Modelling and Forecasting Time Series Sampled at Different Frequencies

José Casals (jcasalsc@cajamadrid.es)
Miguel Jerez (mjerez@ccee.ucm.es) †
Sonia Sotoca (sotoca@ccee.ucm.es)

Departamento de Fundamentos del Analisis Economico II
Universidad Complutense de Madrid
Campus de Somosaguas S/N
28223 Madrid (SPAIN)

Abstract: This paper discusses how to specify an observable high-frequency model for a vector of time series sampled at high and low frequencies. To this end we first study how aggregation over time affects both, the dynamic components of a time series and their observability, in a multivariate linear framework. We find that the basic dynamic components remain unchanged but some of them, mainly those related to the seasonal structure, become unobservable. Building on these results, we propose a structured specification method built on the idea that the models relating the variables in high and low sampling frequencies should be mutually consistent. After specifying a consistent and observable high-frequency model, standard state-space techniques provide an adequate framework for estimation, diagnostic checking, data interpolation and forecasting. Our method has three main uses. First, it is useful to disaggregate a vector of low-frequency time series into high-frequency estimates coherent with both, the sample information and its statistical properties. Second, it may improve forecasting of the low-frequency variables, as the forecasts conditional to high-frequency indicators have in general smaller error variances than those derived from the corresponding low-frequency values. Third, the resulting forecasts can be updated as new high-frequency values become available, thus providing an effective tool to assess the effect of new information over medium term expectations. An example using national accounting data illustrates the practical application of this method.

Keywords: State-space models, Kalman filter, temporal disaggregation, observability, seasonality

---

† Corresponding author. Phone numbers: (+34) 91 394 23 61 (voice) and (+34) 91 394 25 91 (fax).

1. Introduction

This paper discusses how to specify an observable high-frequency model for a vector of time series sampled at high and low frequencies. To avoid cumbersome wordings, we will refer to the low-frequency variables as “annual” and to the high-frequency variables as “quarterly”. The results however are valid for any combination of sampling frequencies.

The need to build such a model arises in two main situations. First, an organization samples quarterly and annual data for several variables. To make a clear presentation to statistically unsophisticated audiences, it wants to estimate the unobserved quarterly values using all available (aggregated and disaggregated) information. Second, an important time series is measured once per year, but indicators about its performance are sampled quarterly. To assess the evolution of the target variable, one wants to compute a quarterly indicator of its fluctuations and forecast its annual value, exploiting in both cases the information provided by the quarterly indicators. Two examples of real world organizations that use these techniques on a regular basis are statistical agencies, which typically need to disaggregate, monitor and forecast important macroeconomic variables, and managers of hydraulic resources, who need to infer high-frequency rainfall data from low-frequency records, see e.g., Onof *et al.* (2000).

There are several ways to model data sampled at different frequencies. For the purposes of this paper we will concentrate in the model-based approach. Each method in this family is characterized by a given high-frequency model, which implies a uniquely determined system of equations relating the sample data with the unknown quarterly figures. Typically, this system is then solved by an extended least-squares procedure. Important examples of this approach are those of Denton (1971), Chow-Lin (1971), Fernández (1981), Litterman (1983) or Santos-Silva & Cardoso (2001). Table 1 summarizes most of these models as particular cases of the regression model:

$$y_t = \boldsymbol{x}_t^T \boldsymbol{\beta} + \varepsilon_t \, ; \;\; \varepsilon_t = \frac{1}{(1-\varphi_1 B)(1-\varphi_2 B)} a_t \qquad (1.1)$$

where y_t denotes the target variable in quarter *t*, $\boldsymbol{x}_t$ is a vector of quarterly indicators, $a_t \sim iid(0, \sigma_a^2)$, and *B* denotes the backshift operator, such that for any sequence w_t: $B^k w_t = w_{t-k}, k = 0, \pm 1, \pm 2, \ldots$

[Insert Table 1]

Note that the models in Table 1 assume different orders of integration for the variable and include no seasonal factors so, either seasonality has been removed beforehand, or it is a common feature between y_t and $\boldsymbol{x}_t$, such that the linear combination $y_t - \boldsymbol{x}_t^T \boldsymbol{\beta}$ has no seasonal structure. When looking at these widely different structures it is natural to ask: how would a specification error affect the disaggregates and forecasts resulting from these models? As time series interpolation and forecasting are particular cases of the same basic inference problem, a natural answer to this question arises by analogy: disaggregates computed using model (1.1) instead of the true data generating process will have in general the same flaws as the forecasts computed with (1.1) in comparison with optimal forecasts. On the other hand, predictive accuracy is critical for time series disaggregation, as large forecast errors generate important revisions of the disaggregates.

In this paper we implement a state-space (SS) approach to model, interpolate and forecast a vector of time series observed at different frequencies. Our starting point consists of breaking up the global problem in four basic questions: How could one specify a quarterly model on the basis of a mixture of quarterly and annual data? How could such model be estimated? How to compute within-the-sample estimates of the unobserved quarterly values? How to forecast the annual variables exploiting the quarterly information available? All these issues, except the first, have been effectively addressed by the SS literature, see e.g., Ansley & Kohn (1983), Harvey & Pierse (1984) and Terceiro (1990, chapters 2 and 5). Due to this wealth of powerful and ready to use tools, other authors, such as Durbin and Quenneville (1997), Nunes (2005) or Proietti (2006), adopted the SS approach to implement different disaggregation proposals.

Therefore we will concentrate in the specification problem. Our approach builds on two basic ideas. First, quarterly and annual models should be mutually consistent, given the aggregation constraint. Second, a statistically adequate annual model, built by standard techniques, provides clear clues about the specification of a quarterly model. To develop the latter idea, Section 2 analyzes the effect of aggregation on the dynamics of a linear system. We find that the system dynamics are not altered by aggregation, but some components may become unobservable. Section 3 characterizes which components loose observability after aggregation and, combining this result with those in Section 2, defines an algorithm to derive the annual model corresponding to a general quarterly representation. Building on previous results about quarterly model aggregation, Section 4 discusses how to specify an observable model for the quarterly values, building on a previously fitted annual model. This discussion results in a structured model-building procedure, which application is illustrated in Section 5 using macroeconomic data. This example also shows that the resulting high-frequency model may provide better forecasts than those resulting from mainstream methods. Section 6 provides some concluding remarks and indicates how to obtain, via Internet, a MATLAB toolbox for time series modelling, which implements all the computational procedures required. The proofs of formal results are included in the appendices.

2. The effect of aggregation on the dynamics and forecasting accuracy of a time series model

Let z_t be an *m*x1 random vector of quarterly values. Without loss of generality (Casals, Sotoca & Jerez, 1999, Theorem 1) we will assume that these values are the observable output of a *steady-state innovations* SS model (hereafter, innovations model):

$$\boldsymbol{x}_{t+1} = \boldsymbol{\Phi}\, \boldsymbol{x}_t + \boldsymbol{\Gamma} \boldsymbol{u}_t + \mathbf{E}\, \boldsymbol{a}_t \tag{2.1}$$

$$\boldsymbol{z}_t = \boldsymbol{H}\boldsymbol{x}_t + \boldsymbol{D}\boldsymbol{u}_t + \boldsymbol{a}_t \tag{2.2}$$

where:

- $\boldsymbol{x}_t$ is a *n*x1 vector of *state variables* or *dynamic components*,
- $\boldsymbol{u}_t$ is a *r*x1 vector of exogenous indicators,
- $\boldsymbol{a}_t$ is a *m*x1 vector of errors, such that $\boldsymbol{a}_t \sim iid\,(\mathbf{0}, \boldsymbol{Q})$.

and the terms $\boldsymbol{\Phi}$, $\boldsymbol{\Gamma}$, $\mathbf{E}$, $\boldsymbol{H}$ and $\boldsymbol{D}$ are real-valued matrices of dimensions *n*x*n*, *n*x*r*, *n*x*m*, *m*x*n* and *m*x*r*, respectively.

We will also assume that model (2.1)-(2.2) is *minimal*. This is a non-restrictive hypothesis meaning that *n* is the smallest number of states required to describe the system dynamics.

2.1. The quarterly model in stacked form

It is difficult to discuss aggregation using model (2.1)-(2.2). To this end, it is more convenient the following "stacked" representation. Let *S* be the seasonal frequency, defined as the number of high-frequency sampling periods (quarters) yielding a single low-frequency (annual) observation. Consider the stacked signal, indicator and error vectors:

$$\boldsymbol{z}_{t:t+S-1} = \begin{bmatrix} \boldsymbol{z}_t \\ \boldsymbol{z}_{t+1} \\ \vdots \\ \boldsymbol{z}_{t+S-1} \end{bmatrix};\ \boldsymbol{u}_{t:t+S-1} = \begin{bmatrix} \boldsymbol{u}_t \\ \boldsymbol{u}_{t+1} \\ \vdots \\ \boldsymbol{u}_{t+S-1} \end{bmatrix};\ \boldsymbol{a}_{t:t+S-1} = \begin{bmatrix} \boldsymbol{a}_t \\ \boldsymbol{a}_{t+1} \\ \vdots \\ \boldsymbol{a}_{t+S-1} \end{bmatrix} \tag{2.3}$$

Without loss of generality we will assume that the aggregation period coincides with the seasonal frequency, so the stacked vectors in (2.3) include all the values subject to aggregation. Under these conditions, the following Proposition holds:

Proposition 1. Model (2.1)-(2.2) can be written equivalently as:

$$\boldsymbol{x}_{T+1} = \overline{\boldsymbol{\Phi}}\boldsymbol{x}_T + \overline{\boldsymbol{\Gamma}}\boldsymbol{u}_{t:t+S-1} + \overline{\mathbf{E}}\boldsymbol{a}_{t:t+S-1} \tag{2.4}$$

$$\boldsymbol{z}_{t:t+S-1} = \overline{\boldsymbol{H}}\boldsymbol{x}_T + \overline{\boldsymbol{D}}\boldsymbol{u}_{t:t+S-1} + \overline{\boldsymbol{C}}\boldsymbol{a}_{t:t+S-1} \tag{2.5}$$

where the index T refers to the aggregated (annual) time unit, such as: $\boldsymbol{x}_{T+1} = \boldsymbol{x}_{t+S}$, $\boldsymbol{x}_T = \boldsymbol{x}_t$, $\boldsymbol{a}_{t:t+S-1} \sim iid(\mathbf{0}, \overline{\boldsymbol{Q}})$ and the matrices in (2.4)-(2.5) are related to those in (2.1)-(2.2) by:

$$\overline{\boldsymbol{\Phi}} = \boldsymbol{\Phi}^S \; ; \; \overline{\boldsymbol{\Gamma}} = \left[\boldsymbol{\Phi}^{S-1}\boldsymbol{\Gamma}, \; \boldsymbol{\Phi}^{S-2}\boldsymbol{\Gamma}, \; \dots, \; \boldsymbol{\Gamma}\right]; \; \overline{\boldsymbol{E}} = \left[\boldsymbol{\Phi}^{S-1}\boldsymbol{E}, \; \boldsymbol{\Phi}^{S-2}\boldsymbol{E}, \; \dots, \; \boldsymbol{E}\right] \tag{2.6}$$

$$\overline{\boldsymbol{H}} = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{S-1} \end{bmatrix}; \; \overline{\boldsymbol{D}} = \begin{bmatrix} \boldsymbol{D} & \mathbf{0} & \dots & \mathbf{0} \\ \boldsymbol{H}\Gamma & \boldsymbol{D} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H\Phi}^{S-2}\boldsymbol{\Gamma} & \boldsymbol{H\Phi}^{S-3}\boldsymbol{\Gamma} & \dots & \boldsymbol{D} \end{bmatrix};$$

$$\overline{\boldsymbol{C}} = \begin{bmatrix} \boldsymbol{I} & \mathbf{0} & \dots & \mathbf{0} \\ \boldsymbol{HE} & \boldsymbol{I} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H\Phi}^{S-2}\boldsymbol{E} & \boldsymbol{H\Phi}^{S-3}\boldsymbol{E} & \dots & \boldsymbol{I} \end{bmatrix}; \; \overline{\boldsymbol{Q}} = \boldsymbol{I} \otimes \boldsymbol{Q} \tag{2.7}$$

Proof. Equation (2.5) is obtained by successively substituting (2.1) in (2.2) and writing the resulting system in matrix form. On the other hand (2.4) is immediately obtained by propagating (2.1) from *t* to *t*+*S*. ■

2.2. Aggregation relationships

Assume now the aggregation relationships:

$$\boldsymbol{z}_T^A = \boldsymbol{J}^A \boldsymbol{z}_{t:t+S-1} \tag{2.8}$$

$$\boldsymbol{z}_T^P = \boldsymbol{J}^P \boldsymbol{z}_{t:t+S-1} \tag{2.9}$$

where $\boldsymbol{z}_T^A$ and $\boldsymbol{z}_T^P$ denote, respectively, the vectors of annual and partially aggregated data observed in year T, including the quarterly values at t, t+1, …, t+S-1. By "partially aggregated data" we mean a sample combining all the observed annual and quarterly values. Note also that there always exists a relationship between the annual and partially aggregated series, see Lütkepohl (1987, Chapter 6):

$$\boldsymbol{z}_T^A = \boldsymbol{J}^* \boldsymbol{z}_T^P \tag{2.10}$$

where (2.8)-(2.10) imply: $\boldsymbol{J}^* \boldsymbol{J}^P = \boldsymbol{J}^A$.

