A Model Selection Approach to Detect Seasonal Unit Roots

Yoshinori Kawasaki *

The Institute of Statistical Mathematics

4-6-7 Minami-Azabu Minato-ku, Tokyo 106, Japan

and

Philip Hans Franses

Econometric Institute and Rotterdam Institute for Business

Economic Studies, Erasmus University Rotterdam

Burg. Oudlaan 50, 3062 PA Rotterdam, The Netherlands

December 9, 1996

Abstract

The popular 'airline' model for a seasonal time series assumes that a variable needs double differencing, i.e. first and seasonal (or annual) differencing. The

*Corresponding author. (kawasaki@ism.ac.jp)

This paper was written while the first author visited Tinbergen Institute, Erasmus University Rotterdam. Parts of a previous version of the paper were presented in seminars at CORE in Louvain-la-Neuve and at the Institute for Advanced Studies in Vienna.

resultant time series can usually be described by a low order moving average model with estimated roots close to the unit circle. This latter feature complicates the standard autoregression-based tests for (seasonal) unit roots which are often used in practice. In this paper we propose an alternative route to detect seasonal unit roots by analyzing (versions of) the basic structural model [BSM]. This BSM can generate data which are (approximately) observationally equivalent to data generated from an airline model. Using Monte Carlo simulations, we show that our method works very well. We illustrate our approach for a large set of macroeconomic time series variables.

Keywords: Seasonal unit roots, Overdifferencing, Information criterion, Structural time series model

1 Introduction and motivation

Since its introduction by Box and Jenkins (1970), the so-called airline model has been extensively used in modelling and forecasting seasonal time series. For a quarterly time series $\{y_t\}$ $(t = 1, 2, \ldots T)$ this airline model is

$$(1 - L)(1 - L^4)y_t = (1 - \theta_1 L)(1 - \theta_4 L^4)\epsilon_t, \tag{1}$$

where L is the lag operator defined by $L^k y_t = y_{t-k}$, where $\{\epsilon_t\}$ is a standard white noise process with variance σ_ϵ^2, and θ_1 and θ_4 are unknown parameters to be estimated. One of the reasons for its empirical success is that, except for σ_ϵ^2, (1) contains only two parameters. As the polynomial on the left hand side can be decomposed as

$$(1 - L)(1 - L^4) = (1 - L)(1 - L)(1 + L)(1 + iL)(1 - iL), \tag{2}$$

it is clear that when $|\theta_1| < 1$ and $|\theta_4| < 1$, the airline model assumes that y_t has 5 roots on the unit circle, i.e. two nonseasonal unit roots 1, and three so-called seasonal unit roots -1 and $\pm i$. In many practical applications, however, the estimated values for θ_1 and θ_4 (which are known to be bounded away from unity) take values which are very close to 1. Hence for these time series it may seem that either the $1 - L$ or the $1 - L^4$ filter, or maybe even the $(1 - L^4)/(1 - L) = 1 + L + L^2 + L^3 \equiv S(L)$ filter, is redundant.

Standard tests for seasonal unit roots as those proposed in Hylleberg et al. (1990) [HEGY] appear not to be very useful to investigate seasonal overdifferencing in the airline model. In fact, the simulation results in Ghysels et al. (1994) convincingly show that the HEGY tests have severe size distortions in case (1) is the data generating process and θ_1 and θ_4 take values close to the unit circle boundary. A cause for these distortions is that the HEGY method assumes that the y_t series can be described by an autoregressive model of some moderate order.

In this paper we aim to overcome these difficulties caused by substantial moving average components by investigating seasonal unit roots in a model that can have close empirical resemblance with the airline model. In fact, the basic structural model [BSM] we analyze can be shown to have the airline model as a special case, see also Harvey (1984). Our investigation relies on the application of information criterion statistics like AIC (Akaike (1973)) and BIC (Schwarz (1978) and Rissanen (1978)). Although our approach can be extended in various ways, we choose to focus on the seasonal unit roots -1 and $\pm i$ in a quarterly time series, also since this appears to be a relevant practical problem.

