by:

$$\begin{bmatrix} \log A_t \\ \log S_t \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} \tag{32}$$

therefore, pre-multiplying the observation equation by:

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \tag{33}$$

yields the following decomposition of the log prices:

$$\begin{bmatrix} \log A_t \\ \log S_t \end{bmatrix} = \begin{bmatrix} .971 & 1 & 1.106 & -.391 \\ .971 & 0 & 1.106 & -.391 \end{bmatrix} \begin{bmatrix} \bar{x}_{1t} \\ \bar{x}_{2t} \\ \bar{x}_{3t} \\ \bar{x}_{4t} \end{bmatrix} + \begin{bmatrix} .110 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{a}_{1t} \\ \hat{a}_{2t} \end{bmatrix} \tag{34}$$

and the corresponding structural components:

$$t_t = .971\bar{x}_{1t} \tag{35}$$

$$c_t = 1.106\bar{x}_{3t} - .391\bar{x}_{4t} \tag{36}$$

$$c_t^A = \bar{x}_{2t} + .110 \tag{37}$$

$$\varepsilon_t^A = \hat{a}_{1t} + \hat{a}_{2t} \tag{38}$$

$$\varepsilon_t^S = \hat{a}_{2t} \tag{39}$$

According to (34), t_t is the common trend, c_t is the common cycle and c_t^A is a stationary component including all the individual features of the log price in Alaraz. A strict application of the decomposition defined in (12) would imply computing a separate component for the constant term, $d_t = .110$. Obviously this would be a meaningless complication, so we add the constant to the cyclical component in (37). Last, the white noise terms, ε_t^A and ε_t^S are, respectively, the irregular components affecting the series of Alaraz and Sandoval, respectively. The estimates of all these components are shown in Figures 4 and 5.

[Insert Figures 4 and 5]

Finally, according to (38)-(39), the covariance matrix of the irregular components is:

$$cov\begin{pmatrix} \varepsilon_t^A \\ \varepsilon_t^S \end{pmatrix} = \begin{bmatrix} .114 & .049 \\ .049 & .052 \end{bmatrix} \tag{40}$$

implying an instantaneous cross-correlation of .636. This high value further confirms that there is a substantial degree of integration between the local markets of Alaraz and Sandoval. The exercise could continue by looking for exogenous variables causing the cofeatures. For example, a climatological indicator could explain part of the common trend and/or the common cycle.

8. CONCLUDING REMARKS.

The structural decomposition proposed in this paper improves existing methods in five aspects.

First, it enforces consistency between the properties of a time series and those of the structural components, as model-based methods do, but avoids the *ad-hoc* independence constraints required by the ARIMA-based or structural modeling approaches. The example in Section 6 shows that this assumption is unnecessary and induces incoherence between the statistical properties of theoretical and empirical components.

Second, the method is independent of particular model specifications, as it is based on the innovations model which encompasses standard processes such as ARIMA, VARMAX or transfer functions, see the Appendix, and structural models or models with errors-in-variables, see Result 1. The example in Section 6 illustrates clearly this point, as the original data generating process is a structural model, with equivalent innovations and ARMA representations.

Third, the flexibility of the model and the state-space methods easily accommodates nonstandard situations defined by missing values, multivariate time series and constraints upon the structural components, such as cointegration or other cofeatures, see the example in Section 7.

Fourth, the signal extraction algorithm employed is a symmetric fixed-interval smoother. This is a substantial advantage over methods using one-sided filters (e.g., forecast decomposition procedures) as: a) all the information available is efficiently processed, so b) the variance of the components will converge fast to its zero steady-state value, and c) all the sequence of empirical components is based on the same information set. However, the modular and open structure of our method makes easy to replace the smoothing algorithm by the Kalman filter, which may be more adequate for on-line applications due to its recursive nature.

Last, the innovations model employed guarantees that the covariance matrix for the components converges to zero. This property assures coherence between the properties of the theoretical and empirical components, provides a rigorous statistical foundation for using the empirical components as observable and mutually independent time series and guarantees that these components will not change as the sample increases. Obviously, this important feature is in practice conditional on the model for the time series. Reestimation of the parameters or changes in the model formulation will affect the components.