The structure of the aggregation matrices in (2.8), (2.9) and (2.10) depends on the aggregation constraints and the nature of the variables. Following Di Fonzo (1990), there are four basic types of quarterly variables: flows, indices, beginning-of-period stocks or end-of-period stocks. Accordingly, the annual samples will be sums of quarterly values, averages or discrete beginning-of-period/end-of-period values. The following examples illustrate some common aggregation structures:

Example 2.1: If all the variables considered are flows then $\boldsymbol{J}^A = [\boldsymbol{I}, \boldsymbol{I}, \ldots, \boldsymbol{I}]$, meaning that each annual figure is the sum of the corresponding quarterly values. On the other hand, $\boldsymbol{J}^A = [\boldsymbol{0}, \boldsymbol{0}, \ldots, \boldsymbol{I}]$ implies that the variables are end-of-period stocks, such that $\boldsymbol{z}_{t+S-1}$ is observed while the corresponding values at t, t+1, …, t+S-2 are not.

Example 2.2: Assume that the first m_1 variables in $\boldsymbol{z}_{t:t+S-1}$, see (2.3), are annual flows, while the remaining m_2 variables (m_1+ m_2=m) are quarterly values. In this case the partial aggregation matrix would be:

$$\underset{\left[(m_1 + S \cdot m_2) \times (S \cdot m_1 + S \cdot m_2)\right]}{\boldsymbol{J}^P} = \begin{bmatrix} \boldsymbol{I}_{m_1} & \mathbf{0} & \boldsymbol{I}_{m_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{m_2} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_{m_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{I}_{m_2} \end{bmatrix} \tag{2.11}$$

where $\boldsymbol{I}_m$ is the $m \times m$ identity and the vectors of quarterly and partially observed values have the following structures:

$$\underset{[(S\cdot m_1+S\cdot m_2)\times 1]}{\boldsymbol{z}_{t:t+S\text{-}1}} = \begin{bmatrix} z_t^{m_1} \\ z_t^{m_2} \\ z_{t+1}^{m_1} \\ z_{t+1}^{m_2} \\ \vdots \\ z_{t+S\text{-}1}^{m_1} \\ z_{t+S\text{-}1}^{m_2} \end{bmatrix} ; \quad \underset{[(m_1+S\cdot m_2)\times 1]}{\boldsymbol{z}_T^P} = \begin{bmatrix} z_t^{m_1}+z_{t+1}^{m_1}+\ldots+z_{t+S\text{-}1}^{m_1} \\ z_t^{m_2} \\ z_{t+1}^{m_2} \\ \vdots \\ z_{t+S\text{-}1}^{m_2} \end{bmatrix} \tag{2.12}$$

Example 2.3: Assuming again that all the variables are flows, the aggregation matrix $\boldsymbol{J}^*$ in (2.10), transforming partially aggregated values into annual values, would be:

$$\underset{[(m_1+m_2)\times(m_1+S\cdot m_2)]}{\boldsymbol{J}^*} = \begin{bmatrix} \boldsymbol{I}_{m_1} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{m_2} & \boldsymbol{I}_{m_2} & \cdots & \boldsymbol{I}_{m_2} \end{bmatrix} \tag{2.13}$$

2.3. Relationship between the quarterly, annual and partially aggregated models

Under the previously stated conditions, the following result formalizes the relationship between the quarterly data model in the form (2.4)-(2.5) and the corresponding models for the annual and partially observed data:

Proposition 2. The models for $\boldsymbol{z}_T^A$ and $\boldsymbol{z}_T^P$ have the state equation (2.4) with the observation equations:

$$\boldsymbol{z}_T^A = \boldsymbol{H}^A\boldsymbol{x}_T + \boldsymbol{D}^A\boldsymbol{u}_{t:t+S\text{-}1} + \boldsymbol{C}^A\boldsymbol{a}_{t:t+S\text{-}1} \tag{2.14}$$

$$\boldsymbol{z}_T^P = \boldsymbol{H}^P\boldsymbol{x}_T + \boldsymbol{D}^P\boldsymbol{u}_{t:t+S\text{-}1} + \boldsymbol{C}^P\boldsymbol{a}_{t:t+S\text{-}1} \tag{2.15}$$

with:

$$\boldsymbol{H}^A = \boldsymbol{J}^A\overline{\boldsymbol{H}} \ ; \ \boldsymbol{D}^A = \boldsymbol{J}^A\overline{\boldsymbol{D}} \ ; \ \boldsymbol{C}^A = \boldsymbol{J}^A\overline{\boldsymbol{C}} \tag{2.16}$$

$$\boldsymbol{H}^P = \boldsymbol{J}^P\overline{\boldsymbol{H}} \ ; \ \boldsymbol{D}^P = \boldsymbol{J}^P\overline{\boldsymbol{D}} \ ; \ \boldsymbol{C}^P = \boldsymbol{J}^P\overline{\boldsymbol{C}} \tag{2.17}$$

Proof. Trivial, as pre-multiplying (2.5) by $\boldsymbol{J}^A$ and $\boldsymbol{J}^P$ immediately yields (2.14) and (2.15) respectively. ■

As a Corollary to Proposition 2, note that the observer of $\boldsymbol{z}_T^A$ can be alternatively obtained by aggregation of $\boldsymbol{z}_T^P$. Pre-multiplying (2.15) by $\boldsymbol{J}^*$ yields:

$$\boldsymbol{z}_T^A = \boldsymbol{J}^*\boldsymbol{z}_T^P = \boldsymbol{J}^*\boldsymbol{H}^P\boldsymbol{x}_T + \boldsymbol{J}^*\boldsymbol{D}^P\boldsymbol{u}_{t:t+S\text{-}1} + \boldsymbol{J}^*\boldsymbol{C}^P\boldsymbol{a}_{t:t+S\text{-}1} \tag{2.18}$$

2.4. The effect of aggregation on predictive accuracy

The following Proposition states that a model efficiently combining annual and quarterly data may provide better forecasts for the annual values than a model using only annual data.

Proposition 3. The forecast error variance of the annual model, given by (2.4) and (2.14), is greater or equal than the error variance of the annual forecasts computed by:

Case (1) aggregation of the forecasts for z_t derived from the quarterly model (2.4)-(2.5) and

Case (2) aggregation of the forecasts for z_T^P derived from the partially aggregated model (2.4) and (2.15).

Proof: See Appendix A.

Proposition 3 compares the ability of three models to predict the observable annual values. These are: (a) the model for the annual sample, (b) the "true" quarterly data generating process, and (c) the model for the partially aggregated sample. Assuming that these models are mutually consistent, given the aggregation constraint, model (b) would provide optimal forecasts, but cannot be empirically specified. Model (a) can be specified by standard methods, but cannot predict better than (b) or (c). Finally, model (c) can be inferred from data using, e.g. the method defined in Section 3, and may provide better forecasts than model (a). It is also more flexible, as it has the ability to update the annual forecast as new quarterly information becomes available.

Proposition 3 generalizes previous results. Wei (1978) gave a proof of Proposition 3, Case (1) for univariate processes, also showing that the loss in forecasting efficiency due to aggregation can be substantial if the nonseasonal component of the model is nonstationary. Conversely, there is no loss in efficiency if the quarterly model is a pure seasonal process. Lütkepohl (1987) also discussed Case (1) in a multivariate stationary framework.

3. The effect of aggregation on the observability of a time series model

3.1. Characterization of the modes that become unobservable after aggregation

As shown in Section 2, the models for quarterly, annual and partially aggregated data share the state equation (2.4). Therefore, aggregation does not affect the states governing a dynamic system. However, it may reduce its observability. Comparing the observation equations (2.5) and (2.14)-(2.15) it is immediate to see that observability loss occurs in two ways: (a) the quarterly observer (2.5) realizes more signals than the aggregated observers (2.14) or (2.15), and (b) the matrices in (2.14)-(2.15) are non-linear functions of the matrices in (2.5), potentially decreasing the observability of some states. To further discuss this issue, Definition 1 particularizes the general concept of observability (Anderson & Moore, 1979) to the notation defined in Section 2.

Definition 1 (observability in the quarterly model). All the components in the quarterly model (2.1)-(2.2) are said to be observable i.i.f. there is no real vector $\boldsymbol{w} \neq \boldsymbol{0}$, with $\boldsymbol{\Phi}\, \boldsymbol{w} = \lambda\, \boldsymbol{w}$ such that $\boldsymbol{H}\boldsymbol{w} = \boldsymbol{0}$, where λ is the eigenvalue associated to the eigenvector $\boldsymbol{w}$.

Definition 2 (unobservable modes in the annual model). Assuming that the variables are flows, so $\boldsymbol{J}^A = [\boldsymbol{I}, \boldsymbol{I}, \ldots, \boldsymbol{I}]$, the annual model (2.4) and (2.14) has unobservable modes i.i.f. there exists a vector $\boldsymbol{w} \neq \boldsymbol{0}$, with $\boldsymbol{\Phi}^S \boldsymbol{w} = \lambda\, \boldsymbol{w}$ such that $\boldsymbol{H}^A \boldsymbol{w} = \boldsymbol{0}$ or, equivalently, $\boldsymbol{H}\sum_{i=0}^{S-1}\boldsymbol{\Phi}^i\, \boldsymbol{w} = \boldsymbol{0}$, see (2.6)-(2.7) and (2.16).

Under these conditions, Theorem 1 characterizes what components of the quarterly model become unobservable after annual aggregation.

Theorem 1 (loss of observability in the annual model). Assuming that the quarterly model (2.1)-(2.2) is observable, the annual model (2.4) and (2.14) includes unobservable modes in any of the following cases:

Case (1) when the annual variables are flows, such that $\boldsymbol{J}^A = [\boldsymbol{I}, \boldsymbol{I}, \ldots, \boldsymbol{I}]$, if there exists an eigenvalue of $\boldsymbol{\Phi}$, λ, such that: $1 + \lambda + \lambda^2 + \ldots + \lambda^{S-1} = 0$

Case (2) if there is a subset of k eigenvalues of $\boldsymbol{\Phi}$, denoted by $\lambda_k(\boldsymbol{\Phi})$, such that for any pair $\lambda_i, \lambda_j \in \lambda_k(\boldsymbol{\Phi}), \lambda_i \neq \lambda_j$, it holds that: $\lambda_i^S = \lambda_j^S$

Case (3) when $\boldsymbol{\Phi}$ has a kxk Jordan block with null eigenvalues, such that the geometric multiplicity of this block grows with its S-th order power.

Proof: See Appendix B.

Theorem 1 characterizes the situation where some modes of the quarterly model, associated to different transition eigenvalues, collapse to a smaller number of modes in the annual model. Consider, e.g., a quarterly univariate model where all the dynamics of a flow variable are given by a seasonal difference $\nabla_S = (1-B)(1+B+B^2+...+B^{S\text{-}1})$. Case (1) implies that, after aggregation, this structure is indistinguishable from a nonseasonal difference $\nabla = (1-B)$. Then observability of S-1 dynamic components of the quarterly model is lost and the dynamics of the corresponding annual model simplify accordingly.

Under Case (2), aggregation reduces the number of system modes, while maintaining the dimension of the state vector. The number of unobservable modes is $k - rank(\boldsymbol{H} \cdot \boldsymbol{W})$, where k is the number of elements in $\lambda_k(\Phi)$ and $\boldsymbol{W}$ is a matrix whose columns are the eigenvectors $\boldsymbol{w}_i$, $i = 1, 2, \ldots, k$ that generate the k-order subspace $\boldsymbol{S_k}$. In the univariate case aggregation keeps only one mode, which corresponds with the larger Jordan block. In the multivariate case the reduction in the number of modes depends on how they affect the different observable time series so, in this case there is no way to compute *a priori* how many modes will become unobservable, but the maximum number of unobservable modes is $k-1$.

Case (3) occurs when the seasonal structure is a pure moving average. In this situation a systematic rule for mode elimination only exists in the univariate case, where there are $1+[(k-1)/S]$ distinguishable modes, being [] the integer part operator and k the dimension of the corresponding Jordan block.

Two important works connected with Theorem 1 are those of Stram & Wei (1986) and Wei & Stram (1990). Specifically, the loss of observability is related with the "hidden periodicity" defined by Stram & Wei (1986, Definition 4.1) for univariate models. In a univariate framework, Case (2) of Theorem 1 is equivalent to hidden periodicity because all the undistinguishable modes created by aggregation affect a single time series. However, in the multivariate case hidden periodicity not always implies loss of observability; assume e.g., that two modes with hidden periodicity affect two different time series. Note also that the mode elimination rules discussed for cases (2) and (3) are coherent in the univariate case with Stram & Wei (1986, Theorems 4.1 and 3.1).

