Forecasting performance of seasonal cointegration models

Mårten Löf*† and Johan Lyhagen*

SSE/EFI *Working Paper Series in Economics and Finance, No. 336*
October 1999

Abstract

Forecasts from seasonal cointegration models are compared with those from a standard cointegration model based on first differences and seasonal dummies. The effects of restricting or not restricting seasonal intercepts in the seasonal cointegration models are examined as well as the recently proposed specification and estimation procedure for the annual frequency by Johansen and Schaumburg (1999). The data generating process used in the Monte Carlo simulation is based on an empirical six-dimensional macroeconomic data set. Results show that the seasonal cointegration model improves forecasting accuracy, compared with the standard cointegration model, even in small samples and if short forecast horizons are considered. Furthermore, the specification suggested by Johansen and Schaumburg seems to work better than the original model presented by Lee (1992). An empirical forecasting example confirm most of the results found in the Monte Carlo study.

Keywords: Seasonal cointegration; Monte Carlo.
JEL: C32, C53.

*Department of Economic Statistics, Stockholm School of Economics, P.O Box 6501 SE-113 83 Stockholm, E-mail: stmlo@hhs.se and stjl@hhs.se.
†Corresponding author

1 Introduction

It is common in applied work to assume an approximately constant seasonal pattern, modelled by seasonal dummies. Recently the unit root and cointegration analysis at the zero frequency have been extended to seasonal frequencies. Nonzero frequency unit roots imply a stochastic or changing seasonal pattern. Assuming a deterministic seasonal pattern in a multivariate setting, when it is in fact evolving over time, could lead to inappropriate inference about short and long run dynamics in the system. These negative effects may be felt strongly when the model is used in forecasting.

Similar to the approach suggested by Johansen (1995) for the zero frequency, Lee (1992) presents a maximum likelihood estimator for seasonally cointegrating relations. The seasonal error correction model [$SECM$], which allows for stochastic seasonality, is however only partially correct when it comes to testing for the cointegrating rank at the annual frequency. A certain parameter restriction has been used in the literature to overcome this problem. If this restriction is relaxed one cannot apply estimation methods based on canonical correlations, as suggested by Lee.

Johansen and Schaumburg (1999) argue that this is a peculiar restriction as it constrains all coefficients at the annual frequency. They present another estimation procedure for the parameters corresponding to this frequency and introduce a general asymptotic theory for the seasonal cointegration model.

Kunst and Franses (1998) argue that deterministic seasonal dummy variables, which are often included unrestrictedly in the $SECM$, should be confined to the seasonal cointegrating relations instead. If cointegration at seasonal frequencies exists, an inclusion of unrestricted seasonal intercepts then implies a growing amplitude in the seasonal cycles, which may not be a realistic assumption in most cases.

In the present study, various seasonal cointegration specifications will be contrasted to the model suggested by Johansen and Schaumburg (1999) [henceforth JS]. We have found three studies on forecasting and seasonal cointegration: Kunst (1993), Reimers (1997) and Kunst and Franses (1998). Kunst (1993) contrasts the $SECM$, where an intercept is included and the restriction on the annual frequency is imposed, to a $VECM$ and also a VAR model in first differences and seasonal dummies. Two examples based on real data and a Monte Carlo experiment show that the benefits from accounting for seasonal cointegration are quite limited. Reimers (1997) uses the conventional seasonal cointegration model suggested by Lee, with another simplifying restriction on the annual frequency and considers a simulated two variable seasonal cointegration model, with a fixed lag length and no seasonal intercepts. The $SECM$ is compared to a traditional $VECM$ in first differences with seasonal dummies. The main conclusion is that models in first differences produce smaller forecast errors for short horizons, but when longer forecasting periods are considered the seasonal cointegration model appears preferable. Kunst and Franses (1998) investigate the forecasting effects of first deleting then either restricting or not restricting the seasonal intercepts as discussed above. Using three empirical data sets and various forecasting periods they show that the suggested restricted seasonal dummy approach yields better forecasts in most cases.

In the present study a seasonally cointegrated data generating process [DGP] based on estimation results from empirical data is used. The system is larger than the simulation studies mentioned above. The forecasting performance of the more general specification for the annual frequency, recently proposed by JS is evaluated. In one part of the analysis the forecasting performance of the various specifications are compared after selecting lag lengths and estimating the cointegrating rank. The macroeconomic data sets for Austria, Germany and the United Kingdom, which are also used in Kunst and Franses (1998), each consists of six variables.

The remainder of this paper is organized as follows. Section 2 presents specifi-

cations of the $SECM$, while Section 3 describes the data and the estimation results for the models to be used as DGP. In Section 4 the Monte Carlo setup is discussed. A comparison of model forecasts is carried out in Section 5. Section 6 presents forecast performance when using the empirical data sets. The final section contains conclusions.