In applications not reported here, we have found our decomposition to be useful in: testing for common dynamic components between seasonal time series, forecasting a time series with long-run constraints on its components and detecting misspecification of the initial model. About the latter use, note that a structural decomposition breaks down the time series into components with simpler properties, according to the dynamics implied by the model. Therefore, if a given component displays unexpected properties (e.g., spectral peaks at frequencies where there should be no information) a probable cause is a wrong specification of the model dynamics.

The procedure described in this article is implemented in the function *e4trend*. It is included in a MATLAB toolbox for time series modeling called E^4, which can be downloaded at *www.ucm.es/info/icae/e4*. This site also includes a complete user manual and other materials, which are freely distributed for teaching and research.

APPENDIX: THE INNOVATIONS REPRESENTATION OF A VARMAX MODEL.

Assume that z_t follows the VARMAX(p,s,q) process:

$$\boldsymbol{F}(B)\boldsymbol{z}_t = \boldsymbol{G}(B)\boldsymbol{u}_t + \boldsymbol{\Xi}(B)\boldsymbol{a}_t \tag{A.1}$$

where $\boldsymbol{a}_t$ is a $m \times 1$ vector of white noise errors, $\boldsymbol{u}_t$ is a $r \times 1$ vector of exogenous variables; the polynomial matrices $\boldsymbol{F}(B)$, $\boldsymbol{G}(B)$ and $\boldsymbol{\Xi}(B)$, are defined by:

$$\boldsymbol{F}(B) = \boldsymbol{I} + \sum_{i=1}^{p} \boldsymbol{F}_i B^i, \quad \boldsymbol{G}(B) = \sum_{i=0}^{s} \boldsymbol{G}_i B^i, \quad \boldsymbol{\Xi}(B) = \boldsymbol{I} + \sum_{i=1}^{q} \boldsymbol{\Xi}_i B^i \tag{A.2}$$

and may contain roots in the unit circle; finally, B denotes the backward-shift operator, such that for any sequence $\boldsymbol{x}_t$: $B^{\pm k}\boldsymbol{x}_t = \boldsymbol{x}_{t \mp k}$.

Under these conditions (Aoki 1990; Terceiro 1990) model (A.1)-(A.2) can be expressed in the equivalent innovations form defining:

$$\boldsymbol{\Phi} = \begin{bmatrix} -\boldsymbol{F}_1 & \boldsymbol{I} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ -\boldsymbol{F}_2 & \boldsymbol{0} & \boldsymbol{I} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\boldsymbol{F}_{k-1} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{I} \\ -\boldsymbol{F}_k & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{G}_1 - \boldsymbol{F}_1\boldsymbol{G}_0 \\ \boldsymbol{G}_2 - \boldsymbol{F}_2\boldsymbol{G}_0 \\ \vdots \\ \boldsymbol{G}_{k-1} - \boldsymbol{F}_{k-1}\boldsymbol{G}_0 \\ \boldsymbol{G}_k - \boldsymbol{F}_k\boldsymbol{G}_0 \end{bmatrix}, \quad \boldsymbol{E} = \begin{bmatrix} \boldsymbol{\Xi}_1 - \boldsymbol{F}_1 \\ \boldsymbol{\Xi}_2 - \boldsymbol{F}_2 \\ \vdots \\ \boldsymbol{\Xi}_{k-1} - \boldsymbol{F}_{k-1} \\ \boldsymbol{\Xi}_k - \boldsymbol{F}_k \end{bmatrix} \tag{A.3}$$

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{D} = \boldsymbol{G}_0 \tag{A.4}$$

where the *state vector* $\boldsymbol{x}_t$ has $n = m \cdot k$ rows, and $k = \max\{p, s, q\}$, and the error term in the innovations model coincides with $\boldsymbol{a}_t$ in (A.1).

REFERENCES.

Anderson, B.D.O., and Moore, J.B. (1979), *Optimal Filtering*, Englewood Cliffs, NJ: Prentice-Hall.

Aoki, M. (1990), *State Space Modeling of Time Series*, Berlin: Springer-Verlag.

Box, G.E.P., Hillmer, S.C., and Tiao, G.C. (1978), "Analysis and Modeling of Seasonal Time Series," in *Seasonal Analysis of Time Series*, ed. A. Zellner, Washington DC: Bureau of the Census.