After discussing the effect of aggregation over the quarterly model dynamics and its observability, it is important to characterize the uniqueness of the correspondence between the models for disaggregated and aggregated data. This is done in the following Theorem.

Theorem 2 (correspondence between quarterly and annual models). If the quarterly model (2.4)-(2.5) is minimal and all its components are observable from annual data, then the annual model (2.4) and (2.14) is unique, allowing for similar transformations, minimal and observable.

Proof: See Appendix C.

The reciprocal proposition is not true in general. There are minimal and observable annual models that necessary correspond to quarterly models with unobservable components. Assume e.g., that the model for annual data is an AR(1) with a negative parameter and that the seasonal frequency S is even. In this case, there is no high-frequency ARMA(1,1) process that adds up to the annual AR(1) model because the S-th power of the transition matrix will always be positive, see (2.6). In the case of ARIMA models this result collapses to Lemma 2 in Wei & Stram (1990).

3.2. Observability and fixed-interval smoothing

In a SS framework, the method of choice to estimate efficiently the system states is the fixed-interval smoother (Anderson & Moore, 1979), which is a two-sided symmetric filter providing estimates of the first and second-order moments of the states conditional on all the information in the sample. The uncertainty of smoothed estimates depends on a property called "detectability", which is defined in the following way.

Definition 3 (detectability). A system is said to be detectable if its unobservable modes are stationary.

Building on this concept, the following result characterizes the effect of undetectable modes on smoothed estimates.

Proposition 4. The variance of fixed-interval smoothing estimates of the states in models (2.4) and (2.14) or (2.15) is finite if and only if all the states are detectable.

Proof: See Appendix D.

In time series disaggregation smoothing is typically employed to estimate the unobserved quarterly values. Therefore detectability is a necessary and sufficient condition to estimate these values with bounded uncertainty, while observability is a sufficient (not necessary) condition.

Infinite smoothed variances arise, for example, when the target variables are flows and their quarterly model includes seasonal roots in the unit circle. In this case the aggregated model has undetectable components and the estimates of seasonal components would have infinite variances.

In practical terms this implies that, if quarterly indicators have seasonal components, it is advisable to remove them before disaggregation.

3.3. An algorithm to obtain the annual representation corresponding to a quarterly model

Combining Propositions 1 and 2 with Theorem 1 and other results from the SS literature, one can devise an algorithm to obtain the reduced-form model for a vector of annual data corresponding to any linear model for the quarterly values, allowing for a general aggregation constraint. This algorithm proceeds as follows:

Step (1): Consider any linear and fixed-coefficients model for the quarterly data. Write the model in the innovations form (2.1)-(2.2). If the model can be written in VARMAX form, this can be done using the expressions given by Terceiro (1990, Section 2.1). In any other case, write the model in a general (non-innovations) SS form and obtain the equivalent innovations representation (Casals, Sotoca & Jerez, 1999, Theorem 1).

Step (2): Obtain the equivalent quarterly representation (2.4)-(2.5) and the annual representation (2.4) and (2.14).

Step (3): If the annual representation is not observable, reduce (2.4) and (2.14) to an equivalent minimal SS realization applying the staircase algorithm (Rosenbrock, 1970).

Step (4): Transform the model obtained in Step (3) to the corresponding innovations form (Casals, Sotoca & Jerez, 1999, Theorem 1).

Step (5): The previous Step yields a model in SS form. If a different representation (e.g., VARMAX) is required, transform the innovations model to the Luenberger observable canonical form (Petkov *et al*. 1991).

Tables 2.a and 2.b show the aggregation of several univariate and bivariate models, illustrating the application of this algorithm and some previous results. Specifically:

1) Models # 1-3 are examples of the observability loss described in Theorem 1 cases (1), (2) and (3), respectively. In particular, aggregation of Model 1 shows how a seasonal difference collapses to a nonseasonal unit root. This result is coherent with Granger & Siklos (1995) and Stram & Wei (1986).

2) Models # 1, 4, 5 and 7-11 show that aggregation does not affect the number of unit roots in a time series, so I(0), I(1) or I(2) quarterly flows yield I(0), I(1) or I(2) annual aggregates. A straightforward implication of this is that, if the quarterly variables are cointegrated, the corresponding annual aggregates will be also cointegrated. This is consistent with the findings of Pierse & Snell (1995), Granger (1990) and Marcellino (1999).

3) Models # 4-5 assume that the quarterly variables have both, regular and seasonal unit roots. In this case, often found when analyzing seasonal data, the corresponding annual model has two unit roots. This suggests that many annual variables should be I(2) while, in practice, annual models are often specified with a single unit root. This apparent contradiction is easy to explain because aggregation may induce a MA root close to unity, see e.g. Model # 5, therefore compensating an AR unit root.

4) Aggregation induces additional MA structure and maintains the order of the stationary AR structure (allowing for observability loss) and typically reducing its persistency. Models # 2 and 8 are clear examples of this. This result is consistent in the univariate case with Amemiya & Wu (1972), Wei (1978) and Stram & Wei (1986). In the multivariate stationary case, it is consistent with the findings of Lütkepohl (1987, Chap. 4 and 6) and Marcellino (1999).

5) Models # 8-11 show that, if there is feedback in the quarterly frequency, there is feedback in the annual frequency.

6) Model # 6 shows the annual model corresponding to a quarterly Chow-Lin AR(1) regression. Therefore, this method is empirically justified only if the annual model relating the target variable and the indicator is a static regression with ARMA(1,1) errors.

7) Models # 7 and 11 show that the algorithm is not restricted to VARMAX or transfer functions, as it can be applied to structural time series models (Harvey, 1989) and VARMAX echelon models (Hannan & Deistler, 1988). In general, it supports any model with an equivalent SS representation.

8) Finally, model # 10 shows that the algorithm can be applied to a general combination of sampling and aggregation frequencies, as it shows how a monthly model aggregates to a quarterly VARMA.

[Insert Tables 2.a and 2.b]

4. An empirical method to specify a high-frequency model

Assume that a linear model has been fitted to all the available variables in the annual frequency. The problem now reduces to devise a systematic method enforcing consistency between the annual model and the unknown quarterly model, given the aggregation constraint and the partially aggregated sample.

Without loss of generality, we will refer to a VARMA specification process, consisting of the successive determination of unit roots, AR and MA dynamics. The basic ideas can be mapped to other model-building methods such as, e.g., structural time series modelling (Harvey, 1989).

4.1. Feasibility of an exact correspondence between the annual and quarterly models

The most rigorous way to specify the quarterly model would consist of obtaining a numerical solution to the equations relating the known annual model and the unknown quarterly model, using the algorithm defined in Section 3.3. We tried this approach and found it unpractical because it is difficult, unrealistic and may be impossible in some cases.

First, it is difficult because the equations relating the SS matrices of the annual and quarterly models are highly nonlinear. Perhaps they can be solved, but we have not been able to devise a procedure to do it consistently.

Second, it is unrealistic because achieving an exact match between the true quarterly data generating process and an empirical annual model would require the ability to model very weak parameters in the annual frequency. For example, consider the models # 2 and 8-10 in Table 2.a. Obviously some parameters in the MA factors are small to be detected by a realistic analysis of the annual time series, so an exact fit between the annual and quarterly models cannot be expected in practice.

Third it may be impossible in some cases because, as stated in the discussion of Theorem 2, a statistically adequate model for the annual data may not have a mathematically consistent quarterly representation.

4.2. A method to enforce approximate consistency

If an exact correspondence between the quarterly and annual models cannot be expected to be found in practice, the only way forward consists of devising a systematic process to achieve an approximate fit and a diagnostic method to assess whether the quarterly model obtained is statistically adequate or not. The following procedure can be used to these purposes.

Step (1) Annual modelling. Specify and estimate a model relating the target annual variable(s) with the annualized values of the quarterly indicator(s). Any model having an equivalent linear SS representation, such as e.g., a transfer function or VARMAX, is adequate for this purpose.

Step (2) Decomposition of the quarterly indicator. Adjust the quarterly indicator(s) to suppress undesired features, such as seasonality and calendar effects.

Step (3) Model specification.

Step (3.1) Set the VAR factor order of the quarterly model to be equal to that of the annual model and, particularly, constrain the number of unit roots to be the same. The foundation of this step results from comparison of the quarterly model (2.4)-(2.5) and the annual model given by (2.4) and (2.14). As both models share the same state equation, they have the same (stationary and nonstationary) autoregressive components.

Step (3.2) Add a VMA(q) structure, with $q \leq n$, being n the size of the state vector in the annual model. This bound to MA dynamics results from the fact that, in a minimal SS representation, the size of the state vector is the maximum of p and q, being p the order of the VAR factor and q the order of the VMA factor.

Step (4) Estimation. Estimate the model specified in Steps (2) and (3) by maximum likelihood and prune insignificant parameters to obtain a parsimonious parameterisation.

Step (5) Diagnostics. Check the final quarterly model by obtaining the corresponding annual representation, using the algorithm described in Section 3.3, and then:

Step (5.1) compare this model with the one specified in Step (1) and

Step (5.2) check whether it filters the annual data to white noise residuals.

Step (6) Forecast accuracy check. If the sample is long enough, compare out-of-the-sample forecasts for the annual values produced by both, the tentative quarterly model and the annual model specified in Step (1).

4.3. Practical suggestions

We have applied the method described above to several real and simulated time series. These exercises provided some useful insights about the practical application of our method:

First, a good characterization of unit roots in Step (1) is critical, as misspecification of these components impacts severely over the quality of final results (Tiao, 1972). When in doubt, over-differencing is safer than under-differencing in our opinion.

Second, the number of MA parameters specified in Step (3.2) may be excessive, depending on the sample size and number of time series. In this case, it is a good idea to constrain the MA matrices to be diagonal and, later, add off-diagonal parameters in a sequence of overfitting experiments.

Third, as Nunes (2005) points out, there are many situations where missing observations arise in a time series disaggregation frameworks. Some of these are: different release dates for some variables, changes in the sampling frequency or non-conformable samples, where the series considered have different starting dates. In these situations the ability of SS methods to take care of missing values is an important asset, as it allows using all the information available in the dataset.

Last, the forecasting accuracy check proposed in Step (6) can be implemented by setting some within-the-sample values to missing and estimating them afterwards. We have found this alternative useful when the sample is too short to reserve some values for out-of-the-sample forecasting.

The example in Section 5 illustrates how the last two ideas can be applied in practice.

5. Disaggregation of Value Added by Industry in Spain (1980-2001).

This Section illustrates the application of the method proposed and its practical advantages in comparison with alternative procedures. To do this, we will disaggregate and forecast the annual series of Value Added by the Industry in Spain (VAI), from 1980 to 2001, using as indicator the quarterly values of a re-balanced Industrial Production Index (IND), from 1980 1st Quarter to 2001 4th Quarter. The latter series is the indicator actually employed by the Agency in charge of the Spanish national accounts.

To clarify when a series is expressed in annual or quarterly frequency we will use an uppercase/lowercase notation, so VAI_T^A and IND_T^A are the values of VAI and the annual average of the indicator in year T (T=1980, 1981, … , 2001), while vai_t and ind_t denote the values of both variables in quarter t (t=1980;1, 1980;2, … , 2001;4).

5.1. Step (1) Annual model

The first step in the analysis consists of modelling the annual values of the target variable and the indicator. A standard analysis (Jenkins & Alavi, 1981) yielded the following model:

$$\begin{bmatrix} (1-B)^2 & 0 \\ 0 & (1-B)^2 \end{bmatrix} \begin{bmatrix} VAI_T^A \\ IND_T^A \end{bmatrix} = \begin{bmatrix} 1-\underset{(.06)}{1.85}\,B & \underset{(.02)}{.74}\,B \\ -\underset{(.05)}{1.48}\,B & 1 \end{bmatrix} \begin{bmatrix} \hat{a}_T^{VAI} \\ \hat{a}_T^{IND} \end{bmatrix} ; (T=1980, 1981, \ldots, 2001)$$

$$\hat{\boldsymbol{\Sigma}}_a = \begin{bmatrix} 2.64 & -- \\ 5.53 & 13.08 \end{bmatrix} ; \boldsymbol{Q}(5) = \begin{bmatrix} 4.84 & 4.31 \\ 6.23 & 4.17 \end{bmatrix} \quad (5.1)$$

where the figures in parentheses are standard errors of the estimates, $\hat{\boldsymbol{\Sigma}}_a$ represents the estimate of the error covariance matrix and $\boldsymbol{Q}(5)$ is the matrix of Ljung-Box statistics, computed using the first five residual auto and cross-correlations. Note that this model implies the existence of feedback between VAI and the indicator. Therefore, if model (5.1) is statistically adequate any procedure assuming unidirectional causality, such as those in Table 1, would be unsuitable for these series.

Model (5.1) provides precise clues about the dynamics of both variables in the quarterly frequency. As discussed in Section 4.2, a consistent quarterly model must have a non-stationary AR structure with two unit roots for each series and an MA term with a maximum order of 4, because the minimal SS representation of model (5.1) requires four state variables.