2 Seasonal Cointegration

Lee (1992) suggests a maximum likelihood estimator for seasonal cointegration relations. This procedure extends the maximum likelihood approach, which is summarized in Johansen (1995). Assuming quarterly data and that $\Delta_4 Y_t$ is stationary, a seasonal $VECM$ model of the following form is considered:

$$\Delta_4 Y_t = \sum_{i=1}^{4} \alpha_i \beta_i^{'} Z_{i,t} + \Gamma_1 \Delta_4 Y_{t-1} + ... + \Gamma_{p-4} \Delta_4 Y_{t-p+4} + \Phi D_t + \varepsilon_t, \qquad (1)$$

where the D_t are deterministic components and where ε_t is i.i.d. $N_n(0, \Omega)$. The process $\Delta_4 Y_t$ above is said to be seasonally cointegrated if and only if at least one of the $\alpha_i \beta_i^{'}$ matrices for $i = 2, 3, 4$ on the right hand side has reduced, non-zero rank. The linear filters which in (1) remove all unit roots except those on the zero, biannual and annual frequencies, respectively are:

$$\begin{aligned} Z_{1,t} &= (L + L^2 + L^3 + L^4) Y_t, \\ Z_{2,t} &= (L - L^2 + L^3 - L^4) Y_t, \\ Z_{3,t} &= (L^2 - L^4) Y_t, \\ Z_{4,t} &= (L - L^3) Y_t. \end{aligned}$$

These filters are the vector equivalents of the univariate transformations used in the so called HEGY-test for seasonal unit roots, see Hylleberg et *al.* (1990) [HEGY]. If the matrices $\alpha_i \beta_i^{'}$ have reduced, but non-zero rank this implies that $\beta_i^{'} Z_{i,t}$ is stationary even though the processes $Z_{i,t}$ are nonstationary. Furthermore, the regressors $Z_{i,t}$ are asymptotically uncorrelated:

$$T^{-1} \sum_{t=1}^{T} Z_{i,t} Z_{j,t}^{\prime} \xrightarrow{P} 0, i \neq j,$$

implying that the cointegration vectors and adjustment coefficients can be found by removing the reduced rank restriction on the other frequencies by concentrating out the associated regressors.

Strictly speaking, model (1) is only partially correct at frequency $\pi/2$. As mentioned in the introduction, it is often suggested that one should impose the restriction $\alpha_4 \beta_4^{'} = 0$, i.e. that $Z_{4,t}$ has no influence on $\Delta_4 Y_t$. If this restriction is relaxed one cannot apply the estimation method that uses canonical correlations.

JS, who argue that this is a strong restriction, which is not justified from a theoretical point of view, refines the theory for seasonal cointegration in the general case. For quarterly data, where the process $\Delta_4 Y_t$ is assumed to be stationary, they propose the following error correction model:

$$\begin{aligned} \Delta_4 Y_t &= \sum_{i=1}^{2} \alpha_i \beta_i^{'} X_{i,t} + 2(\alpha_R \beta_R^{'} + \alpha_I \beta_I^{'}) X_{R,t} + 2(\alpha_R \beta_I^{'} - \alpha_I \beta_R^{'}) X_{I,t} \\ &\quad + \Gamma_1 \Delta_4 Y_{t-1} + ... + \Gamma_{p-4} \Delta_4 Y_{t-p+4} + \Phi D_t + \varepsilon_t, \qquad (2) \end{aligned}$$

where the processes $X_{1,t}, ..., X_{I,t}$ are:

$$\begin{aligned} X_{1,t} &= \frac{1}{4}Z_{1,t}, \\ X_{2,t} &= -\frac{1}{4}Z_{2,t}, \\ X_{R,t} &= \frac{1}{4}Z_{3,t}, \\ X_{I,t} &= -\frac{1}{4}Z_{4,t}, \end{aligned}$$

respectively. It can be seen that the restriction $\alpha_4\beta_4^{'} = 0$ mentioned above would imply that $\alpha_R\beta_I^{'} - \alpha_I\beta_R^{'} = 0$ in (2), which is a peculiar restriction on the coefficients at the annual frequency.

The estimation of $\alpha_i\beta_i^{'}$ for $i = 1, 2$ uses canonical correlations in analogy to the Johansen procedure and hence does not require any detailed explanation here. However the estimation of β_R and β_I is nonstandard.