Box, G.E.P., Pierce, D.A., and Newbold, P. (1987), "Estimating Trend and Growth Rates Analysis in Seasonal Time Series," *Journal of the American Statistical Association*, 82, 397, 276-282.

Burman, J.P. (1980), "Seasonal Adjustment by Signal Extraction," *Journal of the Royal Statistical Society*, series A, 143, 321-337.

Casals, J., Sotoca, S. and Jerez, M. (1999), "A Fast and Stable Method to Compute the Likelihood of Time Invariant State-Space Models," *Economics Letters*, 65, 329-337.

Casals, J., Jerez, M., and Sotoca, S. (2000), "Exact Smoothing for Stationary and Nonstationary Time Series," *International Journal of Forecasting*, 16, 1, 59-69.

De Souza, C.E., Gevers, M.R., and Goodwin, G.C. (1986), "Riccati Equations in Optimal Filtering of Nonstabilizable Systems having Singular State Transition Matrices," *IEEE Transactions on Automatic Control*, Vol. AC-31, 9, 831-838.

Engle, R.F., and Kozicki, S. (1993), "Testing for Common Features," *Journal of Business and Economic Statistics*, 11, 369-380.

Espasa, A., and Peña, D. (1995), "The Decomposition of Forecast in Seasonal ARIMA Models," *Journal of Forecasting*, 14, 7, 565-583.

Findley, D.B., Monsell, B.C., Bell, W.R., Otto, M.C., and Chen, B. (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program," *Journal of Business and Economic Statistics*, 16, 127-177.

Golub, G.H., and Van Loan, C.F. (1996), *Matrix Computations*. Baltimore: John Hopkins University Press.

Gómez, V., and Maravall, A. (1996), "Programs TRAMO and SEATS: Instructions for the User," Working paper 9628, Madrid: Bank of Spain (*http://www.bde.es/servicio/software/software.htm*).

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

Hillmer, S.C., and Tiao, G.C. (1982), "An ARIMA-Model-Based Approach to Seasonal Adjustment," *Journal of the American Statistical Association*, 77, 63-70.

Hodrick, R.J., and Prescott, E.C. (1997), "Post-war U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29, 1-16.

Ionescu, V., Oara, C., and Weiss, M. (1997), "General Matrix Pencil Techniques for the Solution of Algebraic Riccati Equations: A Unified Approach." *IEEE Transactions on Automatic Control*, 42, 8, 1085-1097.

Jenkins, G.M., and Alavi, A.S. (1981), "Some Aspects of Modelling and Forecasting Multivariate Time Series," *Journal of Time Series Analysis*, 2, 1, 1-47.

Petkov, P. Hr., Christov, N.D., and Konstantinov, M.M. (1991), *Computational Methods for Linear Control Systems*, Englewood Cliffs, NJ: Prentice-Hall.

Pole, A., West, M ., and Harrison, J. (1994), *Applied Bayesian Forecasting and Time Series Analysis*, London: Chapman & Hall.

Shiskin, J., Young, A.H., and Musgrave, J.C. (1967), "The X-11 Variant of the Census Method II Seasonal Adjustment Program," *Technical Paper*, Washington DC: Bureau of the Census.

Terceiro, J. (1990), *Estimation of Dynamic Econometric Models with Errors in Variables*. Berlin: Springer-Verlag.

West, M. (1997), "Time Series Decomposition and Analysis in a Study of Oxygen Isotope Records," *Biometrika*, 84, 489-494.

Young, P., Pedregal, D.J., and Tych, W. (1999), "Dynamic Harmonic Regression," *Journal of Forecasting*, 18, 369-394.