5.2. Step (2) Decomposition of the quarterly indicator

The second step requires modelling the indicator in the quarterly frequency to estimate the components useful for disaggregation. To this end, we will use the following model:

$$ind_t = \underset{(.03)}{1.02}L_t - \underset{(.35)}{1.85}E_t - \underset{(1.72)}{5.94}S_t^{92.4} + \hat{N}_t \;;\; (t = 1980;1, 1980;2, \ldots, 2001;4)$$

$$(1-B)(1-B^4)\hat{N}_t = (1-\underset{(.06)}{.83}B^4)\hat{a}_t \;;\; \hat{\sigma}_a^2 = 3.21 \;; Q(15) = 12.80 \quad (5.2)$$

where L_t is the number of non-holidays in quarter *t*, E_t is a dummy variable to account for Easter effects and $S_t^{92.4}$ is an intervention variable capturing a persistent level change from 1992 4[th] Quarter onwards.

Model (5.2) implies that the indicator can be split into: (a) trend, (b) seasonal component, (c) irregular component, (d) calendar effects associated to the number of non-holidays and Easter; and (e) the step 1992 4[th] Quarter effect, see Casals, Jerez & Sotoca (2002). Figure 1 shows these components. In the light of previous results, components (b) and (d) are useless for disaggregation. By adding the remaining components we obtain a seasonally and calendar-adjusted quarterly indicator series.

[Insert Figure 1]

5.3. Steps (3) and (4) Specification and estimation of the quarterly model

Building on the results of previous steps, we now estimate a doubly integrated VMA(4) process for the value added and the adjusted indicator in the quarterly frequency. After pruning insignificant parameters we obtained the following model:

$$\begin{bmatrix} (1-B)^2 & 0 \\ 0 & (1-B)^2 \end{bmatrix} \begin{bmatrix} vai_t \\ \widetilde{ind}_t \end{bmatrix} = \begin{bmatrix} 1-\underset{(.06)}{1.48}\,B+\underset{(.06)}{.36}\,B^2 & \underset{(.003)}{.044B} \\ \underset{(.06)}{-.84B} & 1-\underset{(.02)}{.78B} \end{bmatrix} \begin{bmatrix} \hat{a}_t^{vai} \\ \hat{a}_t^{\widetilde{ind}} \end{bmatrix}; \hat{\boldsymbol{\Sigma}}_a = \begin{bmatrix} .06 & -- \\ .40 & 3.64 \end{bmatrix} \quad (5.3)$$

$$(t = 1980;1, 1980;2, \ldots, 2001;4)$$

where vai_t denotes the unobserved value of VAI in quarter *t* and $\widetilde{ind}_t$ is the corresponding adjusted indicator. Finally, by applying a fixed-interval smoothing to the sample (Casals Jerez & Sotoca 2000) we obtain the disaggregates and forecasts shown in Figure 2.

[Insert Figure 2]

5.4. Step (5) Diagnostics

We now obtain the annual representation corresponding to model (5.3) using, the algorithm described in Section 3.3. It is:

$$
\begin{bmatrix} (1-B)^2 & 0 \\ 0 & (1-B)^2 \end{bmatrix} \begin{bmatrix} VAI_T^A \\ \widetilde{IND}_T^A \end{bmatrix} = \begin{bmatrix} 1-1.52B-.23B^2 & .70B+.10B^2 \\ -1.18B-.19B^2 & 1+.14B-.07B^2 \end{bmatrix} \begin{bmatrix} \hat{a}_T^{VAI} \\ \hat{a}_T^{\widetilde{IND}} \end{bmatrix}
$$

$$
\hat{\boldsymbol{\Sigma}}_a = \begin{bmatrix} 2.54 & -- \\ 5.53 & 13.79 \end{bmatrix}; \; \boldsymbol{Q}(5) = \begin{bmatrix} 7.38 & 7.81 \\ 9.49 & 8.48 \end{bmatrix} \qquad (T = 1980, 1981, \ldots, 2001) \tag{5.4}
$$

where $\widetilde{IND}_T^A$ denotes the annual average of the adjusted indicator in year *T*. Models (5.4) and (5.1) differ mainly in the additional second-order MA parameters in (5.4). We tried to fit an annual model including these parameters and the corresponding estimates resulted insignificant. Perhaps this additional MA structure is due to the disaggregated indicator information included in (5.3) and excluded in (5.1). On the other hand, the residuals obtained by filtering the annual series using (5.4) are stationary, normal and do not show important autocorrelations. Therefore, we accept that model (5.3) is statistically adequate and roughly conformable with (5.1).

Previous results in this example show that our method can be applied to real disaggregation problems. The remaining Subsections highlight its advantages when dealing with non-conformable samples and in terms of forecasting power.

5.5. Step (6) Forecast accuracy and non-conformable samples

Statistical Bureaux typically have long records for the target variable, say GDP, and shorter ones for the indicators. In this situation, standard techniques constrain the analysis to the common time window, thus assuming a substantial information loss. However the SS methods employed here allow for missing values, so all the available information can be used. To illustrate this idea, we delete the first four observations of the quarterly indicator and the last annual value of VAI. Re-estimating model (5.3) with this non-conformable sample yields the following results:

$$
\begin{bmatrix} (1-B)^2 & 0 \\ 0 & (1-B)^2 \end{bmatrix} \begin{bmatrix} vai_t \\ \widetilde{ind}_t \end{bmatrix} = \begin{bmatrix} \underset{(.04)}{1-1.48\,B} \underset{(.03)}{+.34\,B^2} & \underset{(.003)}{.045B} \\ \underset{(.04)}{-.89B} & \underset{(.02)}{1-.78B} \end{bmatrix} \begin{bmatrix} \hat{a}_t^{vai} \\ \hat{a}_t^{\widetilde{ind}} \end{bmatrix}; \; \hat{\boldsymbol{\Sigma}}_a = \begin{bmatrix} .06 & -- \\ .41 & 3.74 \end{bmatrix} \tag{5.5}
$$

$$
(t = 1980;1, 1980;2, \ldots, 2001;4)
$$

Estimates in (5.5) are remarkably close to those in (5.3). Table 3 shows the original sample information and the numerical results obtained with models (5.3) and (5.5). Note that: (a) the forecast provided by (5.5) for the 2001 annual value is very accurate, (b) quarterly interpolations and out-of-sample forecasts obtained with both models are very similar, and (c) the indicator retropolations computed with (5.5) and the non-conformable sample have acceptable errors.

[Insert Table 3]

5.6 Comparison with mainstream methods

Table 4 shows the root mean squared errors (RMSEs) obtained by forecasting the last five end-of-year values of VAI, using Model (5.3) and the main alternative methods. In all cases, the forecasts are conditional to the true indicator values for the same years. For example, to compute the 1997 values each model was estimated using annual VAI and quarterly indicator values up to 1996, and was then used to predict the end-of-year value of VAI using: (a) the past of VAI up to 1996 and (b) past values of the indicator up to 1997. The same procedure was applied for 1998, 1999, 2000 and 2001 extending the sample in each case with one, two, three and four additional years of data.

[Insert Table 4]

Note that Model (5.3) produced the best forecasts, with a 8-9% advantage over the second-best and much larger gains in comparison with the remaining methods. Obviously, this comparison is not fair, as model (5.3) was carefully fitted to the data while the alternative forecasts are mechanical. Therefore previous results should probably be seen as a re-statement of the idea that a model fitted by a trained human is usually able to beat automatic forecasting systems.

6. Concluding remarks

This paper makes three main contributions to the literature about aggregation of time series:

First, it encompasses and extends many previous results about: (a) the effect of aggregation on the dynamics of ARIMA (Amemiya & Wu, 1972; Brewer, 1973; Stram & Wei, 1986) and VARMA processes (Marcellino, 1999), (b) the observability loss due to aggregation (Wei & Stram, 1990) and (c) the negative effect of aggregation on forecasting accuracy in the univariate (Wei, 1978) and multivariate cases (Lütkepohl, 1987).

Second, it proposes an algorithm to aggregate any quarterly model to the corresponding annual reduced-form. This method is very general, as the only formal requirement for the quarterly model is that it has an equivalent state-space representation.

Third, it proposes a method to build an observable high-frequency model from a partially aggregated sample, allowing for general sampling frequencies and aggregation constraints. This method emphasizes the idea that the quarterly model should be both, consistent with a statistically adequate model fitted to annual data, and compliant with standard diagnostic tests.

When presenting this paper to professional audiences a common question was: why would I want to use this complex method instead a simpler mechanical procedure? We think that simple methods are adequate in complex situations, e.g., when one has to disaggregate thousands of time series with limited resources, to produce official statistics following a rigid calendar. This situation, which is very common in statistical agencies, requires robust, fast and mechanic methods. On the other hand, forecasters often concentrate in a smaller set of time series and may want to use the model with the higher forecasting power. The example in Section 5 shows that, for these needs, our modelling approach has the potential to beat mechanical methods.

The procedures described in this article are implemented in a MATLAB toolbox for time series modelling called E^4, which can be downloaded at *www.ucm.es/info/icae/e4*. The source code for all the functions in the toolbox is freely provided under the terms of the GNU General Public License. This site also includes a complete user manual and other materials.

References

Amemiya, T. & Wu, R.Y. (1972). The Effect of Aggregation on Prediction in the Autoregressive Model, *Journal of the American Statistical Association*, 339, 628-632.

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

Ansley, C.F. & Kohn, R. (1983). Exact Likelihood of Vector Autoregressive-Moving Average Process with Missing or Aggregated Data, *Biometrika*, 70, 1, 275–278.

Ansley, C. F. & Kohn, R. (1989). Filtering and Smoothing in State Space Models with Partially Diffuse Initial Conditions, *Journal of Time Series Analysis*, 11, 275–293.

Bitmead, R.R., Gevers, M.R. Petersen, I.R. & Kaye R.J. (1985). Monotonicity and Stabilizability Properties of the Solutions of the Riccati Difference Equation: Propositions, Lemmas, Theorems, Fallacious Conjectures and Counterexamples, *System Control Letters*, 5, 309-315.

Brewer, K.W. (1973). Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models, *Journal of Econometrics*, 1, 133-154.

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

Casals, J. Jerez, M. & Sotoca, S. (2002). An Exact Multivariate Model-based Structural Decomposition, *Journal of the American Statistical Association*, 97, 458, 553-564.

Chow, G.C. & Lin, A.L. (1971). Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series, *The Review of Economics and Statistics*, 53, 372-375.

De Jong, P. (1991). The diffuse Kalman filter, *The Annals of Statistics*, 19, 1073–1083.

Denton, F.T. (1971). Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization, *Journal of the American Statistical Association*, 66, 333, 99-102.

Di Fonzo, T. (1990). The Estimation of M Disaggregate Time Series when Contemporaneous and Temporal Aggregates are Known, *The Review of Economics and Statistics*, 72, 1, 178-182.

Durbin, J. & Quenneville, B. (1997). Benchmarking by State Space Models, *International Statistical Review / Revue Internationale de Statistique*, 65, 1, 23-48.

Fernández, R.B. (1981). A Methodological Note on the Estimation of Time Series, *The Review of Economics and Statistics*, 63, 471-476.

Granger, C.W.J. (1990). Aggregation of Time-Series Variables: A Survey, in *Disaggregation in Econometric Modelling*, eds. T. Barker and M.H. Pesaran, pp. 17-34, Routledge, London.

Granger, C.W.J. & Siklos, P.R. (1995). Systematic Sampling, Temporal Aggregation, Seasonal Adjustment and Cointegration. Theory and Evidence, *Journal of Econometrics*, 66, 357-369.

Hannan, E. J. & Deistler, M. (1988). *The Statistical Theory of Linear Systems*. New York: John Wiley,

Harvey, A.C. & Pierse, R.G. (1984). Estimating Missing Observations in Economic Time Series, *Journal of the American Statistical Association*, 79, 385, 125-131.

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

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

Kalman, R.E. (1963). Mathematical Description of Linear Systems, *SIAM Journal of Control*, 1, 152-192.

Litterman, R.B. (1983). A Random Walk, Markov Model for the Distribution of Time Series, *Journal of Business and Economic Statistics*, 1, 169-173.

Lütkepohl, H. (1987). *Forecasting Aggregated Vector ARMA Processes*, Berlin: Springer-Verlag.

Marcellino, M. (1999). Some Consequences of Temporal Aggregation in Empirical Analysis, *Journal of Business and Economic Statistics*, 1, 129-136.

Nunes, L.C. (2005). Nowcasting Quarterly GDP Growth in a Monthly Coincident Indicator Model, *Journal of Forecasting*, 24, 575-592.

Onof, C. Wheater, H. Chandler, D. Isham, V. Cox, D.R. Kakou, A. Northrop, P. Oh, L. & Rodriguez-Iturbe, I. (2000). Spatial-Temporal Rainfall Fields: Modelling and Statistical Aspects, *Hydrology and Earth System Sciences*, 4, 4, 581-601.