After regressing $\Delta_4 Y_t$, $X_{j,t}$, $j = 1$, 2, R, I and ε_t on lagged values of $\Delta_4 Y_t$ and D_t and defining the residuals as R_{0t}, R_{jt} and η_t, respectively then removing the restriction of reduced rank on $\alpha_i\beta_i^{'}$ for $i = 1, 2$ by regressing R_{0t}, R_{Rt} and R_{It} on R_{1t} and R_{2t}, it is shown in JS that the resulting residuals (U_{0t}, U_{Rt} and U_{It}) satisfy the following equation:

$$\begin{aligned} U_{0t} &= 2(\alpha_R\beta_R^{'} + \alpha_I\beta_I^{'})U_{Rt} + 2(\alpha_R\beta_I^{'} - \alpha_I\beta_R^{'})U_{It} + \xi_t \\ &= 2\left(\alpha_R - \alpha_I\right)\begin{pmatrix} \beta_R & -\beta_I \\ \beta_I & \beta_R \end{pmatrix}^{'}\begin{pmatrix} U_{Rt} \\ U_{It} \end{pmatrix} + \xi_t \\ &= \widetilde{\alpha}\boldsymbol{\beta}^{'}U_{1t} + \xi_t, \end{aligned} \tag{3}$$

asymptotically. Defining the product moments as $S_{ij} = (1/T)\Sigma_{t=1}^{T} U_{it}U_{jt}^{'}$ for $i, j = 0, 1$, we have that for fixed values of $\boldsymbol{\beta}$ the concentrated likelihood function with respect to $\widetilde{\alpha}$ and hence the variance-covariance matrix Ω_ξ take the form, apart from a constant:

$$L_{\max}^{-\frac{2}{T}}(\boldsymbol{\beta}) = |\widehat{\Omega}_\xi| = |S_{00}|\frac{|\boldsymbol{\beta}^{'}(S_{11} - S_{10}S_{00}^{-1}S_{01})\boldsymbol{\beta}|}{|\boldsymbol{\beta}^{'}S_{11}\boldsymbol{\beta}|}. \tag{4}$$

The maximization of (4) cannot be solved as an eigenvalue problem as in the zero and biannual frequency cases since $\boldsymbol{\beta}$ itself has complex structure while the product matrices $S_{11} - S_{10}S_{00}^{-1}S_{01}$ and S_{11} do not have such a structure. JS use a switching algorithm proposed by Boswijk (1995) where the maximum likelihood estimator of $\boldsymbol{\beta}$ is calculated iteratively: Isolate β_R and β_I by using a normalized form ($\Omega_\xi^{-1/2}U_{0t} = \widetilde{U}_{0t}$) of U_{0t}, namely:

$$\begin{aligned} \widetilde{U}_{0t} &= 2\Omega_\xi^{-1/2}(\alpha_R\beta_R^{'} + \alpha_I\beta_I^{'})U_{Rt} + \\ &\quad 2\Omega_\xi^{-1/2}(\alpha_R\beta_I^{'} - \alpha_I\beta_R^{'})U_{I,t} + \Omega_\xi^{-1/2}\xi_t \\ &= \alpha_R^N\beta_R^{'}U_{Rt} - \alpha_I^N\beta_R^{'}U_{It} + \alpha_R^N\beta_I^{'}U_{It} - \alpha_I^N\beta_I^{'}U_{Rt} + \widetilde{\xi}_t, \end{aligned} \tag{5}$$

where $\Omega_\xi^{-1/2}\xi_t = \widetilde{\xi}_t \overset{iid}{\sim} N_p(0, I)$, and then vectorize (5) by using the formula $vec(\alpha_i^N\beta_i^{'}U_{it}) = (U_{it}^{'} \otimes \alpha_i^N)vec(\beta_i^{'})$. Since $\widetilde{U}_{0t}$ is a vector and hence $vec(\widetilde{U}_{0t}) = \widetilde{U}_{0t}$, this yields:

$$\widetilde{U}_{0t} = U_{2t}\begin{bmatrix} vec(\beta_R^{'}) \\ vec(\beta_I^{'}) \end{bmatrix} + \widetilde{\xi}_t \tag{6}$$

where $U_{2t} = \Big[(U_{Rt}^{'} \otimes \alpha_R^N) - (U_{It}^{'} \otimes \alpha_I^N) \quad (U_{It}^{'} \otimes \alpha_R^N) + (U_{Rt}^{'} \otimes \alpha_I^N)\Big]$. β_R and β_I can now be found from:

$$\begin{bmatrix} vec(\beta_R^{'}) \\ vec(\beta_I^{'}) \end{bmatrix} = \left(\sum_{t=1}^{T} U_{2t} U_{2t}^{'}\right)^{-1} \sum_{t=1}^{T} U_{2t} U_{0t}^{'}. \tag{7}$$

For a given value of $\boldsymbol{\beta}$, which is generated randomly in the first iteration, estimates of $\widetilde{\alpha} = (S_{01}\boldsymbol{\beta}(\boldsymbol{\beta}^{'} S_{11}^{-1}\boldsymbol{\beta})^{-1})/2$ and $\Omega_\xi = S_{00} - S_{01}\boldsymbol{\beta}(\boldsymbol{\beta}^{'} S_{11}^{-1}\boldsymbol{\beta})^{-1})\boldsymbol{\beta}^{'} S_{10}$ are computed. In a second step U_{0t} is normalized and vectorized as described above, yielding a new estimate of $\boldsymbol{\beta}$ for which we can compute a new likelihood function. This procedure is repeated until a suitable convergence criterion is satisfied.