Table 1. Correspondence between the state variables and the structural components

<table>
<tr><td colspan="2">eig($\bar{\boldsymbol{\Phi}}$)</td><td>Spectral peaks at</td><td>Structural component</td></tr>
<tr><td rowspan="4">Real</td><td>$\lambda_j = 1$</td><td>$f_j = 0$</td><td>Trend</td></tr>
<tr><td rowspan="3">$\lambda_j = a_j,\ -1 \le a_j < 1$</td><td>$f_j = 0$ if $0 < a_j < 1$</td><td>Cycle</td></tr>
<tr><td colspan="2">If $a_j = 0$, then x_j is either a redundant state, with no effect on the structural components, or a lagged uncorrelated error which does not generate spectral peaks at any frequency. In this case, it should be included in the Cycle</td></tr>
<tr><td>$f_j = 1/2$ if $-1 \le a_j < 0$</td><td>Seasonal if $1/2 \in \boldsymbol{F_S}$, Cycle otherwise</td></tr>
<tr><td>Complex</td><td>$\lambda_{j,k} = a_j \pm b_j i$</td><td>$f_{j,k} = (2\pi)^{-1} \arctan(b_j / a_j)$</td><td>Seasonal if $f_{j,k} \in \boldsymbol{F_S}$, Cycle otherwise</td></tr>
</table>

Table 2: Definition of the data generating process

Component	Structural model representation	ARIMA representation
Trend	$T_{t+1} = T_t + \Delta_t$ $\Delta_{t+1} = \Delta_t + \eta_t$	$(1 - B)^2 T_{t+1} = \eta_t$
Seasonal	$(1 + B + B^2 + B^3) S_{t+1} = \omega_t$	$(1 + B + B^2 + B^3) S_{t+1} = \omega_t$
Time series	$z_t = T_t + S_t + e_t$	$(1 - B)(1 - B^4) z_t$ $= (1 - .933B + .091B^2 - .047B^3 - .585B^4 + .548B^5) a_t$ $\approx (1 - .933B)(1 - .585B^4) a_t \;;\; a_t \sim \text{nid}(0, 1.824)$

Table 3. Structure of the data generating process in block-diagonal innovations form

Outputs	Inputs						eig($\bar{\mathbf{\Phi}}$)	Freq. (f_j)	Component
	$\bar{x}_{1t}$	$\bar{x}_{2t}$	$\bar{x}_{3t}$	$\bar{x}_{4t}$	$\bar{x}_{5t}$	a_t			
			$\bar{\mathbf{\Phi}}$			$\bar{\boldsymbol{E}}$			
$\bar{x}_{1\,t+1}$	1.000	1.000	0	0	0	.188	1	0	Trend
$\bar{x}_{2\,t+1}$	.000	1.000	0	0	0	.019	1	0	
$\bar{x}_{3\,t+1}$	0	0	.489	1.461	0	-.070	$\pm i$	.25	Seasonal
$\bar{x}_{4\,t+1}$	0	0	-.848	-.489	0	-.116			
$\bar{x}_{5\,t+1}$	0	0	0	0	-1	.203	-1	.5	
			$\bar{\boldsymbol{H}}$			—			
z_t	1.000	.000	.619	-.342	-.577	1			

Table 4: Comparison between the models for the structural components in ARIMA notation

Component	Data generating process	Innovations model
Trend	$(1 - B)^2 T_t = \eta_{t-1}$ $\eta_{t-1} \sim \text{nid}(0\,,\,1/1600)$	$(1 - B)^2 t_t = (1 - .899B).188\gamma_{t-1}$
Seasonal	$(1 + B + B^2 + B^3) S_t = \omega_{t-1}$ $\omega_{t-1} \sim \text{nid}(0\,,\,.1)$	$(1 + B + B^2 + B^3) s_t$ $= (1 + 1.402B + 2.347B^2)(-.120)\gamma_{t-1}$
Irregular	$e_t \sim \text{nid}(0\,,\,1)$	$\gamma_t \sim \text{nid}(0\,,\,1.824)$

Table 5. Structure of model (29) in block-diagonal innovations form

Outputs	Inputs							eig($\bar{\mathbf{\Phi}}$)	Component
	$\bar{x}_{1t}$	$\bar{x}_{2t}$	$\bar{x}_{3t}$	$\bar{x}_{4t}$	$u_t = 1$	$\hat{a}_{1t}$	$\hat{a}_{2t}$		
	$\bar{\mathbf{\Phi}}$				$\bar{\mathbf{\Gamma}}$	$\bar{\boldsymbol{E}}$			
$\bar{x}_{1t+1}$	1	0	0	0	0	.522	.844	1	Trend
$\bar{x}_{2t+1}$	0	.231	0	0	.026	.231	0	.231	Cycle
$\bar{x}_{3t+1}$	0	0	.333	-.866	0	.052	.223	.013±.496i	Cycle (harmonic)
$\bar{x}_{4t+1}$	0	0	.403	-.307	0	-.140	.103		
	$\bar{\boldsymbol{H}}$				$\boldsymbol{D}$	—			
z_{1t}	0	1	0	0	.110	1	0		
z_{2t}	.971	0	1.106	-.391	0	0	1		