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

Pierse, R.G. & Snell, A.J. (1995). Temporal Aggregation and the Power of Tests for a Unit Root, *Journal of Econometrics*, 65, 333-345.

Proietti, T. (2006). Temporal Disaggregation by State Space Methods: Dynamic Regression Methods Revisited, *Econometrics Journal*, 9, 357-372.

Rosenbrock, M.M. (1970). *State-Space and Multivariable Theory*, New York: John Wiley.

Santos-Silva, J.M.C. & Cardoso, F. (2001). The Chow-Lin Method using Dynamic Models, *Economic Modelling*, 18, 269-280.

Stram, D.O. & Wei, W.W.S. (1986). Temporal Aggregation in the ARIMA Process, *Journal of Time Series Analysis*, 7, 4, 279-292.

Tiao, G.C. (1972). Asymptotic Behavior of Time Series Aggregates, *Biometrika,* 59, 521-531.

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

Wei, W.W.S. (1978). Some Consequences of Temporal Aggregation in Seasonal Time Series Models, in: Zellner, A. (ed.), *Seasonal Analysis of Economic Time Series,* Washington DC: Bureau of the Census, 433-448.

Wei, W.W.S. & Stram, D.O. (1990). Disaggregation of Time Series Models, *Journal of the Royal Statistical Society, B Series,* 52, 3, 453-467.

Figure 1. Decomposition of the quarterly indicator. The adjusted indicator includes the trend, irregular and level change components and excludes seasonality and calendar affects. Disaggregates based in this indicator can be interpreted as calendar and seasonally adjusted quarterly estimates of VAI.

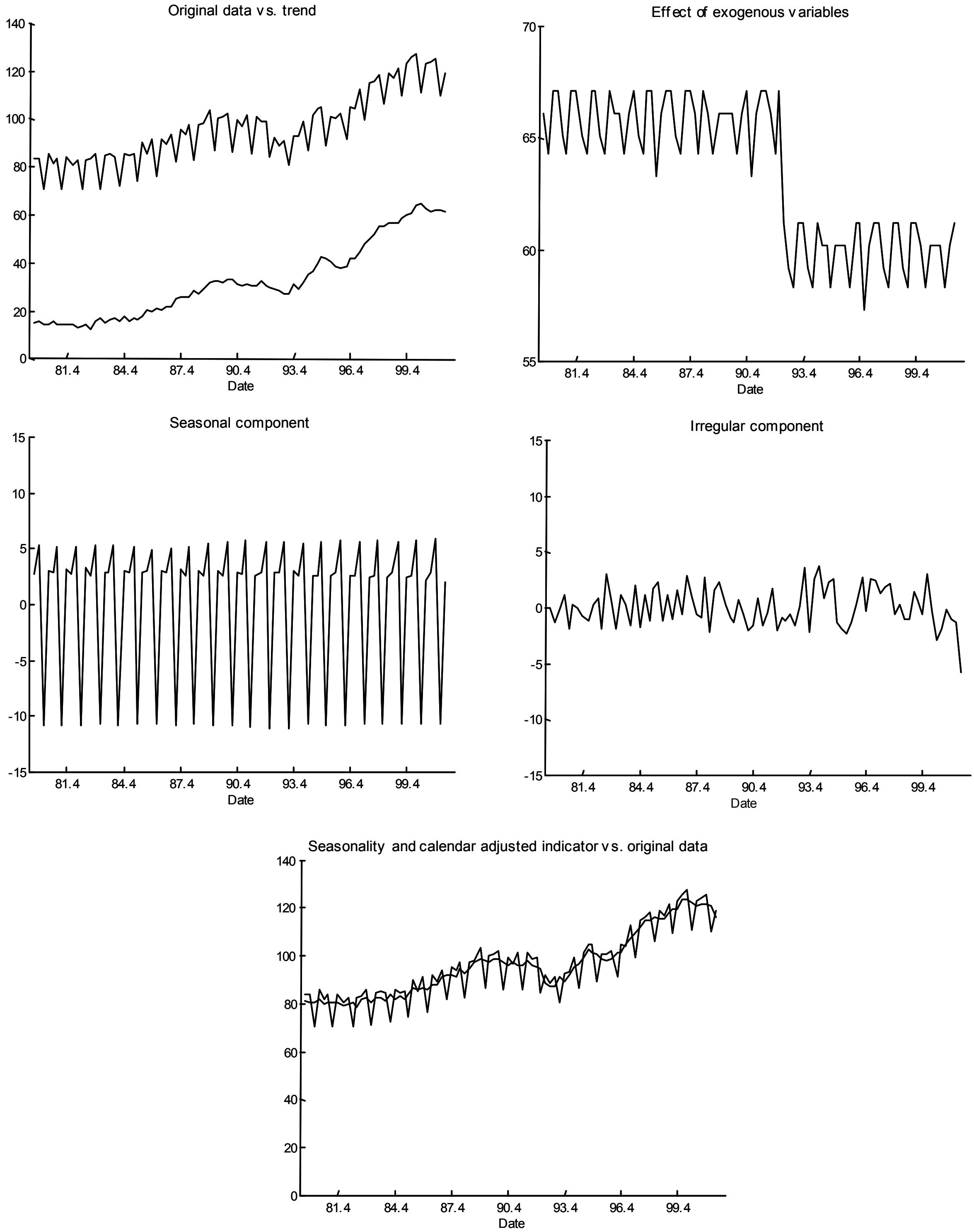

Figure 2. Standardized plot of quarterly estimates of VAI (thick line) and adjusted indicator (thin line). The last values are forecasts computed from 2002.1 to 2002.4 quarters using model (5.3).

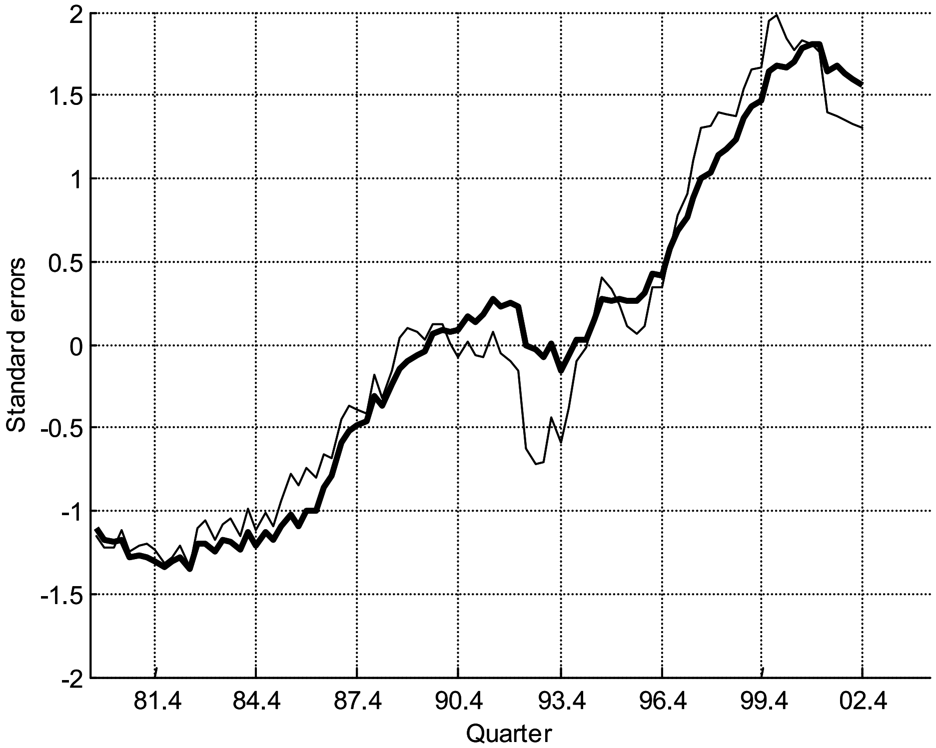

Table 1: Restrictions on model (1.1) assumed by different methods. The symbol (*) means that the corresponding parameter is to be estimated.

Method	β	φ_1	φ_2
Denton (1971)	1	1	0
Chow-Lin (1971)	(*)	0	0
Chow-Lin (1971) / AR(1)	(*)	(*)	0
Fernández (1981)	(*)	1	0
Litterman (1983)	(*)	1	(*)

Table 2.a: Aggregation of several univariate models. The columns labelled "states" show the number of dynamic components in the minimal SS representation of the corresponding model. Therefore, the difference between the number of states in the quarterly and annual models is the number of dynamic components that become unobservable after aggregation. The quarterly variables z_t and z_{1t} are assumed to be flows, so their annual aggregate is the sum of the corresponding quarterly values. If the quarterly indicator in model # 6 (z_{2t}) were to be aggregated as an annual average, the coefficients in the annual transfer function should be multiplied by 4.

#	Quarterly Model	States	Annual Model	States
1	$(1-B^4)z_t = a_t$; $\sigma_a^2 = 1$	4	$(1-B)z_t^A = a_t^A$; $\sigma_A^2 = 4.00$	1
2	$(1-.8B)(1+.8B)z_t = a_t$; $\sigma_a^2 = 1$	2	$(1-.410B)z_t^A = (1+.160B)a_t^A$; $\sigma_A^2 = 7.99$	1
3	$z_t = (1-.6B^4)a_t$; $\sigma_a^2 = 1$	4	$z_t^A = (1-.600B)a_t^A$; $\sigma_A^2 = 4.00$	1
4	$(1-B)(1-B^4)z_t = a_t$; $\sigma_a^2 = 1$	5	$(1-B)^2 z_t^A = (1+.240B)a_t^A$; $\sigma_A^2 = 41.60$	2
5	$(1-B)(1-B^4)z_t = (1-.8B)(1-.6B^4)a_t$; $\sigma_a^2 = 1$	5	$(1-B)^2 z_t^A = (1-.997B+.238B^2)a_t^A$; $\sigma_A^2 = 7.05$	2
6	$z_{1t} = .5\, z_{2t} + \frac{1}{1-.8B} a_t$; $\sigma_a^2 = 1$	1	$z_{1t}^A = .500\, z_{2t}^A + \frac{1+.228B}{1-.410B} a_{1t}^A$; $\sigma_a^2 = 23.103$	1
7	$(1-B)T_{t+1} = u_t$; $\sigma_u^2 = .01$ $(1+B+B^2+B^3)S_{t+1} = v_t$; $\sigma_v^2 = .01$ $z_t = T_t + S_t + \varepsilon_t$; $\sigma_\varepsilon^2 = 1$	4	$(1-B)z_t^A = (1-.669B)a_t^A$; $\sigma_A^2 = 5.844$	1

Table 2.b: Aggregation of several bivariate models. All the quarterly variables z_{1t} and z_{2t} are assumed to be flows, so their annual aggregate is the sum of the corresponding quarterly values. If a variable were to be aggregated as an average of the quarterly values, the model would require an appropriate re-scaling.

#	High-frequency Model	States	Low-frequency Model	States
8	$\begin{bmatrix} 1-.7B & -.8B \\ 0 & 1 \end{bmatrix}\begin{bmatrix} (1-B^4)z_{1t} \\ (1-B^4)z_{2t} \end{bmatrix}=\begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}; \mathbf{\Sigma}_a=\begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}$	10	$\begin{bmatrix} 1-.240B & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} (1-B)z_{1t}^A \\ (1-B)z_{2t}^A \end{bmatrix}=\begin{bmatrix} 1-.039B & 1.461B \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a_{1t}^A \\ a_{2t}^A \end{bmatrix}$ $\mathbf{\Sigma}_A=\begin{bmatrix} 35.36 & 7.62 \\ 7.62 & 4.00 \end{bmatrix}$	3
9	$\begin{bmatrix} (1-B^4)z_{1t} \\ (1-B^4)z_{2t} \end{bmatrix}=\begin{bmatrix} 1 & -.8B \\ 1.2B & 1 \end{bmatrix}\begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}; \mathbf{\Sigma}_a=\begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}$	8	$\begin{bmatrix} (1-B)z_{1t}^A \\ (1-B)z_{2t}^A \end{bmatrix}=\begin{bmatrix} 1-.080B & -.053B \\ .289B & 1+.015B \end{bmatrix}\begin{bmatrix} a_{1t}^A \\ a_{2t}^A \end{bmatrix}; \mathbf{\Sigma}_A=\begin{bmatrix} 4.09 & 1.41 \\ 1.41 & 13.00 \end{bmatrix}$	2
10	$\begin{bmatrix} (1-B^{12})z_{1t} \\ (1-B^{12})z_{2t} \end{bmatrix}=\begin{bmatrix} 1 & -.8B \\ 1.2B & 1 \end{bmatrix}\begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}; \mathbf{\Sigma}_a=\begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}$	24	$\begin{bmatrix} (1-B^4)z_{1t}^A \\ (1-B^4)z_{2t}^A \end{bmatrix}=\begin{bmatrix} 1-.100B & -.075B \\ .365B & 1+.024B \end{bmatrix}\begin{bmatrix} a_{1t}^A \\ a_{2t}^A \end{bmatrix}; \mathbf{\Sigma}_A=\begin{bmatrix} 3.22 & 1.03 \\ 1.03 & 9.27 \end{bmatrix}$	8
11	$\begin{bmatrix} 1-B & 0 \\ -.5 & 1 \end{bmatrix}\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix}=\begin{bmatrix} 1-.4B & .8B \\ -.5 & 1 \end{bmatrix}\begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}; \mathbf{\Sigma}_a=\begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}$	1	$\begin{bmatrix} 1-B & 0 \\ -.5 & 1 \end{bmatrix}\begin{bmatrix} z_{1t}^A \\ z_{2t}^A \end{bmatrix}=\begin{bmatrix} 1-.745B & 1.808B \\ -.5 & 1 \end{bmatrix}\begin{bmatrix} a_{1t}^A \\ a_{2t}^A \end{bmatrix}; \mathbf{\Sigma}_A=\begin{bmatrix} 51.91 & 29.55 \\ 29.55 & 19.58 \end{bmatrix}$	1

Table 3. Original data and interpolations obtained with full and non-conformable samples. Underlined values correspond to interpolations, retropolations or forecasts. The column "Indicator" refers to the calendar and seasonally adjusted values computed using model (5.2).