Kunst and Franses (1998) argue that deterministic seasonal dummy variables, which are often included unrestrictedly in (1) to handle the deterministic part of seasonality, should be confined to the seasonal cointegrating relations instead. This is because unrestricted seasonal intercepts in the $SECM$ may lead to diverging seasonal trends, which is unlikely in most cases. They propose the following model instead:

$$\begin{aligned} \Delta_4 Y_t \quad = \quad & \mu + \alpha_1 \beta_1^{'} Z_{1,t} + \alpha_2[\beta_2^{\prime} Z_{2,t} + \rho_2(cos\pi(t-1))] + \\ & \alpha_3[\beta_3^{\prime} Z_{3,t} + (\rho_3, \rho_4)(cos\ \frac{\pi}{2}(t-1) \quad cos\ \frac{\pi}{2}(t-2))] + \\ & \Gamma_1 \Delta_4 Y_{t-1} + ... + \Gamma_{p-4} \Delta_4 Y_{t-p+4} + \Phi D_t + \varepsilon_{t,} \end{aligned} \tag{8}$$

where ρ_2, ρ_3 and ρ_4 are the restricted seasonal intercepts. We restrict the seasonal dummy variables in the same way for the JS specification in the section below.

3 Tests for Seasonal Integration and Cointegration

Gross domestic product (Y), private consumption (C), gross fixed investment (I), goods exports (X), real wages (W) are transformed into natural logarithms. The real interest rates (R) are given in percentage points. The three data sets are plotted in appendix A 5-7. The motivation for using these data series is that the neoclassical growth theory suggests several long-run relations among these variables. The data set for Austria covers the time period 1964:1 to 1994:4, whereas the time series for Germany run from 1960:1 to 1988:1 and for the UK from 1957:1 to 1994:1. Unit roots in the univariate time series are tested using the HEGY-test. The test procedure investigates whether the seasonal difference $(1 - B^4)$, which assumes the presence of four unit roots, is the appropriate filter compared with other nested filters.

Results of the tests for seasonal and nonseasonal unit roots appear in Table 3.1. All auxiliary regressions include an intercept (I), seasonal dummies (D) and a deterministic trend (T), except those for the real interest rates where an exclusion of the trend seems to be a more appropriate specification. All variables seem to contain unit roots at the zero frequency, except UK consumption and the real interest rate in Austria. The results are more mixed at the seasonal frequencies. All roots at the biannual frequency, except for consumption, are rejected for Austria. On the other hand we find evidence of four unit roots at this frequency for both Germany and UK. Turning to the annual frequency we find evidence of three, one and four stationary variables for Austria, Germany and UK, respectively. Since unit roots are most frequent in the German data set it is sensible to conclude that cointegration at all frequencies is most likely to be found there.

Table 3.1. Seasonal integration tests.

	Austria			*Germany*			*UK*		
Var.	t_{π_1}	t_{π_2}	$F_{\pi_{3,4}}$	t_{π_1}	t_{π_2}	$F_{\pi_{3,4}}$	t_{π_1}	t_{π_2}	$F_{\pi_{3,4}}$
Y	-1.78	**-3.47**	1.81	-1.52	**-2.83**	3.29	-3.21	-1.38	4.43
C	-1.28	-1.83	3.14	-1.21	-1.45	3.41	**-3.49**	-1.85	3.73
I	-2.29	**-4.96**	**9.79**	-2.83	-2.68	6.21	-2.68	-2.53	**13.3**
X	-1.49	**-5.18**	**48.6**	-0.96	**-5.13**	**39.7**	-1.12	-2.12	**6.82**
W	-1.69	**-3.31**	2.58	-1.18	-1.62	2.18	-2.16	**-3.05**	**19.9**
R	**-3.01**	**-2.95**	**7.20**	-2.77	-2.03	4.73	-1.94	**-5.38**	**22.3**
Lags of $\Delta_4 y_t$ and additional regressors (μ).									
	Lags	μ		Lags	μ		Lags	μ	
Y	1-4	I,D,T		1	I,D,T		1-5	I,D,T	
C	1-4	I,D,T		1-4	I,D,T		1-5	I,D,T	
I	1-2	I,D,T		1-3	I,D,T		1	I,D,T	
X	-	I,D,T		-	I,D,T		1-4	I,D,T	
W	1-6	I,D,T		1-4	I,D,T		1	I,D,T	
R	1-3	I,D		1	I,D		1-5	I,D	