Our simulation results in section 4 show that, when translating our outcomes in terms of size and power, our method has good size and power properties. This is in contrast with the recent approach in Psaradakis (1996) which aims to test for unit roots in time series with nearly deterministic seasonal variation using the prewhitening techniques proposed in Maekawa (1994). In fact, the empirical size of these novel tests are close to the nominal size, but the resultant tests appear to have virtually no power. Our simulations in section 4 however show that we can make proper decisions even when the alternative models are very close to the null model.

The outline of this paper is as follows. In section 2, we highlight some features of the airline model and the BSM. In section 3, we present an outline of our model selection approach. In section 4, we evaluate our method using Monte Carlo simulations, and based on these results we formulate simple decision rules. In section 5, we apply our approach to a set of 22 macroeconomic time series variables and we compare the results with those obtained using the HEGY method in various other studies. In section 6, we conclude our paper with some remarks.

2 Models

In this section we compare the airline model in (1) and the BSM with respect to their autocorrelation properties in order to show that the BSM extends the airline model in various directions. For the airline model in (1), it is easy to show that the autocorrelations at lag k, denoted as ρ_k, are

$$\rho_1 = -\theta_1/(1+\theta_1^2) \tag{3}$$

$$\rho_2 = 0 \tag{4}$$

$$\rho_3 = \rho_5 = \theta_1\theta_4/\{(1+\theta_1^2)(1+\theta_4^2)\} \tag{5}$$

$$\rho_4 = -\theta_4/(1+\theta_4^2) \tag{6}$$

$$\rho_j = 0 \quad \text{for} \quad j = 6, 7, \ldots \tag{7}$$

When $\theta_1 = \theta_4 = 1$, i.e., when the double differencing filter $(1-L)(1-L^4)$ amounts to overdifferencing, the nonzero ρ_k's take the values $\rho_1 = \rho_4 = -\frac{1}{2}$ and $\rho_3 = \rho_5 = \frac{1}{4}$ so that $\sum_{k=1}^{\infty} \rho_k = -\frac{1}{2}$.

The BSM for y_t is given by a set of three equations, i.e.,

$$y_t = \mu_t + s_t + w_t, \quad w_t \sim \text{NID}(0, \sigma_w^2) \tag{8}$$

$$(1-L)^2\mu_t = u_t, \quad u_t \sim \text{NID}(0, \sigma_\mu^2) \tag{9}$$

$$(1+L+L^2+L^3)s_t = v_t, \quad v_t \sim \text{NID}(0, \sigma_s^2) \tag{10}$$

where the error processes w_t, u_t and v_t are also mutually independent. This approach to modelling nonstationary trending and seasonal time series became popular through the work of Kitagawa (1981) and Harvey (1985), inter alia. The basic structural model seems to be one of the most widely accepted terms to refer to the set of equations (8) – (10). Although (9) is a second order random walk instead of a local linear trend in Harvey (1989, p.172), we still refer to (8) – (10) as the BSM.

Substituting (9) and (10) in (8) yields that y_t can be described by

$$(1-L)(1-L^4)y_t = q_t \tag{11}$$

where q_t is a moving average process of order 5 [MA(5)]. The autocovariances γ_k, $k = 0, 1, 2, \ldots$, for q_t are

$$\gamma_0 = 4\sigma_\mu^2 + 6\sigma_s^2 + 4\sigma_w^2 \tag{12}$$

$$
\begin{aligned}
\gamma_1 &= 3\sigma_\mu^2 - 4\sigma_s^2 - 2\sigma_w^2 && (13)\\
\gamma_2 &= 2\sigma_\mu^2 + \sigma_s^2 && (14)\\
\gamma_3 &= \sigma_\mu^2 + \sigma_w^2 && (15)\\
\gamma_4 &= -2\sigma_w^2 && (16)\\
\gamma_5 &= \sigma_w^2 && (17)\\
\gamma_j &= 0 \quad \text{for } j = 6, 7, \ldots. && (18)
\end{aligned}
$$