	Original sample		**Full sample**			**Non-conformable sample**		
Obs.	**Annual VAI**	**Indicator**	**Interpolation**	**Annual sum**	**Indicator**	**Interpolation**	**Annual sum**	**Indicator**
1980.1		81.421	14.993		81.421	15.009		82.869
1980.2		80.451	14.810		80.451	14.896		82.237
1980.3		80.524	14.744		80.524	14.778		81.584
1980.4	59.341	81.877	14.794	59.341	81.877	14.658	59.341	80.968
1981.1		80.196	14.474		80.196	14.502		80.196
1981.2		80.710	14.531		80.710	14.519		80.710
1981.3		80.823	14.483		80.823	14.474		80.823
1981.4	57.895	80.299	14.407	57.895	80.299	14.400	57.895	80.299
...								
2000.1		123.520	23.326		123.520	23.342		123.520
2000.2		124.005	23.417		124.005	23.431		124.005
2000.3		122.033	23.374		122.033	23.371		122.033
2000.4	93.620	121.147	23.502	93.620	121.147	23.476	93.620	121.147
2001.1		121.823	23.753		121.823	23.690		121.823
2001.2		121.550	23.801		121.550	23.718		121.550
2001.3		121.006	23.830		121.006	23.734		121.006
2001.4	94.711	115.986	23.327	94.711	115.986	23.217	94.359	115.986
2002.1			23.414		115.664	23.311		115.723
2002.2			23.295		115.343	23.199		115.461
2002.3			23.177		115.021	23.088		115.198
2002.4			23.059	92.944	114.699	22.976	92.574	114.936

Table 4: Ranking of RMSEs for end-of year forecasts of annual VAI from 1997 to 2001. RMSEs are computed for forecasts of both, VAI levels and growth rates. The columns RMSE% show the corresponding RMSEs normalized so that the RMSE of Model (5.5) is 100%

		Levels		Annual growth rates	
Rig	**Model**	**RMSE**	**RMSE %**	**RMSE**	**RMSE %**
1	Model (5.5)	0.65	100.0%	0.71	100.0%
2	Litterman (1983) with constant	0.72	109.6%	0.77	108.4%
3	Fernández (1981) with constant	0.96	147.5%	1.03	145.2%
4	Chow Lin (1971) AR(1) with constant	1.21	185.2%	1.30	182.5%
5	Fernández (1981) without constant	1.44	219.8%	1.54	216.0%
6	Litterman (1983) without constant	1.47	224.9%	1.57	220.9%
7	ADL(1,0) without constant	1.53	234.0%	1.73	242.4%
8	ADL(1,0) with constant	1.61	247.0%	1.75	245.3%
9	Chow Lin (1971) AR(1) without constant	1.82	278.9%	1.98	277.2%
10	Boot, Fibs and Lisman (1967)	3.04	466.2%	3.39	475.2%
11	Denton (1971)	18.01	2758.4%	20.26	2842.3%

APPENDIX A: PROOF OF PROPOSITION 3

A.1. Previous results

Result 1. Consider the algebraic Riccati equation of the Kalman filter:

$$\boldsymbol{P}_{t+1|t} = \Phi\, \boldsymbol{P}_{t|t-1}\Phi^T + \boldsymbol{EQE}^T - \boldsymbol{K}_t\boldsymbol{B}_t\boldsymbol{K}_t^T \tag{A.1}$$

where:

$$\boldsymbol{K}_t = (\Phi\, \boldsymbol{P}_{t|t-1}\boldsymbol{H}^T + \boldsymbol{EQ})\boldsymbol{B}_t^{-1} \tag{A.2}$$

$$\boldsymbol{B}_t = \boldsymbol{H}\,\boldsymbol{P}_{t|t-1}\boldsymbol{H}^T + \boldsymbol{Q} \tag{A.3}$$

Under these conditions, if $\boldsymbol{P}^*_{t|t-1} \geq \boldsymbol{P}_{t|t-1}$ and $\boldsymbol{Q}^* \geq \boldsymbol{Q}$ then $\boldsymbol{P}^*_{t+1|t} \geq \boldsymbol{P}_{t+1|t}$. See the proof in Bitmead *et al.* (1985)

Result 2. Let be $\boldsymbol{M}$ a ($m^* \times m$) matrix where $rank(\boldsymbol{M}) = m^*$ with $m^* < m$, and $\boldsymbol{A}$, a symmetric positive-definite ($m \times m$) matrix, with $rank(\boldsymbol{A}) = m$. Under these conditions:

$$\boldsymbol{A} - \boldsymbol{M}^T(\boldsymbol{MA}^{-1}\boldsymbol{M}^T)\boldsymbol{M} \geq 0 \tag{A.4}$$

$$rank[\boldsymbol{A} - \boldsymbol{M}^T(\boldsymbol{MA}^{-1}\boldsymbol{M}^T)\boldsymbol{M}] = m - m^* \tag{A.5}$$

Proof. It is immediate to see that:

$$\boldsymbol{A} - \boldsymbol{M}^T(\boldsymbol{MA}^{-1}\boldsymbol{M}^T)\boldsymbol{M} = \boldsymbol{A} - \boldsymbol{CAC}^T = (\boldsymbol{I}-\boldsymbol{C})\boldsymbol{A}(\boldsymbol{I}-\boldsymbol{C})^T \tag{A.6}$$

where $\boldsymbol{C} = \boldsymbol{M}^T(\boldsymbol{MA}^{-1}\boldsymbol{M}^T)^{-1}\boldsymbol{MA}^{-1}$. Therefore, positive-definiteness of $\boldsymbol{A}$ assures that $(\boldsymbol{I}-\boldsymbol{C})\boldsymbol{A}(\boldsymbol{I}-\boldsymbol{C})^T$ is a positive-definite matrix also.

A.2. Proof of Proposition 3

Case 1) Consider the high-frequency model in the stacked form (2.4)-(2.5). The corresponding *k*-step ahead forecast error variance is given by:

$$var(\boldsymbol{e}_K\,|\Omega_T) = \bar{\boldsymbol{H}}\,\boldsymbol{P}_{T+K-1|T}\bar{\boldsymbol{H}}^T + \bar{\boldsymbol{C}}\bar{\boldsymbol{Q}}\bar{\boldsymbol{C}}^T \tag{A.7}$$

where *T* is the forecast origin and Ω_T denotes the quarterly information available at time *T*. According to (A.7), the variance of the annual forecast obtained by aggregation would be:

$$var(\boldsymbol{e}_K^A | \Omega_T) = \boldsymbol{J}^A var(\boldsymbol{e}_K | \Omega_T) \left(\boldsymbol{J}^A\right)^T \tag{A.8}$$

On the other hand, the Kalman filter covariance matrices $\boldsymbol{P}_{T+K\text{-}1|T}$ in (A.7) result from the recursion:

$$\boldsymbol{P}_{T+K|T} = \bar{\Phi}\, \boldsymbol{P}_{T+K\text{-}1|T} \bar{\Phi}^T + \bar{\boldsymbol{E}}\bar{\boldsymbol{Q}}\bar{\boldsymbol{E}}^T \tag{A.9}$$

and successive substitution in (A.9) yields immediately:

$$\boldsymbol{P}_{T+K|T} = \bar{\Phi}^{K-1} \boldsymbol{P}_{T+1|T} (\bar{\Phi}^{K-1})^T + \sum_{i=0}^{K-2} \bar{\Phi}^i \bar{\boldsymbol{E}}\bar{\boldsymbol{Q}}\bar{\boldsymbol{E}}^T (\bar{\Phi}^i)^T \tag{A.10}$$

see, e.g., Anderson & Moore (1979).

On the other hand, in the low-frequency model (2.4)-(2.14) the *k*-step ahead forecast error variance is:

$$var(\boldsymbol{e}_K^A | \Omega_T^A) = \boldsymbol{J}^A \bar{\boldsymbol{H}} \boldsymbol{P}_{T+K\text{-}1|T} \bar{\boldsymbol{H}}^T \left(\boldsymbol{J}^A\right)^T + \boldsymbol{J}^A \bar{\boldsymbol{C}}\bar{\boldsymbol{Q}}\bar{\boldsymbol{C}}^T \left(\boldsymbol{J}^A\right)^T \tag{A.11}$$

where Ω_T^A denotes the annual information available at time *T*.

Expressions (A.7) and (A.11) have the same mathematical structure. Then, any difference between both variances will be due to the covariances $\boldsymbol{P}_{T+K\text{-}1|T}$. Assuming that the initial condition for the propagation of $\boldsymbol{P}$ in both cases is the same, $\boldsymbol{P}_1$, we obtain for the quarterly model (2.4)-(2.5):

$$\boldsymbol{P}_{2|1} = \bar{\Phi}\, \boldsymbol{P}_1 \bar{\Phi}^T + \bar{\boldsymbol{E}}\bar{\boldsymbol{Q}}\bar{\boldsymbol{E}}^T - \boldsymbol{K}_1 \boldsymbol{B}_1 \boldsymbol{K}_1^T \tag{A.12}$$

and for the annual model (2.4) and (2.14):

$$\boldsymbol{P}_{2|1}^A = \bar{\Phi}\, \boldsymbol{P}_1 \bar{\Phi}^T + \bar{\boldsymbol{E}}\bar{\boldsymbol{Q}}\bar{\boldsymbol{E}}^T - \boldsymbol{K}_1 \boldsymbol{B}_1 \left(\boldsymbol{J}^A\right)^T \left[\boldsymbol{J}^A \boldsymbol{B}_1 \left(\boldsymbol{J}^A\right)^T\right]^{-1} \boldsymbol{J}^A \boldsymbol{B}_1 \boldsymbol{K}_1^T \tag{A.13}$$

where $\boldsymbol{K}_1$ is the Kalman filter gain in *t*=1. Therefore, from (A.12) and (A.13):

$$\boldsymbol{P}_{2|1}^{\mathrm{A}} - \boldsymbol{P}_{2|1} = \boldsymbol{K}_1 \boldsymbol{B}_1 \left\{ \boldsymbol{B}_1^{-1} - \left(\boldsymbol{J}^A\right)^T \left[\boldsymbol{J}^A \boldsymbol{B}_1 \left(\boldsymbol{J}^A\right)^T\right]^{-1} \boldsymbol{J}^A \right\} \boldsymbol{B}_1 \boldsymbol{K}_1^T \tag{A.14}$$

and, using Result 1, it is immediate to see that $\boldsymbol{P}_{2|1}^{\mathrm{A}} - \boldsymbol{P}_{2|1} \geq 0$ because