For the cointegration tests we use the same lag lengths as Kunst and Franses (1998), indicated by k. The ranks of the matrices $\alpha_i \beta_i^{'}$ $(i = 1, 2)$ and $\widetilde{\alpha}\boldsymbol{\beta}^{'}$ are determined using the trace test, where the null hypothesis is: at most r cointegrating vectors against the alternative of at most $r + 1$ vectors. Table 3.2 summarizes the result using the model with unrestricted zero frequency intercept and other seasonal intercepts restricted. Numbers written in bold indicate significance at the 5% level. Critical values for the zero frequency are from Table 15.3 in Johansen (1995) and for the biannual frequency from Tables 1a-1f in Franses and Kunst (1996). For the annual frequency we use the critical values from Table 2 in JS. Two long-run cointegrating vectors are found on the zero frequency for Germany and three for Austria and the UK. On frequency π we find evidence of six vectors for Austria and identify two cointegrating vectors for Germany. For the UK we find no evidence of cointegration at frequency π. Finally at frequency $\pi/2$ three vectors are found for Austria, two for Germany and one for the UK. Table 3.4 show our results for the Lee specification with unrestricted seasonal intercepts, while Table 3.5 shows the results for the same model with restricted seasonal intercepts. Lee and Siklos (1995) present critical values for the unrestricted seasonal intercepts case, but they only consider smaller systems. Critical values for Table 3.4 are therefore based on our own calculations. For the biannual and annual frequencies we consult Tables 1a-1f in Franses and Kunst (1999). It can be seen in Table 3.4 that we find evidence of one more cointegrating vector on the biannual frequency for Germany and the UK as compared to the results in Table 3.2. For the annual frequency there is some evidence of a fourth vector in Austria. The conclusions from Table 3.5 are the same as from Table 3.2. Having chosen the rank for the different frequencies we test for further reduction of the model. A hypothesis that simplifies the cointegration analysis on frequency $\pi/2$ is the one implying real structure (H_{REAL}), i.e. $\beta_I = 0$ in (2). However this hypothesis is strongly rejected (see Table 3.3). Based on these results, where the biannual frequency for Austria seems to have full rank and no evidence for cointegrating relationships at the same frequency for the UK are found, the German data set is used in the Monte Carlo forecasting analysis.

Table 3.2. Cointegration tests. Restricted seasonal intercepts [JS].

	Austria, $k=0$			*Germany*, $k=0$			UK, $k=5$		
p-r	0	π	$\pi/2$	0	π	$\pi/2$	0	π	$\pi/2$
1	**161.3**	**161.4**	**284.7**	**132.2**	**128.9**	**287.3**	**140.7**	95.9	**218.7**
2	**88.1**	**113.1**	**174.6**	**68.8**	**82.7**	**175.3**	**73.6**	58.2	134.2
3	**48.8**	**73.8**	**106.1**	38.2	40.8	97.8	**47.7**	39.4	74.7
4	22.5	**42.7**	56.2	17.2	16.8	50.4	25.7	22.7	41.0
5	6.8	**20.9**	28.7	8.4	8.6	21.7	8.7	11.3	17.7
6	1.5	**9.7**	9.4	1.9	4.1	6.7	0.8	3.2	5.8

Table 3.3. Tests for real restriction, $\beta_I = 0$.

	LR_{REAL}	$\overset{a}{\sim} \chi^2(r_{\pi/2}(p - r_{\pi/2}))$
Austria	43.3	$p_{value} = 0.00$
Germany	68.8	$p_{value} = 0.00$
UK	20.6	$p_{value} = 0.00$

Table 3.4. Cointegration tests. Unrestricted seasonal intercepts.

	Austria, $k=0$			*Germany*, $k=0$			UK, $k=5$		
p-r	0	π	$\pi/2$	0	π	$\pi/2$	0	π	$\pi/2$
1	**161.3**	**181.0**	**200.1**	**132.2**	**140.8**	**174.5**	**140.7**	**104**	**130.7**
2	**88.1**	**116.5**	**122.8**	**68.8**	**94.1**	**95.1**	**73.6**	64.5	71.7
3	**48.8**	**77.2**	**77.8**	38.2	**52.3**	54.5	**47.7**	42.0	40.6
4	22.5	**46.1**	**40.1**	17.2	24.3	29.8	25.7	25.3	22.4
5	6.8	**21.0**	18.9	8.4	7.9	15.2	8.7	10.5	7.7
6	1.4	**9.7**	6.2	1.9	3.2	4.0	0.8	2.8	1.7

Table 3.5. Cointegration tests. Restricted seasonal intercepts.