When $\sigma_\mu^2 = 0$ and $\sigma_s^2 = 0$, it is easy to see that (11) reduces to the airline model with $\theta_1 = 1$ and $\theta_4 = 1$. Strictly speaking, when $\sigma_\mu^2 \neq 0$, the q_t process in (11) is invertible. This can be understood from the fact that for q_t in (11) holds that

$$\sum_{k=1}^{\infty} \rho_k = -\frac{1}{2} + \frac{8\sigma_\mu^2}{\gamma_0} \qquad (19)$$

i.e., the theoretical autocorrelations only sum to $-\frac{1}{2}$ when $\sigma_\mu^2 = 0$. One can observe from (12) – (18) that the autocorrelation function of (11) can come close to that of the airline model with θ_1 and θ_4 close to unity. Hence, the BSM is flexible enough to generate a wide range of time series data, amongst which are those that one may want to describe by an airline model.

In this paper we exploit this link between the airline model and the BSM in order to investigate the presence of seasonal unit roots when (1) seems to give a good data description. In other words, we indirectly examine whether the airline model amounts to overdifferencing at the seasonal frequencies by analyzing the BSM. Our null hypothesis is (8) – (10) and our alternative model is (8) and (9) with

$$(1 + aL)(1 + bL^2)s_t = v_t \qquad (20)$$

where $(0 <)a \neq 1$ and $(0 <)b \neq 1$. In the next section we describe our model selection approach.

3 Model Selection Approach

Throughout this paper we consider four models to be compared. They share equations (8) and (9), i.e. we assume that all models consist of three components and that the trend component μ_t and the observational noise component w_t are common. Only the seasonal components differ according to the number of assumed seasonal unit roots. In sum, we consider

$$\text{Model 0} \quad : \quad (1+aL)(1+bL^2)s_t = v_t \tag{21}$$

$$\text{Model 1} \quad : \quad (1+L)(1+bL^2)s_t = v_t \tag{22}$$

$$\text{Model 2} \quad : \quad (1+aL)(1+L^2)s_t = v_t \tag{23}$$

$$\text{Model 3} \quad : \quad (1+L)(1+L^2)s_t = v_t \quad . \tag{24}$$

In models 0,1 and 2, a and b are unknown hyperparameters, so they have to be estimated as well as σ_s^2. Strictly speaking, model 0, for example, consists of (8), (9) and (21), but there will be no confusion even if we simply refer to equation (21) as model 0 instead of mentioning the whole set of equations. Model 3 is the standard BSM as discussed in the previous section. In addition we can refer to model 0 through 2 as relaxed BSMs. This makes sense because model 3 is a restricted version of the other models.

Our basic strategy to determine the number of seasonal unit roots in y_t is as follows. For given data, we estimate all the four models (21) through (24). If we introduce unnecessary parameters (a and/or b), that will be reflected in the values of information criteria. On the other hand, if seasonal unit roots do not exist, the information criterion statistics of incorrectly restricted models (model 3 or even model 1 and 2) should be inferior. Although many variants of information criteria are proposed since Akaike (1973), the most popular criteria in practice seem to be AIC and BIC,

$$\begin{aligned} \text{AIC} &= -2\hat{\ell} + 2k \\ \text{BIC} &= -2\hat{\ell} + k \log T \end{aligned}$$

where $\hat{\ell}$ denotes the maximized log-likelihood, and k is the number of estimated unknown parameters, see Akaike (1973), and Schwarz (1978), Rissanen (1978), respectively. It is not clear a priori if we should rely on a single information criterion statistic in our decision. Hence it is worthwhile to investigate how AIC and BIC behave in relevant situations. The corresponding simulation results are reported in the next section.