$\boldsymbol{B}_1^{-1} - \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}} \left[\boldsymbol{J}^A \boldsymbol{B}_1 \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}}\right]^{-1} \boldsymbol{J}^A \geq 0$. Therefore, for t=2:

$$\boldsymbol{P}_{3|2}^{\mathbf{A}} \geq \bar{\Phi}\, \boldsymbol{P}_{2|1} \bar{\Phi}^{\boldsymbol{T}} + \bar{\boldsymbol{E}}\bar{\boldsymbol{Q}}\bar{\boldsymbol{E}}^{\boldsymbol{T}} - \boldsymbol{K}_2 \boldsymbol{B}_2 \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}} \left[\boldsymbol{J}^A \boldsymbol{B}_2 \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}}\right]^{-1} \boldsymbol{J}^A \boldsymbol{B}_2 \boldsymbol{K}_2^{\boldsymbol{T}} \tag{A.15}$$

and Result 2 assures that: $\boldsymbol{P}_{3|2}^{\mathbf{A}} - \boldsymbol{P}_{3|2} = \boldsymbol{K}_2 \boldsymbol{B}_2 \left\{\boldsymbol{B}_2^{-1} - \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}} \left[\boldsymbol{J}^A \boldsymbol{B}_2 \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}}\right]^{-1} \boldsymbol{J}^A\right\} \boldsymbol{B}_2 \boldsymbol{K}_2^{\boldsymbol{T}} \geq 0$ Therefore, by induction we obtain:

$$\boldsymbol{P}_{t|t\text{-}1}^{\mathbf{A}} - \boldsymbol{P}_{t|t\text{-}1} = \boldsymbol{K}_{t\text{-}1} \boldsymbol{B}_{t\text{-}1} \left\{\boldsymbol{B}_{t\text{-}1}^{-1} - \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}} \left[\boldsymbol{J}^A \boldsymbol{B}_{t\text{-}1} \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}}\right]^{-1} \boldsymbol{J}^A\right\} \boldsymbol{B}_{t\text{-}1} \boldsymbol{K}_{t\text{-}1}^{\boldsymbol{T}} \geq 0 \tag{A.16}$$

and, as the Riccati equation of the Kalman filter converges to its steady-state solution and it holds that:

$$\boldsymbol{P}_E^A - \boldsymbol{P}_E \geq \boldsymbol{K}_E \boldsymbol{B}_E \left\{\boldsymbol{B}_E^{-1} - \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}} \left[\boldsymbol{J}^A \boldsymbol{B}_E \left(\boldsymbol{J}^A\right)^{\boldsymbol{T}}\right]^{-1} \boldsymbol{J}^A\right\} \boldsymbol{B}_E \boldsymbol{K}_E^{\boldsymbol{T}} \geq 0 \tag{A.17}$$

■

Case 2) Consider the partially aggregated model (2.4)-(2.15). Its k-step ahead forecast error variance is given by:

$$var(\boldsymbol{e}_K^P \left| \Omega_T^P \right.) = \boldsymbol{H}^P \boldsymbol{P}_{T+K\text{-}1|T} (\boldsymbol{H}^P)^{\boldsymbol{T}} + \boldsymbol{C}^P \bar{\boldsymbol{Q}} (\boldsymbol{C}^P)^{\boldsymbol{T}} \tag{A.18}$$

where Ω_T^P denotes the partially aggregated information available at time T. As (a) the partially aggregated model and the annual model share the same state equation and (b) the underlying quarterly model is the same, Expression (A.11) is also valid in this case. Substituting $\boldsymbol{J}^A = \boldsymbol{J}^* \boldsymbol{J}^P$ in (A.11) yields:

$$\begin{aligned} var(\boldsymbol{e}_K^A \left| \Omega_T^P \right.) &= \boldsymbol{J}^* \boldsymbol{J}^P \bar{\boldsymbol{H}} \boldsymbol{P}_{T+K\text{-}1|T} \bar{\boldsymbol{H}}^{\boldsymbol{T}} (\boldsymbol{J}^* \boldsymbol{J}^P)^{\boldsymbol{T}} + \boldsymbol{J}^* \boldsymbol{J}^P \bar{\boldsymbol{C}} \bar{\boldsymbol{Q}} \bar{\boldsymbol{C}}^{\boldsymbol{T}} (\boldsymbol{J}^* \boldsymbol{J}^P)^{\boldsymbol{T}} \\ &= \boldsymbol{J}^* \boldsymbol{H}^P \boldsymbol{P}_{T+K\text{-}1|T} (\boldsymbol{H}^P)^{\boldsymbol{T}} (\boldsymbol{J}^*)^{\boldsymbol{T}} + \boldsymbol{J}^* \boldsymbol{C}^P \bar{\boldsymbol{Q}} (\boldsymbol{C}^P)^{\boldsymbol{T}} (\boldsymbol{J}^*)^{\boldsymbol{T}} \end{aligned} \tag{A.19}$$

and, consequently:

$$\boldsymbol{P}_E^A - \boldsymbol{P}_E^P \geq \boldsymbol{K}_E\, \boldsymbol{B}_E \left\{\boldsymbol{B}_E^{-1} - (\boldsymbol{J}^*)^{\boldsymbol{T}} \left[\boldsymbol{J}^* \boldsymbol{B}_E (\boldsymbol{J}^*)^{\boldsymbol{T}}\right]^{-1} \boldsymbol{J}^*\right\} \boldsymbol{B}_E \boldsymbol{K}_E^{\boldsymbol{T}} \geq 0 \tag{A.20}$$

■

APPENDIX B: PROOF OF THEOREM 1

According to definitions 1) and 2), the annual model (2.4) and (2.14) has unobservable modes if there exists at least a vector $\boldsymbol{w} \neq \boldsymbol{0}$ such that $\boldsymbol{\Phi}^{S}\boldsymbol{w} = \lambda \boldsymbol{w}$ and:

$$\boldsymbol{H}\sum_{i=0}^{S\text{-}1}\boldsymbol{\Phi}^{i}\ \boldsymbol{w} = \boldsymbol{0} \tag{B.1}$$

B.1 Proof of Case 1).

Expression (B.1) implies:

$$\boldsymbol{Hw} + \boldsymbol{H\Phi w} + \boldsymbol{H\Phi}^{2}\boldsymbol{w} + \ldots \boldsymbol{H\Phi}^{S-1}\boldsymbol{w} = \boldsymbol{0} \tag{B.2}$$

If $\boldsymbol{w}$ is an eigenvector of $\boldsymbol{\Phi}$, then it is also an eigenvector of the powers of $\boldsymbol{\Phi}$, so $\boldsymbol{\Phi}^{i}\boldsymbol{w} = \lambda^{i}\boldsymbol{w}$ $(i = 1, 2, \ldots)$ and, therefore:

$$\boldsymbol{Hw} + \boldsymbol{H}\lambda\boldsymbol{w} + \boldsymbol{H}\lambda^{2}\boldsymbol{w} + \ldots \boldsymbol{H}\lambda^{S\text{-}1}\boldsymbol{w} = \boldsymbol{Hw}\left[1 + \lambda + \lambda^{2} + \ldots + \lambda^{S\text{-}1}\right] = \boldsymbol{0} \tag{B.3}$$

but, as the quarterly model is assumed to be observable, then $\boldsymbol{Hw} \neq \boldsymbol{0}$ and (B.3) holds i.i.f. $1 + \lambda + \lambda^{2} + \ldots + \lambda^{S\text{-}1} = 0$ ■

B.2 Proof of Case 2).

Under the conditions of Case 2), the k eigenvectors associated to $\lambda_k(\boldsymbol{\Phi})$ generate a subspace, denoted by $\boldsymbol{S}_{\lambda,k}$, such that for any $\boldsymbol{w} \in \boldsymbol{S}_{\lambda,k}$, it holds that $\boldsymbol{\Phi}^{S}\boldsymbol{w} = \lambda\boldsymbol{w}$, where $\lambda = \lambda_i^{S}, \forall \lambda_i \in \lambda_k(\boldsymbol{\Phi})$.

Let $\boldsymbol{S}_{I}$ be the intersection between $\boldsymbol{S}_{\lambda,k}$ and the null space of matrix $\boldsymbol{H}$, then $dim(\boldsymbol{S}_{I}) = dim\left[\boldsymbol{S}_{\lambda,k} \cap null(\boldsymbol{H})\right] = k - rank(\boldsymbol{HW})$, where $\boldsymbol{W}$ is a matrix which columns are the eigenvectors spanning $\boldsymbol{S}_{\lambda,k}$. In the univariate case, the number of unobservable modes is exactly $dim(\boldsymbol{S}_{I}) = k - rank(\boldsymbol{HW}) = k - 1$, because $\boldsymbol{H}$ is a row matrix and the quarterly model is observable. In the multivariate case there is no exact rule but, in general, $k - rank(\boldsymbol{HW})$ modes become unobservable and the maximum number of unobservable modes is $k - 1$. ■

B.3 Proof of Case 3).

Under the conditions of Case 3), it is easy to see that the S-th power of any k-dimension Jordan block with null eigenvalues breaks into several blocks. The dimension of the larger sub-block will then be $\left[1+(k-1)/S\right]$, where "[]" denotes the integer part of a real argument.

Fragmentation of the Jordan block implies an increase in the geometric multiplicity associated to the null eigenvalues, so there are several linearly independent eigenvectors associated to each null eigenvalue. This situation is therefore similar to the one considered in Case 2). ■

APPENDIX C: PROOF OF THEOREM 2

Consider the matrices in (2.1)-(2.2) and (2.4)-(2.5), denoted by $\boldsymbol{R}_1 = (\Phi, \Gamma, \boldsymbol{E}, \boldsymbol{H}, \boldsymbol{D}, \boldsymbol{Q})$ and $\boldsymbol{R}_2 = (\bar{\Phi}, \bar{\Gamma}, \bar{\boldsymbol{E}}, \bar{\boldsymbol{H}}, \bar{\boldsymbol{D}}, \bar{\boldsymbol{C}}, \bar{\boldsymbol{Q}})$. If (2.1)-(2.2) is minimal, then there exists a biunivocal correspondence between both representations, such that $\boldsymbol{R}_2 = F(\boldsymbol{R}_1)$, being *F*() a biyective application.

Assume that this is not true. Then there are at least two realizations, $\boldsymbol{R}_1$ and $\boldsymbol{R}_1^*$, with the same canonical representation, such that $\boldsymbol{R}_2 = F(\boldsymbol{R}_1) = F(\boldsymbol{R}_1^*)$. Then $\boldsymbol{R}_1$ and $\boldsymbol{R}_1^*$ are output-equivalent. However, if $\boldsymbol{R}_1$ is minimal, then $\boldsymbol{R}_1^*$ can only be a similar transformation of $\boldsymbol{R}_1$. As $\boldsymbol{R}_1$ and $\boldsymbol{R}_1^*$ have the same canonical representation, then the only similar transformation is identity and $\boldsymbol{R}_1 = \boldsymbol{R}_1^*$.

Also, there exists a biunivocal correspondence between $\boldsymbol{R}_1$ and the matrices characterizing the annual model (2.4)-(2.14), denoted by $\boldsymbol{R}_3 = (\bar{\Phi}, \bar{\Gamma}, \bar{\boldsymbol{E}}, \boldsymbol{H}^A, \boldsymbol{D}^A, \boldsymbol{C}^A, \bar{\boldsymbol{Q}})$, such that $\boldsymbol{R}_3 = G(\boldsymbol{R}_1)$, being *G*() a biyective application.

Assume again that this is not true. Then there would be at least two quarterly realizations, $\boldsymbol{R}_1$ and $\boldsymbol{R}_1^*$, with the same canonical representation, such that $\boldsymbol{R}_3 = G(\boldsymbol{R}_1) = G(\boldsymbol{R}_1^*)$. However we know that $F(\boldsymbol{R}_1) \neq F(\boldsymbol{R}_1^*)$ and, as $F(\boldsymbol{R}_1)$ and $G(\boldsymbol{R}_1)$ share the same state equation, then $F(\boldsymbol{R}_1)$ and $F(\boldsymbol{R}_1^*)$ should have different quarterly observation equations that yield the same annual observation equation. Then there are two quarterly models with the same annual realization, so: $\boldsymbol{R}_1 = (\Phi, \Gamma, \boldsymbol{E}, \boldsymbol{H}, \boldsymbol{D}, \boldsymbol{Q})$ and $\boldsymbol{R}_1^* = (\Phi, \Gamma, \boldsymbol{E}, \boldsymbol{H}^*, \boldsymbol{D}^*, \boldsymbol{Q})$, such that: $F(\boldsymbol{R}_1) = (\bar{\Phi}, \bar{\Gamma}, \bar{\boldsymbol{E}}, \bar{\boldsymbol{H}}, \bar{\boldsymbol{D}}, \bar{\boldsymbol{C}}, \bar{\boldsymbol{Q}})$, $F(\boldsymbol{R}_1^*) = (\bar{\Phi}, \bar{\Gamma}, \bar{\boldsymbol{E}}, \bar{\boldsymbol{H}}^*, \bar{\boldsymbol{D}}^*, \bar{\boldsymbol{C}}^*, \bar{\boldsymbol{Q}})$ and $\boldsymbol{R}_3 = G(\boldsymbol{R}_1) = G(\boldsymbol{R}_1^*) = (\bar{\Phi}, \bar{\Gamma}, \bar{\boldsymbol{E}}, \boldsymbol{H}^A, \boldsymbol{D}^A, \boldsymbol{C}^A, \bar{\boldsymbol{Q}})$, where:

$$\boldsymbol{H}^A = \boldsymbol{J}^A \bar{\boldsymbol{H}} = \boldsymbol{J}^A \bar{\boldsymbol{H}}^* \tag{C.1}$$

$$\boldsymbol{C}^A = \boldsymbol{J}^A \bar{\boldsymbol{C}} = \boldsymbol{J}^A \bar{\boldsymbol{C}}^* \tag{C.2}$$

$$\boldsymbol{D}^A = \boldsymbol{J}^A \bar{\boldsymbol{D}} = \boldsymbol{J}^A \bar{\boldsymbol{D}}^* \tag{C.3}$$

and, assuming that the variables are flows, $\boldsymbol{J}^A = [\boldsymbol{I}, \boldsymbol{I}, \ldots, \boldsymbol{I}]$.