	Austria, $k=0$			*Germany*, $k=0$			UK, $k=5$		
p-r	0	π	$\pi/2$	0	π	$\pi/2$	0	π	$\pi/2$
1	**161.3**	**161.4**	**187.9**	**132.2**	**128.9**	**171.2**	**140.7**	95.9	**130.2**
2	**88.1**	**113.1**	**111.1**	**68.8**	**82.7**	**93.6**	**73.6**	58.2	71.5
3	**48.8**	**73.8**	**67.2**	38.2	40.8	56.8	**47.7**	39.4	40.4
4	22.5	**42.7**	41.7	17.2	16.8	32.0	25.7	22.7	22.1
5	6.8	**20.9**	21.3	8.4	8.6	17.4	8.7	11.3	7.6
6	1.5	**9.7**	7.8	1.9	4.1	5.9	0.8	3.2	1.6

4 The Monte Carlo setup

The data generating processes are based on parameter estimates from the above seasonal cointegration model specifications using the German data set. In a first step we compare the forecast accuracy for the seasonal cointegration models when seasonality is viewed as being purely stochastic, i.e. only a single intercept is included. The seasonal cointegration specifications we compare are those of JS (JSA) and Lee (1992) (LEE), respectively. Both specifications are contrasted with a traditional $VECM$ in first differences, including seasonal dummies (JD). We also compare the models when a deterministic seasonal component is introduced. The compared specifications in this case are the JS specification with restricted seasonal dummies (JSB), the Lee (1992) model with unrestricted seasonal dummies ($LEEUD$) and restricted seasonal dummies ($LEERD$). For the two latter model specifications

we impose the restriction $\alpha_4\beta_4^{'} = 0$. The forecast errors from the three seasonal cointegration models in fourth differences are again compared with those from a cointegrated model (Johansen's procedure) in first differences with a constant and three centered seasonal dummies (JD). Series of length T = 100 and T = 200 are generated from the various DGPs with $N(0,1)$ serially and mutually uncorrelated errors. Furthermore, 100 presample observations and 20 postsample observations are generated. These last 20 observations are saved and used for ex-ante forecasting. In each case 1000 replications (N) are generated.

In the first part of the analysis we do not test for the proper lag order for the model specifications. Instead we set $k = 0$ for the seasonal cointegration models and $k = 3$ for model in first differences. With this lag structure the zero frequency is exactly comparable for all model specifications. Moreover, we do not test for the cointegrating rank at different frequencies in the first part. We consider instead the cointegrating ranks $r_0 = 1, 2$, $r_\pi = 1, 2$ and $r_{\pi/2} = 1, 2$, for the zero, biannual and annual frequencies, respectively. In the second part of the analysis we choose the lag order according to the Schwarz Criterion (SBC). We also conduct tests for the cointegrating rank with relevant critical values for every replication.

The figures reported are computed in the following way. For each model we calculate squared forecast errors, in levels (as to preserve comparability), for all six equations. We consider up to 20-step ahead forecasts (only 1-8 step ahead forecast errors are shown in the figures). We then compute the squared average forecast error over the equations for each horizon. Finally we compute the average of these after conducting 1000 replications.

5 Monte Carlo Results

Figures 1 - 4 in the appendix summarize the Monte Carlo study, where we start by analyzing the forecasting accuracy when the DGPs only include an intercept and the seasonality is thus generated as being purely stochastic. The two DGPs considered are JSA and LEE. These two seasonal cointegration specifications are contrasted with JD. A comparison of the two seasonal cointegration approaches when $r_0 = 1$, $r_\pi = 1$ and $r_{\pi/2} = 1$, summarized in figure 1, indicates that JSA produces the best forecasts. If the DGP is based on estimates using the JS specification, LEE produces even larger forecasts errors than the model in first differences for short forecast horizons, which is in line with the results found by Reimers (1997). However, if we consider a DGP based on estimates of $\alpha_i\beta_i^{'}$ ($i = 1, 2$ and 3) using the original Lee (1992) model, it can be seen that JSA generates exactly the same system average forecast errors even though the smaller sample size is considered. This is an indication that the general specification of the annual frequency proposed by JS may result in a good forecasting behavior even if the $\alpha_4\beta_4^{'} = 0$ restriction in (1) is clearly motivated. The case when r_0, r_π and $r_{\pi/2}$ are set to be equal to two results in almost the same conclusions (not shown here). However, when the DGP is based on LEE estimates of the adjustment- and long-run parameters, JSA produces slightly larger forecast errors than the DGP based model (LEE), which is estimated using canonical correlations at each frequency. Even though the difference is not big it could be an indication that the estimation procedure for the annual frequency proposed by JS works best if there are few cointegrating relations to be estimated.