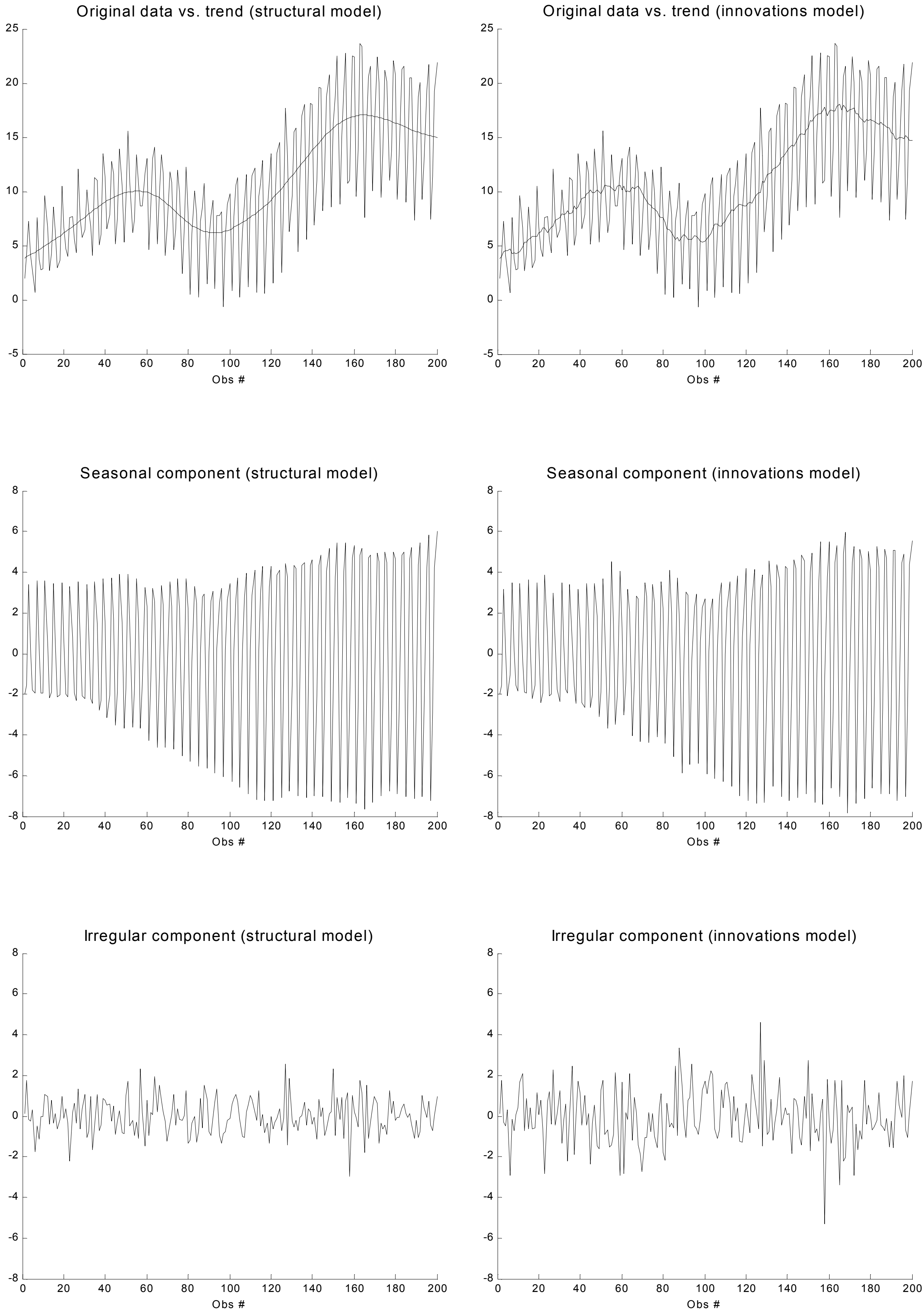


Figure 1. Comparison between the components obtained from the structural model and the innovations model. Note that: a) the trend and seasonal components obtained in both cases are very similar; b) the estimate of the trend resulting from the innovations model is more irregular because the smoothness of the structural model trend results from the ratios between the variances in (22), which are explicit in the data generating process but not in the innovations model; and c) the irregular components have different volatilities as expected, see Table 4.