Condition (C.1) implies that $\boldsymbol{H}^A = \boldsymbol{H}\sum_{i=0}^{S\text{-}1}\Phi^i = \boldsymbol{H}^*\sum_{i=0}^{S\text{-}1}\Phi^i$, so $(\boldsymbol{H} - \boldsymbol{H}^*)\sum_{i=0}^{S\text{-}1}\Phi^i = \boldsymbol{0}$. This implies that, if $\bar{\boldsymbol{H}} \neq \bar{\boldsymbol{H}}^*$, then $\sum_{i=0}^{S\text{-}1}\Phi^i$ is a rank-deficient matrix and there exists an eigenvalue of Φ, λ, such that $\sum_{i=0}^{S\text{-}1}\lambda\Phi^i = \boldsymbol{0}$. Therefore a contradiction arises as the annual model would not be observable (Theorem 1, Case 1).

Finally, conditions (C.2) and (C.3) are now easy to prove. As $\boldsymbol{R}_1$ and $\boldsymbol{R}_1^*$ share the matrices $\Phi, \Gamma, \boldsymbol{E}, \boldsymbol{Q}$ and *H*, see (C.1), then $\boldsymbol{C} = \boldsymbol{C}^* = \boldsymbol{I}$ and $\boldsymbol{D} = \boldsymbol{D}^*$, see (2.7). ■

APPENDIX D: PROOF OF PROPOSITION 4

D.1. Previous results

Result 1 (observability staircase form). This Result, due to Kalman (1963), states that for any SS model, characterized by the pair $(\boldsymbol{H}, \Phi)$, there exists a similar transformation $\boldsymbol{T}$, such that $\boldsymbol{H}^* = \boldsymbol{H}\boldsymbol{T}^{-1}$, $\Phi^* = \boldsymbol{T}\Phi\boldsymbol{T}^{-1}$, that results in a model $(\boldsymbol{H}^*, \Phi^*)$ with the following structure:

$$\Phi^* = \begin{bmatrix} \Phi_{11}^N & \Phi_{12}^N & 0 & 0 \\ 0 & \Phi_{22}^N & 0 & 0 \\ 0 & 0 & \Phi_{11}^E & \Phi_{12}^E \\ 0 & 0 & 0 & \Phi_{22}^E \end{bmatrix} \tag{D.1}$$

$$\boldsymbol{H}^* = \begin{bmatrix} 0 & \boldsymbol{H}_2^N & 0 & \boldsymbol{H}_2^E \end{bmatrix} \tag{D.2}$$

where:

1) $\Phi^N = \begin{bmatrix} \Phi_{11}^N & \Phi_{12}^N \\ 0 & \Phi_{22}^N \end{bmatrix}$, $\Phi^E = \begin{bmatrix} \Phi_{11}^E & \Phi_{12}^E \\ 0 & \Phi_{22}^E \end{bmatrix}$ characterize, respectively, the nonstationary and stationary subsystems, so: $\lambda(\Phi^N) \geq 1$, $\lambda(\Phi^E) < 1$, where $\lambda()$ denotes any eigenvalue of the corresponding argument matrix.

2) $\begin{bmatrix} \Phi_{22}^N & 0 \\ 0 & \Phi_{22}^E \end{bmatrix}$ and $\begin{bmatrix} \boldsymbol{H}_2^N & \boldsymbol{H}_2^E \end{bmatrix}$ characterize the observable subsystem.

3) Φ_{11}^N corresponds to the states associated with non-detectable modes.

4) Φ_{11}^E corresponds to the states associated with detectable, but unobservable, modes.

Result 2 (variance of smoothed estimates when the system has unit roots). Casals, Jerez & Sotoca (2000) show that for any SS model with unit eigenvalues, corresponding to a time series with unit AR roots, the exact covariance of fixed-interval smoothed estimates of the states is:

$$\boldsymbol{P}_{t|N} = \boldsymbol{P}_{t|N}^* + \boldsymbol{V}_t \boldsymbol{P}_{1|N} \boldsymbol{V}_t^T \tag{D.3}$$

where:

$$P_{1|N} = (P_1 + S)^{-1},\ P_1 = \mathrm{cov}(x_1) \tag{D.4}$$

$$S = \sum_{t=1}^{N} \Phi_t H B_t^{-1} H^T \Phi_t^T \tag{D.5}$$

$$\Phi_t = (\Phi - K_t H)\Phi_{t-1},\ \Phi_1 = I \tag{D.6}$$

$$V_t = (I - P_{t|t-1} R_{t-1})\Phi_t \tag{D.7}$$

$$R_{t-1} = H^T B_t^{-1} H + (\Phi - K_t H)^T R_t (\Phi - K_t H) \tag{D.8}$$

being $P_{t|N}^{*}$ the smoother covariance derived from a Kalman filter with null initial conditions.

Application of Result 2 to a non-stationary subsystem requires the initial state covariance given in (D.4), P_1, to show unbounded uncertainty. Therefore (De Jong, 1991; Ansley & Kohn, 1989) its expression must be:

$$P_1 = \mathrm{diag}(k\Pi, N),\ k \gg 0,\ \Pi > 0 \text{ and } N \text{ the solution of } N = \Phi^S N (\Phi^S)^T + Q^S \tag{D.9}$$

D.2. Proof of Proposition 4

Using Result 1, the aggregated models (2.4) and (2.14), or (2.4) and (2.15), can be written in observability staircase form. If this form has a non-stationary subsystem, the smoother must be initialized according to (D.9) and the matrix S, see (D.5), has the following structure:

$$S = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & S_2^{NN} & 0 & S_2^{NE} \\ 0 & 0 & 0 & 0 \\ 0 & S_2^{EN} & 0 & S_2^{EE} \end{bmatrix} \tag{D.10}$$

where $\begin{bmatrix} S_2^{NN} & S_2^{NE} \\ S_2^{EN} & S_2^{EE} \end{bmatrix}$ is positive-definite. Denoting $R = \Pi^{-1}$, $V = N^{-1}$ and partitioning Π, R, N and V as:

$$\Pi = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix};\ R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix};\ N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix};\ V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \tag{D.11}$$

the matrix $P_{1|N}$, see (D.4), can be written in the following block-form:

$$P_{1|N} = \begin{bmatrix} \frac{1}{k}R_{11} & \frac{1}{k}R_{12} & 0 & 0 \\ \frac{1}{k}R_{21} & \frac{1}{k}R_{22} + S_2^{NN} & 0 & S_2^{NE} \\ 0 & 0 & V_{11} & V_{12} \\ 0 & S_2^{EN} & V_{21} & V_{22} + S_2^{EE} \end{bmatrix}^{-1} \tag{D.12}$$

with $k \gg 0$. Applying the partitioned-matrix inversion lemma to (D.12) yields:

$$P_{1|N} = \begin{bmatrix} P_{1|N}^{NN} & P_{1|N}^{NE} \\ P_{1|N}^{EN} & P_{1|N}^{EE} \end{bmatrix} \tag{D.13}$$

where:

$$P_{1|N}^{NN} = \begin{bmatrix} \left[\frac{1}{\mathrm{k}}R_{11} - \frac{1}{\mathrm{k}^2}R_{12}\left(\frac{R_{22}}{\mathrm{k}} + W\right)^{-1} R_{21}\right]^{-1} & -R_{11}^{-1}R_{12}\left(\frac{\Pi_{22}^{-1}}{\mathrm{k}} + W\right)^{-1} \\ = & \left(\frac{\Pi_{22}^{-1}}{\mathrm{k}} + W\right)^{-1} \end{bmatrix} \tag{D.14}$$

$$P_{1|N}^{EE} = \begin{bmatrix} \left[V_{11} - V_{12}\left(V_{22} + Y\right)^{-1} V_{21}\right]^{-1} & -V_{11}^{-1}V_{12}\left(N_{22}^{-1} + Y\right)^{-1} \\ = & \left(N_{22}^{-1} + Y\right)^{-1} \end{bmatrix} \tag{D.15}$$

$$P_{1|N}^{EN} = \begin{bmatrix} V_{11}^{-1}V_{12}\left(N_{22}^{-1} + S_2^{EE}\right)^{-1} S_2^{EN}\left(\frac{\Pi_{22}^{-1}}{\mathrm{k}} + W\right)^{-1} R_{21}R_{11}^{-1} & -V_{11}^{-1}V_{12}\left(N_{22}^{-1} + S_2^{EE}\right)^{-1} S_2^{EN}\left(\frac{\Pi_{22}^{-1}}{\mathrm{k}} + W\right)^{-1} \\ \left(N_{22}^{-1} + S_2^{EE}\right)^{-1} S_2^{EN}\left(\frac{\Pi_{22}^{-1}}{\mathrm{k}} + W\right)^{-1} R_{21}R_{11}^{-1} & \left(N_{22}^{-1} + S_2^{EE}\right)^{-1} S_2^{EN}\left(\frac{\Pi_{22}^{-1}}{\mathrm{k}} + W\right)^{-1} \end{bmatrix} \tag{D.16}$$

being $W = S_2^{NN} - S_2^{NE}(V_{22} + S_2^{SS})^{-1} S_2^{EN}$ and $Y = S_2^{EE} - S_2^{EN}\left(\frac{R_{22}}{k} + S_2^{NN}\right)^{-1} S_2^{NE}$

In these conditions the limit values of the blocks in (D.13), as *k* tends to infinity, are:

$$\lim_{k\to\infty} P_{1|N}^{NN} = \begin{bmatrix} kR_{11}^{-1} & -R_{11}^{-1}R_{12}W^{-1} \\ = & W^{-1} \end{bmatrix} \tag{D.17}$$

$$\lim_{k\to\infty} \boldsymbol{P}_{1|N}^{EE} = \begin{bmatrix} \left[\boldsymbol{V}_{11} - \boldsymbol{V}_{12}(\boldsymbol{V}_{22} + \boldsymbol{Y})^{-1}\boldsymbol{V}_{21}\right]^{-1} & -\boldsymbol{V}_{11}^{-1}\boldsymbol{V}_{12}(\boldsymbol{N}_{22}^{-1} + \boldsymbol{Y})^{-1} \\ = & (\boldsymbol{N}_{22}^{-1} + \boldsymbol{Y})^{-1} \end{bmatrix} \tag{D.18}$$

$$\lim_{k\to\infty} \boldsymbol{P}_{1|N}^{EN} = \begin{bmatrix} \boldsymbol{V}_{11}^{-1}\boldsymbol{V}_{12}(\boldsymbol{N}_{22}^{-1}\boldsymbol{S}_2^{EE})^{-1}\boldsymbol{S}_2^{EN}\boldsymbol{W}^{-1}\boldsymbol{R}_{21}\boldsymbol{R}_{11}^{-1} & -\boldsymbol{V}_{11}^{-1}\boldsymbol{V}_{12}(\boldsymbol{N}_{22}^{-1} + \boldsymbol{S}_2^{EE})^{-1}\boldsymbol{S}_2^{EN}\boldsymbol{W}^{-1} \\ (\boldsymbol{N}_{22}^{-1} + \boldsymbol{S}_2^{EE})^{-1}\boldsymbol{S}_2^{EN}\boldsymbol{W}^{-1}\boldsymbol{R}_{21}\boldsymbol{R}_{11}^{-1} & (\boldsymbol{N}_{22}^{-1} + \boldsymbol{S}_2^{EE})^{-1}\boldsymbol{S}_2^{EN}\boldsymbol{W}^{-1} \end{bmatrix} \tag{D.19}$$

Therefore, the blocks (1,2), (2,1) and (2,2) in (D.13) converge to finite values. Only the (1,1) block diverge to infinity but this is irrelevant to Proposition 4, as it corresponds to the non-detectable modes, see Result 1. ■

Note that the initial state has a persistent effect over the smoothed estimates of the states because the (1,1) block of $\boldsymbol{P}_{t|N}$, see (D.3), is:

$$\boldsymbol{P}_{t|N}^{(1,1)} = \boldsymbol{P}_{t|N}^{*(1,1)} + (\Phi_{11}^{N})^{t-1}\boldsymbol{P}_{1|N}^{(1,1)}(\Phi_{11}^{N})^{t-1} + \ldots \tag{D.20}$$

taking into account (D.7)-(D.8) and the fact that when the system is in observability staircase form the matrix $\boldsymbol{R}_t$ is:

$$\boldsymbol{R}_t = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \boldsymbol{R}_{t2}^{NN} & 0 & \boldsymbol{R}_{t2}^{NE} \\ 0 & 0 & 0 & 0 \\ 0 & \boldsymbol{R}_{t2}^{EN} & 0 & \boldsymbol{R}_{t2}^{EE} \end{bmatrix} \tag{D.21}$$

As $\lambda(\Phi_{11}^{N}) \geq 1$, it is immediate to see that initial conditions will affect all the sequence of smoothed estimates, no matter the sample size. Therefore, the infinite variance of a diffuse prior would be propagated to the smoothed estimates along the whole sample.