The rest of this section provides a brief summary of the estimation of unknown hyperparameters and unobserved variables like $\{\mu_t\}$. Generally, it is quite common to give a state space representation to a time series model, which enables us to use recursive methods to calculate the log-likelihood function. Especially this holds for the unobserved components model or the structural time series model. If we define the state vector x_t as $x_t = (\mu_t, \mu_{t-1}, s_t, s_{t-1}, s_{t-2})$, it is easy to see that the system equation (or the transition equation, (25) below) and the observation equation (or the measurement equation, (26) below) are linear with respect to the state vector, and also that they are linear with respect to the system and observational noise. More explicitly, we have

$$x_t = Fx_{t-1} + G\tilde{v}_t \tag{25}$$

$$y_t = Hx_t + w_t \tag{26}$$

where F, G and H are defined as

$$F = \left(\begin{array}{cc|ccc} 2 & -1 & & O & \\ 1 & 0 & & & \\ \hline & & -a & -b & -ab \\ & O & 1 & 0 & 0 \\ & & 0 & 1 & 0 \end{array}\right), \quad G = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$H = (1, 0, 1, 0, 0).$$

The mean of $\tilde{v}_t$ is assumed to be zero, and the covariance matrix of $\tilde{v}_t$ is the diagonal matrix with elements σ_μ^2 and σ_s^2 by assumption, see (9) and (10). The F matrix shown above is for model 0. By restricting a and/or b to 1, we have state space forms for the other models.

When we assume the Gaussian distribution for every noise process (see (8)–(10)), the above modelling approach invariably exploits the Kalman filter. To obtain maximum likelihood estimates of the hyperparameters, the fixed interval smoother yields a rather precise distribution of the state vector, containing the unobserved components, Anderson and Moore (1979) and Harvey (1989), inter alia.

We also note that this kind of modelling can be characterized as a smoothness prior approach to 'ill-posed' problems, see Akaike (1980) and Kitagawa (1981). Stated briefly, we have $2T+1$ unknown parameters in (8) while we have only T observations. Therefore, we introduce priors which allow smooth time transitions for parameters μ_t and s_t, and we control only a few number of hyperparameters such as σ_μ^2 and σ_s^2.

We close this section with a remark on the support of the hyperparameters a and b in seasonal components of the relaxed BSMs. Upon estimation, we impose nonnegativity of a and b. If a is negative, it is no longer related with the biannual frequency but its inverse corresponds with a nonseasonal root of a seasonal polynomial. In the same way, if b is negative, $1+bB^2$ loses its connection with annual seasonal frequencies. Intuitively, it is also natural to expect the estimates of a and/or b to be less than 1 or at most 1. Although most of the empirical results in section 5 appear to meet this expectation, we did not impose such restrictions as $a<1$ or $b<1$ in estimating them.

4 Monte Carlo Simulations

In this section, after describing the Monte Carlo design, we examine several simulation results and decision rules. When both AIC and BIC select the same model by minimum AIC and BIC respectively, we select that model. If two criteria give a split decision, we show that the combination of the results from minimum AIC and BIC still seems to provide accurate decision rules.

First, we give our Monte Carlo design in detail. All data are generated by the BSM or relaxed BSMs based on their state space representations. As an initial state vector, we used the result of one empirical data analysis. After fitting the BSM to Japanese consumption data, we obtained the filter mean at the end point (automatically smoother mean) as $x_{T|T} = (0.86971882, 0.86493286, -0.03838920, -0.03977958, 0.07376490)'$. Fixing the innovation variance of the trend component to $\sigma_\mu^2 = 1.0 \times 10^{-6}$, we consider three different combinations of the seasonal innovation variance and the observational noise variance.

(A) $\sigma_s^2 = \sigma_w^2 = 1.0 \times 10^{-4}$

(B) $\sigma_s^2 = 1.0 \times 10^{-4}$ and $\sigma_w^2 = 5.0 \times 10^{-4}$

(C) $\sigma_s^2 = 1.0 \times 10^{-4}$ and $\sigma_w^2 = 1.0 \times 10^{-5}$.

Notice that with these values of the variances, the simulated data can easily be described by an airline model. Let λ be the trade-off parameter (see Gersch and Kitagawa (1988) for example) defined by $\lambda = \sigma_s^2/\sigma_w^2$. Then, we have (A) $\lambda = 1$, (B) $\lambda = 0.2$ and (C) $\lambda = 10$. In case (B) the signal is more buried in the observational noise than others, while in case (C) we have less measurement errors. As we will see in our empirical study, most of macroeconomic variables seem to be like case (C).