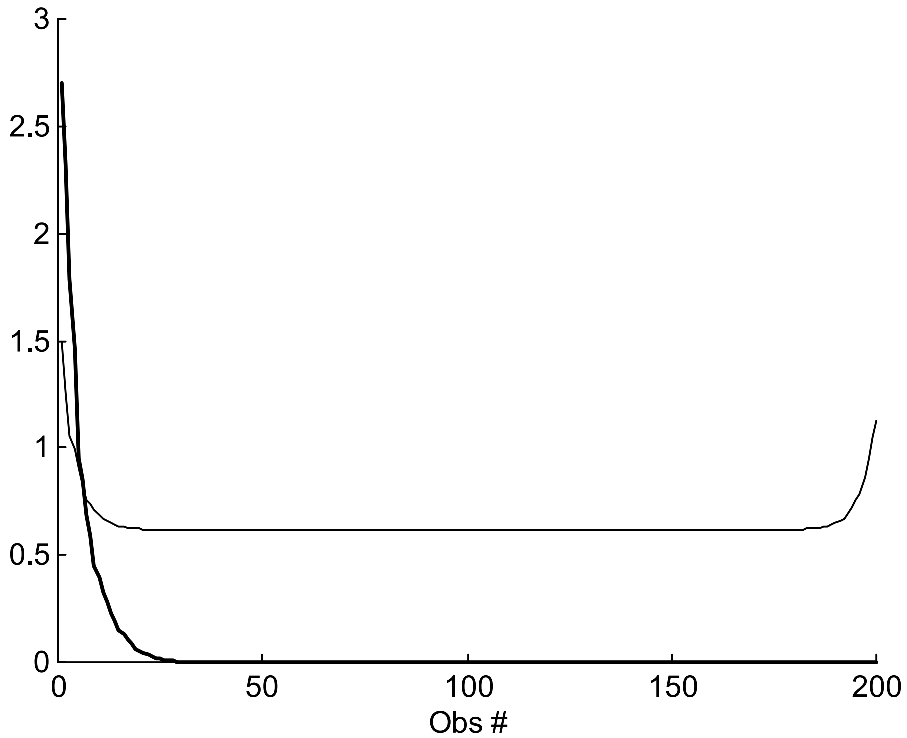


Figure 2. Trace of the smoother covariance matrices corresponding to the structural model (thin line) and the innovations model (thick line).