When seasonality is generated as being both deterministic and stochastic we compare three seasonal cointegration specifications, namely, the Johansen and Schaumburg (1999) model with restricted seasonal dummies JSB, the Lee (1992) specification with restricted seasonal dummies, proposed by Kunst (1998) and Franses, $LEERD$ and the unrestricted seasonal dummy model $LEEUD$. These are again

contrasted with a cointegration model in first differences with deterministic seasonal dummies JD. The results, summarized in figures 3-4, are comparable with those mentioned above. The forecast errors generated by JSB are always very close to those generated by $LEERD$ even though the latter specification is used as DGP. However, the opposite is not true for any combination of ranks at different frequencies in the DGPs. Again, the two Lee specifications produce larger forecasts errors than JD for short forecast horizons, especially if the larger sample size is considered (not shown here). Finally, if the seasonal cointegration model with unrestricted seasonal intercepts is used as DGP that same model produces larger forecast errors than JSB and $LEERD$, which confirms the results of Kunst and Franses (1998).

If we consider the case when the lag order is chosen and tests for the cointegrating rank are done in every replication we get the same results (not shown here). If the DGP is based on estimates using the JS specification, $LEERD$ and $LEEUD$ produces larger forecasts errors compared to JSB and if we consider a DGP based on estimates of $\alpha_i \beta_i^{'}$ ($i = 1, 2$ and 3) using the original LEE specification with or without restricted seasonal intercepts JSB performs almost equally well, even if the smaller sample size is considered.

We also want to study the forecasting performance of the JS specification if the hypothesis real structure (H_{REAL}), i.e. $\beta_I = 0$ at the annual frequency in (2) is satisfied. Our DGP in this case are the JS specification with real structure imposed and we consider the case where seasonality is generated as being purely stochastic and the case when seasonality is generated as being both deterministic and stochastic. We estimate and make forecasts from models with and without the restriction imposed and from models where we first test for the restriction and then make forecasts according to the result. Comparing the forecasting performance for this three approaches reveal no big differences, but the approach to pretest for the restriction and then make forecasts seems to outperform the second approach, if longer forecast horizons are considered and more cointegrating relations are to be estimated. The first approach, where we just estimate and make forecasts from the DGP based model gives, not surprisingly, the best results.

6 An Empirical Forecasting Example

In this section we investigate the forecasting performance of some of the models presented in previous sections, namely JSB, $LEERD$, $LEEUD$ and JD, using the empirical data set for Germany, which covers the time period 1960:1 - 1988:1. We impose the ranks $r_0 = 2$, $r_\pi = 2$ and $r_{\pi/2} = 2$ for the seasonal cointegration models and $r_0 = 2$ for the model in first differences. Furthermore, we set $k = 0$ for the seasonal cointegration models and $k = 3$ for JD. All squared forecast errors correspond to level variables. In the first step we save eight observations and make ex ante forecasts. In the next step the sample is extended by one quarter, but we do not reestimate the equations. These extensions of the sample are through 1987:4, where we generate the last one-step ahead forecast error. Hence there are eight one-step ahead forecast errors, seven two-step ahead forecast errors and so on. The eight-step ahead forecasts are then based on a single observation for each equation. The results are summarized by the RMSE statistic (root mean squared error) for 1, 2, 4 and 8-step ahead forecasts in Table 4.1 below.

Table 4.1. Forecast errors evaluated with RMSE*100.

Horizon:	1	2	4	8		1	2	4	8
JSB					LEEUD				
Y	**1.26**	1.22	1.44	**0.08**	Y	1.32	**0.93**	1.30	0.34
C	1.15	1.13	0.92	**0.11**	C	0.91	0.82	0.86	0.61
I	4.12	3.09	**3.51**	7.67	I	3.86	**2.66**	3.52	**6.96**
X	**2.37**	**3.12**	**4.22**	**4.37**	X	2.83	3.13	4.69	5.69
W	1.79	1.94	1.88	**2.58**	W	1.80	1.76	1.88	3.36
R	1.55	1.76	2.01	**1.46**	R	1.30	1.32	**1.69**	1.97
LEERD					JD				
Y	1.35	0.97	**1.29**	0.14	Y	1.31	1.79	1.81	0.02
C	**0.86**	**0.76**	**0.81**	0.15	C	1.22	1.50	1.31	0.93
I	**3.79**	2.73	3.63	7.53	I	4.60	4.50	3.64	4.84
X	2.73	3.17	4.67	5.11	X	1.90	2.53	4.23	4.95
W	**1.75**	**1.74**	**1.85**	3.30	W	1.92	2.43	2.26	2.47
R	**1.27**	**1.32**	1.70	1.60	R	2.60	2.73	1.40	1.60