For each case (A) – (C), we consider 16 combinations of (a, b), see Tables 1 – 3. In unreported preliminary research, we examined lower values of a and b like 0.8 or 0.7. However we find only the cases close to unit roots and the unit roots case itself worth reporting here. In fact, the further away from 1 the stationary roots of the seasonal polynomial are, the AIC and BIC invariably detect the true data generating process.

For each case of the 16 DGPs, the number of experiments is 100, and the sample size is common throughout this experiment and is set to 150. Although we know the true initial state vector, we estimate it by backward filtering in each trial. This increases the penalty term in AIC and BIC but it obviously has no effect on model selection where only the difference of AIC and BIC matters.

We can summarize what we learn from Tables 1 – 3 as follows.

- The larger the trade-off parameter λ, the easier it is to detect the true seasonal structure. This result is not unexpected because a large λ means we have relatively small observational noise compared to the signal (or the seasonal pattern).

- AIC faces problems when $a = 1$, i.e. there is a biannual unit root in the seasonal polynomial. In the first row of every table, we can recognize the known tendency of AIC to prefer overparametrized models.

- The performance of the BIC is generally satisfactory, although BIC does not perform well when we have stationary roots at the annual frequency extremely close to the unit circle.

- It is relatively easier to detect the annual frequency unit roots than the biannual unit roots. Except for the three unit roots case, both AIC and BIC give the correct decision. Probably this is partly because the parameter b reflects the contributions of two seasonal frequencies due to the aliasing effect.

- When λ is large and $a \leq 0.99$ and $b \leq 0.98$, our method has a 'power' of 1, i.e., we always select the appropriate alternative model.

In the $a = 1$ and/or $b = 0.99$ cases, decisions based on AIC and on BIC rarely coincide. Therefore, we propose the rules of thumb in such split decision cases as given in Table 4. Though it is a rough indication, it seems to be fairly valid when we have a large λ value. In the next section we will observe that λ is large for many empirical data.

Before closing this section, we make a few more remarks on these and unreported tables. In both cases where AIC or BIC perform worse, we can see a similar tendency even if other roots are far from unity. For example, even if $b = 0.9$ or $b = 0.7$, the AIC cannot detect the true seasonal structure in the $a = 1$ case. Our second remark is that the $b = 0.99$ case is much closer to unit circle than the $a = 0.99$ case. Hence $1 + b^2B^2$ may be more natural parametrization than $1 + bB^2$.

5 Empirical Analysis

In this section we illustrate our approach for a selected set of 22 quarterly macroeconomic time series variables, namely 8 UK series previously analyzed in Osborn (1990), 6 US time series in Franses (1996) and GNP data of 8 countries in Hylleberg et al. (1993). In these studies, the data are all analyzed using the AR–based HEGY method. We summarize our results in Tables 5 to 7.

For each series, we report AIC and BIC values for all four models (21) – (24). In the third panel of each table, the number of seasonal unit roots by our method (upper) and the result of the above mentioned studies (lower) are shown. The trade-off parameter defined in section 4, estimated for the model selected using our method, is reported in the final panel. Except for UK import data (UKIMPORT), every series has a large λ value, so our decision rules are expected to perform well. The estimated seasonal innovation variance in UKIMPORT is extremely small, and the posterior mean of $\{s_t\}$ looks almost deterministic. This suggests that a model with constant seasonal dummies might be more suitable for this series. Also notice that the finding of model 3 by both criteria can be interpreted as the variance σ_s^2 equals zero for these data.

The empirical results in Tables 5 to 7 can be summarized as follows. Our BSM-based method and the AR-based HEGY method find agreement for only 4 of the 22 time series. In most cases, i.e. 10, the BSM-based method finds more unit roots than the HEGY method does. We find that model 3, which is the BSM in (8)– (10), is preferred in 8 of the 22 cases. The no seasonal unit root model is selected only twice (versus 7 times using the HEGY method). Hence, in general we find empirical evidence that confirm the simulation findings in Ghysels et al. (1994), i.e. we find more unit roots when we allow that the data can be approximately described by an airline model.