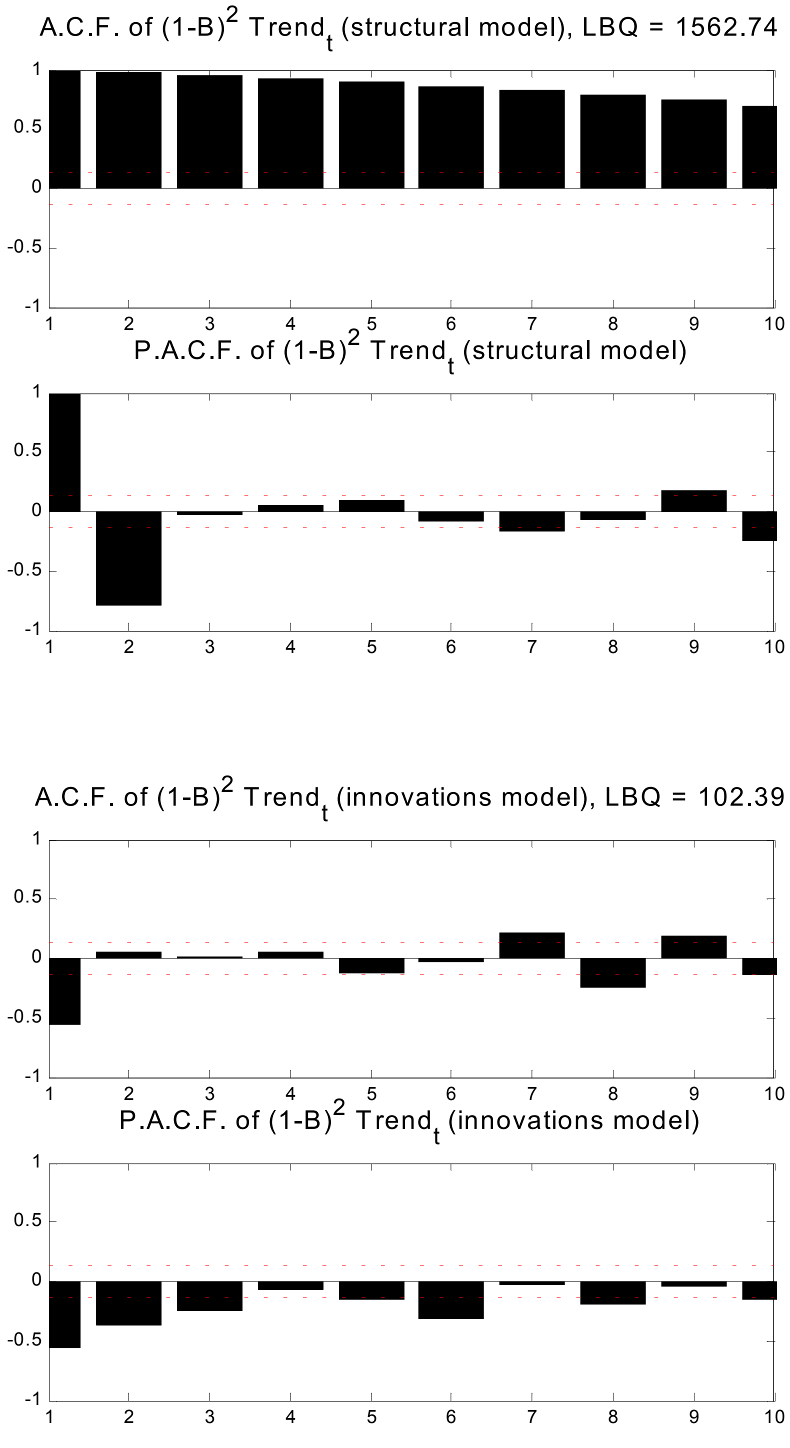


Figure 3. Autocorrelation analysis of the second-order differences of the trends implied by the structural model and the innovations model. The value "LBQ" appearing at the header of each plot is the Ljung-Box Q statistic computed with the first 10 lags of the sample autocorrelation function.