6 Conclusion

In this paper we proposed a model selection approach to detect the number of seasonal unit roots, especially for series which may well be described by the familiar airline model. Our Monte Carlo simulations suggest that our method performs well. Even in cases extremely close to the unit root case, the combination of AIC and BIC still provides a useful strategy to determine the number of seasonal unit roots. Our empirical analysis shows that our method and the HEGY method can lead to substantially different results, where our method tends to finds more seasonal unit roots, as expected.

References

Akaike, H., 1973, Information theory and an extention of the maximum likelihood principle, *in* B. Pertoc and F. Caski (eds), *Second International Symposium in Information Theory*, Akademiai Kiado, Budapest, pp. 267–281.

Akaike, H., 1980, Likelihood and the Bayes procedure, *in* J. Bernardo, M. D. Groot, D. Lindley and A. Smith (eds), *Bayesian Statistics*, University Press Valencia, Spain, pp. 143–166.

Anderson, B.D.O. and Moore, J.B., 1979, *Optimal filtering*, Prentice Hall, New Jersey.

Box, G.E.P. and Jenkins, G.M., 1970, *Time series analysis, forecasting and control*, Holden-Day, San Francisco.

Franses, P.H., 1996, Recent advances in modelling seasonality, *Journal of Economic Surveys* **10**, 299–345.

Gersch, W. and Kitagawa, G., 1988, Smoothness priors in time series, *in* J. Spall (ed.), *Bayesian Analysis of Time Series and Dymnamic Systems*, Marcel Dekker, New York, pp. 431–476.

Ghysels, E., Lee, H.S. and Noh, J., 1994, Testing for unit roots in seasonal time series, *Journal of Econometrics* **62**, 415–442.

Harvey, A.C., 1984, A unified view of statistical forecasting procedures, *Journal of Forecasting* **3**, 245–283.

Harvey, A.C., 1985, Trends and cycles in macroeconomic time series, *Journal of Business and Economic Statistics* **3**, 216–227.

Harvey, A.C., 1989, *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, Cambridge.

Hylleberg, S., Engle, R.F., Granger, C.W.J. and Yoo, B.S., 1990, Seasonal integration and cointegration, *Journal of Econometrics* **44**, 215–238.

Hylleberg, S., Jørgensen, C. and Sørensen, N.K., 1993, Seasonality in macroeconomic time series, *Empirical Economics* **18**, 321–335.

Kitagawa, G., 1981, A nonstationary time series model and its fitting by a recursive filter, *Journal of Time Series Analysis* **2**, 103–116.

Maekawa, K., 1994, Prewhitened unit root test, *Economic Letters* **45**, 145–153.

Osborn, D., 1990, A survey of seasonality in UK macroeconomic variables, *Journal of Forecasting* **6**, 327–336.

Psaradakis, Z., 1996, Testing for unit roots in time series with nearly deterministic seasonal variation, *Unpublished manuscript*, Department of Economics, Birkbeck College, University of London.

Rissanen, J., 1978, Modeling by shortest data description, *Automatica* **14**, 465–471.

Schwarz, G., 1978, Estimating the dimension of a model, *Annals of Statistics* **6**, 461–464.

Table 1: Simulation results for case (A) : $\lambda = 1$

	$b \rightarrow$	1.00		0.99		0.98		0.97	
$a \downarrow$		AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
1.00	model 0	0	0	94	1	95	2	90	3
	model 1	0	0	*4	*2	*5	*94	*10	*97
	model 2	94	0	0	0	0	4	0	0
	model 3	*6	*100	0	97	0	0	0	0
0.99	model 0	7	0	*100	*18	*100	*98	*100	*100
	model 1	0	0	0	0	0	0	0	0
	model 2	*93	*100	0	82	0	1	0	0
	model 3	0	0	0	0	0	1	0	0
0.98	model 0	21	0	*100	*26	*100	*99	*100	*100
	model 1	0	0	0	0	0	0	0	0
	model 2	*79	*100	0	74	0	1	0	0
	model 3	0	0	0	0	0	0	0	0
0.97	model 0	17	0	*100	*40	*100	*100	*100	*100
	model 1	0	0	0	0	0	0	0	0
	model 2	*83	*100	0	60	0	0	0	0
	model 3	0	0	0	0	0	0	0	0

The cells with an asterisk indicate how many times minimum AIC and BIC select the true DGP out of 100 experiments.