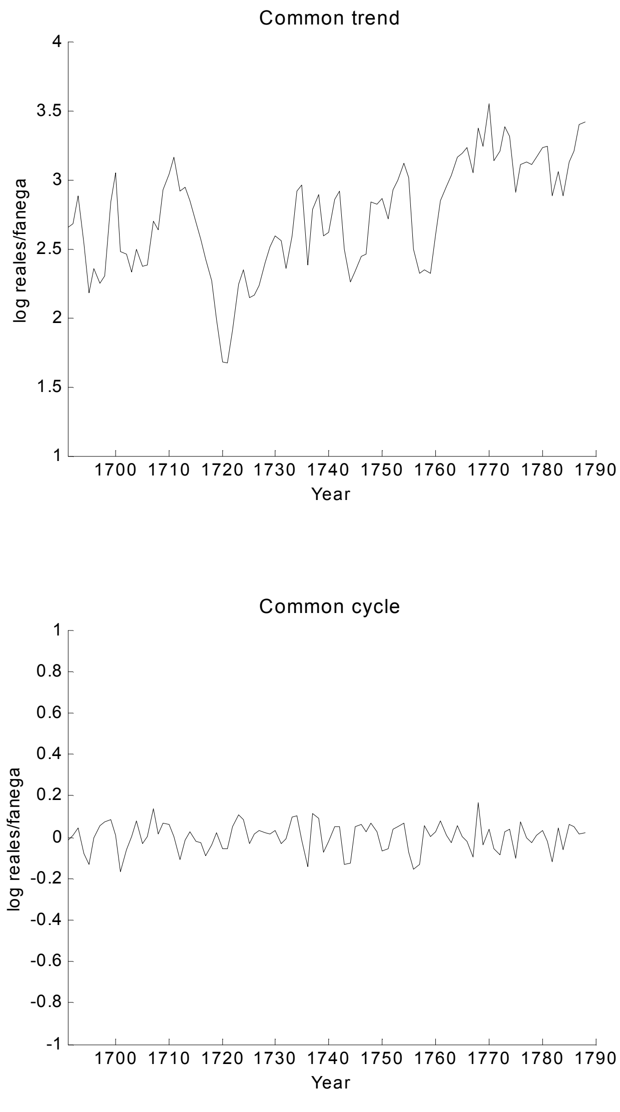
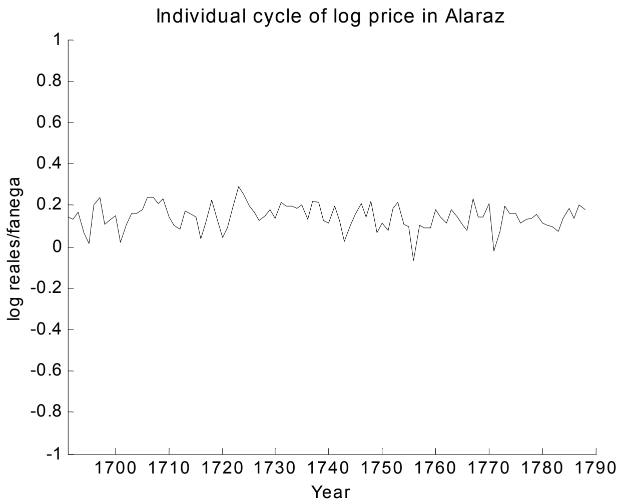
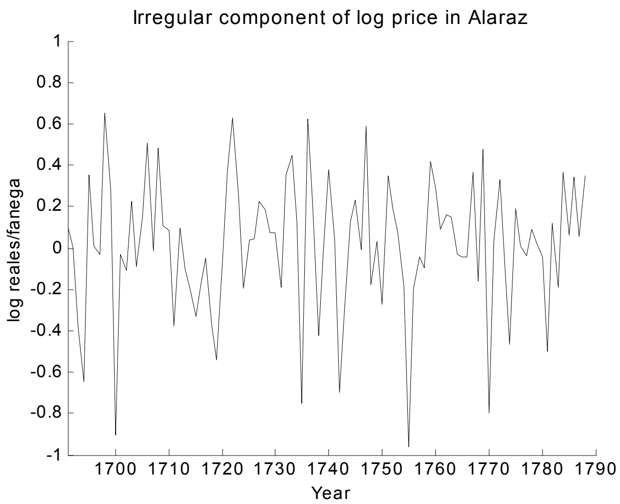


Figure 4. Common features of wheat prices (in logs).





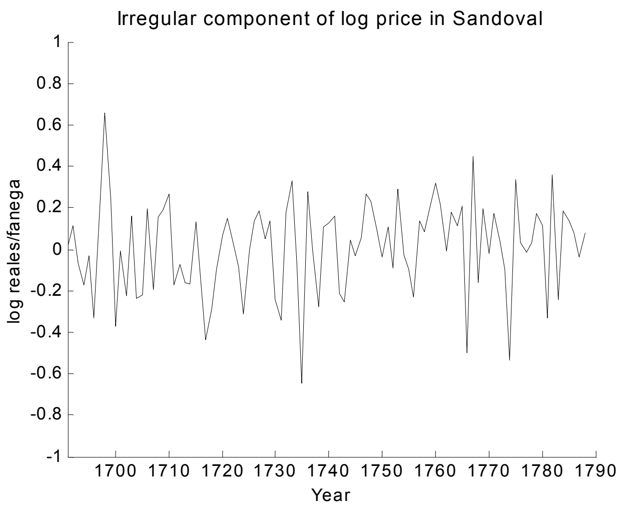


Figure 5. Individual features of (log) wheat prices.