Table 2: Simulation results for case (B) : $\lambda = 0.2$

	$b \rightarrow$	1.00		0.99		0.98		0.97	
$a \downarrow$		AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
1.00	model 0	20	0	85	5	85	5	73	11
	model 1	0	0	*13	*8	*15	*81	*27	*87
	model 2	58	7	2	1	0	0	0	0
	model 3	*22	*93	0	86	0	14	0	2
0.99	model 0	24	0	*99	*32	*100	*80	*100	*92
	model 1	0	0	0	0	0	7	0	7
	model 2	*76	*94	1	56	0	8	0	0
	model 3	0	6	0	12	0	5	0	1
0.98	model 0	41	3	*100	*44	*100	*89	*100	*99
	model 1	0	0	0	0	0	0	0	0
	model 2	*59	*97	0	56	0	11	0	1
	model 3	0	0	0	0	0	0	0	0
0.97	model 0	48	0	*99	*48	*100	*97	*96	*94
	model 1	0	0	0	0	0	0	0	1
	model 2	*52	*100	1	51	0	3	4	5
	model 3	0	0	0	1	0	0	0	0

The cells with an asterisk indicate how many times minimum AIC and BIC select the true DGP out of 100 experiments.

Table 3: Simulation results for case (C) : $\lambda = 10$

$a \downarrow$	$b \rightarrow$	1.00		0.99		0.98		0.97	
		AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
1.00	model 0	0	0	99	0	100	0	100	0
	model 1	0	0	*1	*0	*0	*100	*0	*100
	model 2	100	0	0	0	0	0	0	0
	model 3	*0	*100	0	100	0	0	0	0
0.99	model 0	0	0	*100	*0	*100	*100	*100	*100
	model 1	0	0	0	0	0	0	0	0
	model 2	*100	*100	0	100	0	0	0	0
	model 3	0	0	0	0	0	0	0	0
0.98	model 0	0	0	*100	*1	*100	*100	*100	*100
	model 1	0	0	0	0	0	0	0	0
	model 2	*100	*100	0	99	0	0	0	0
	model 3	0	0	0	0	0	0	0	0
0.97	model 0	2	0	*100	*14	*100	*100	*100	*100
	model 1	0	0	0	0	0	0	0	0
	model 2	*98	*100	0	86	0	0	0	0
	model 3	0	0	0	0	0	0	0	0

The cells with an asterisk indicate how many times minimum AIC and BIC select the true DGP out of 100 experiments.

Table 4: Rules of thumb in split decision cases

AIC	BIC	decision
0	1	1
0	2	0
0	3	1
1	3	1
2	3	3

Table 5: Empirical results : UK data

		UKEXPORT	UKGDP	UKIMPORT	UKINVPUB
AIC values	Model 3	-477.88^*	-643.67^*	-459.69^*	-246.06
	Model 2	-476.02	-642.94	-458.67	-245.20
	Model 1	-476.76	-641.88	-458.12	-252.98
	Model 0	-475.10	-641.16	-456.27	-257.20^*
BIC values	Model 3	-472.04^*	-637.85^*	-453.86^*	-240.70
	Model 2	-467.29	-634.20	-449.96	-237.05
	Model 1	-468.02	-633.15	-449.38	-244.94
	Model 0	-463.45	-629.52	-444.62	-246.47^*
Decision					
– BSM-based method		3	3	3	0
– HEGY-AR method		0	0	0	3
trade-off param. λ		6.49	624.79	1.17×10^{-13}	41.36

		UKINVTOT	UKNONDUR	UKTOTCO	UKWORKFO
AIC values	Model 3	-464.99	-635.30^*	-601.24	-1069.36
	Model 2	-468.97^*	-634.83	-601.77^*	-1089.88^*
	Model 1	-463.01	-634.39	-600.86	-1082.06
	Model 0	-467.16	-633.95	-601.47	-1088.04
BIC values	Model 3	-459.16	-629.48^*	-595.42^*	-1063.54
	Model 2	-460.23^*	-626.09	-593.04	-1081.14^*
	Model 1	-454.27	-625.65	-592.13	-1073.32
	Model 0	-455.50	-622.30	-589.82	-1076.39
Decision					
– BSM-based method		2	3	3	2
– HEGY-AR method		0	3	3	0
trade-off param. λ		241.59	4322.37	2052.51	8780.94

The decision based on the HEGY tests is obtained from Osborn (1990). The data concern quarterly observations on exports, GDP, imports, public investment, total investment, consumption nondurables, total consumption and workforce. All data are in natural logs.

Table 6: Empirical results : US data

		USCONSTO	USDUR	USINDPRO
AIC values	Model 3	-778.21	-397.85	-607.17
	Model 2	-778.58	-397.34	-608.85^*
	Model 1	-780.67	-400.67^*	-606.35
	Model 0	-781.06^*	-400.47	-608.66
BIC values	Model 3	-771.83^*	-391.46^*	-601.47^*
	Model 2	-769.00	-387.76	-600.29
	Model 1	-771.09	-391.09	-597.80
	Model 0	-768.29	-387.69	-597.26
Decision				
– BSM-based method		1	1	3
– HEGY-AR method		1	1	1
trade-off param. λ		3800.73	282.03	96756.33

		USMONEY	USNONDUR	USSERVI
AIC values	Model 3	-753.97	-734.65	-1141.26
	Model 2	-753.04	-733.86	-1142.84
	Model 1	-754.33^*	-734.86^*	-1141.95
	Model 0	-753.42	-734.09	-1143.91^*
BIC values	Model 3	-748.15^*	-728.26^*	-1134.87^*
	Model 2	-744.30	-724.29	-1133.26
	Model 1	-745.59	-725.28	-1132.37
	Model 0	-741.77	-721.32	-1131.13
Decision				
– BSM-based method		1	1	1
– HEGY-AR method		0	3	0
trade-off param. λ		24271.13	2320.05	22012.87

The decision based on the HEGY tests is obtained from Franses (1996). The data concern quarterly observation on total consumption, consumption durables, industrial production, money (M1), consumption of nondurables and consumption of services. All data are in natural logs.

Table 7: Empirical results : GNP data

		JAPAN	UK	ITALY	TAIWAN
AIC values	Model 3	-261.59	-618.00^*	-209.74	-418.51
	Model 2	-262.95	-617.60	-208.94	-420.63
	Model 1	-261.83	-616.17	-211.35^*	-422.83
	Model 0	-263.18^*	-615.79	-210.64	-424.80^*
BIC values	Model 3	-256.62^*	-612.25^*	-205.55^*	-413.16
	Model 2	-255.49	-608.97	-202.65	-412.61
	Model 1	-254.37	-607.54	-205.07	-414.81^*
	Model 0	-253.22	-604.29	-202.26	-414.11
Decision					
– BSM-based method		1	3	1	1
– HEGY-AR method		3	1	3	3
trade-off param. λ		5359.25	765.51	998.55	1904.05

		NETHER	GERMANY	CANADA	SWEDEN
AIC values	Model 3	-170.22	-428.30	-421.89	-260.44^*
	Model 2	-168.85	-435.17	-435.51^*	-258.91
	Model 1	-171.19^*	-429.64	-421.09	-259.01
	Model 0	-170.06	-437.01^*	-434.94	-257.48
BIC values	Model 3	-166.65^*	-422.88	-416.55	-255.92^*
	Model 2	-163.50	-427.04^*	-427.49^*	-252.12
	Model 1	-165.84	-421.51	-413.07	-252.22
	Model 0	-162.93	-426.17	-424.25	-248.43
Decision					
– BSM-based method		1	0	2	3
– HEGY-AR method		3	2	3	2
trade-off param. λ		5095.84	2817.95	3981.91	694.63

The decision based on the HEGY tests is obtained from Hylleberg et al. (1993). All data are in natural logs.