If we start to compare the three seasonal cointegration specifications the results in Table 4.1 indicate that JSB is better than $LEERD$ and $LEEUD$ if longer forecast horizons are considered (the smallest RMSEs generated from the seasonal cointegration models are indicaded in bold). The results for JSB are worse if we consider one, two and four-step ahead forecasts where $LEERD$ generates better forecasts than JSB. Notable are that 20 out of the 24 smallest RMSE statistics are from the two specifications with restricted seasonal dummies. If a third cointegrating vector is imposed at the biannual frequency in $LEEUD$, which we found evidence for in Section 3, that model behaves better and generate smaller errors for shorter horizons. When we include the model in first differences (JD) in the analysis we see that it works better for longer forecast horizons (underlined figures indicate smaller RMSEs than the seasonal cointegration models). If the whole sequence from one to eight-steps ahead forecasts are considered we find that 11 out the 48 smallest RMSE statistics are generated from the model in first differences and five of those comes from the goods exports equation. Notable is that the results of the HEGY-tests indicate no seasonal unit roots for that particular time series, which may explain why JD generates better forecasts for goods exports. The result from this empirical part thus seem to confirm most of the results from the Monte Carlo study.

7 Conclusions

The forecasting performance of the seasonally cointegrated model of Johansen and Schaumburg (1999) [JS] is compared to related specifications and to a standard model based on first differences and seasonal dummies. We examine data sets from Austria, Germany and the UK, each including six variables, i.e. gross domestic product, private consumption, gross fixed investment, goods exports, real wages and the real interest rate. Having examined the integration and cointegration properties for the three data sets we select a data generating processes that is based on parameter estimates generated from the different seasonal cointegration model specifications using the German data set. In the first part of the Monte Carlo study, where lag lengths and cointegrating ranks are set in advance, the general specification proposed by JS has the best forecasting performance. There is no big differences in forecasting performance between that specification and the specifica-

tion proposed by Lee (1992) even if the $\alpha_4\beta_4^{'} = 0$ restriction is valid i.e. the latter model is used as DGP. In some cases The JS specification actually performs equally well. In the second part of the analysis, lag order selection and testing procedures for cointegrating rank are included in each replication and the results are quite similar. In short, these results show that different seasonal cointegration specifications do improve the forecasting accuracy compared with the standard cointegration model, also in small samples and if short forecast horizons are considered. Furthermore, the model suggested by JS seems to work better than the original model presented by Lee (1992). We also study the forecasting performance of the JS specification if the hypothesis of a real structure (H_{REAL}), i.e. $\beta_I = 0$ at the annual frequencies is satisfied. We estimate and make forecasts from models with and without the restriction imposed and from models where we pretest for the restriction and then make forecasts according to the result. Comparing the forecasting performance for these three approaches reveals no big differences, but the approach to first test for the restriction and then make forecasts seems to outperform the second approach, if longer forecast horizons are considered and more cointegrating relations are to be estimated. In the final section we investigate the forecasting performance using the empirical data set of Germany. A comparison of the three seasonal cointegration specifications indicate that JSB is better than $LEERD$ and $LEEUD$ if longer forecast horizons are considered. However, for one, two and four-step ahead forecasts $LEERD$ generates better forecasts than JSB. The model in first differences generate larger forecast errors as compared to the seasonal cointegration models as a whole in most cases, except for goods exports. A possible explanation for this could be that this particular time series do not seem to contain any seasonal unit roots, according to the HEGY-test.

Acknowledgments

We are grateful Lars-Erik Öller for insightful comments and to Robert Kunst who generously allowed us to use the data sets. Financial support for Lyhagen from Tore Browaldhs Foundation for Scientific Research and Education and The Swedish Foundation for International Cooperation in Research and Higher Education (STINT) are gratefully acknowledged.

References

Boswijk HP (1995) Identifiability of cointegrated systems. Discussion paper, TI 95,78.

Dahl Pedersen I (1996) A practical implementation of the seasonal cointegration theory. Masters Thesis, University of Copenhagen.

Engle RF, Granger CWJ, Hylleberg S, Lee HS (1993) Seasonal cointegration: The Japanese consumption function. Journal of Econometrics 55: 275-298.

Franses PH and Kunst RM (1999) On the role of seasonal intercepts in seasonal cointegration. Oxford Bulletin of Economics and Statistics 61: 409-434.

Hylleberg S, Engle RF, Granger CWJ, Yoo BS (1990) Seasonal integration and cointegration. Journal of Econometrics 44: 215-238.

Johansen S (1995) Likelihood-based inference in cointegrated vector autoregressive models. Oxford: Oxford University Press.

Johansen S and Schaumburg E (1999) Likelihood analysis of seasonal cointegration. Journal of Econometrics 88: 301-339.

Kunst RM and Franses PH (1998) The impact of seasonal constants on forecasting seasonally cointegrated time series. Journal of Forecasting 17: 109-124.

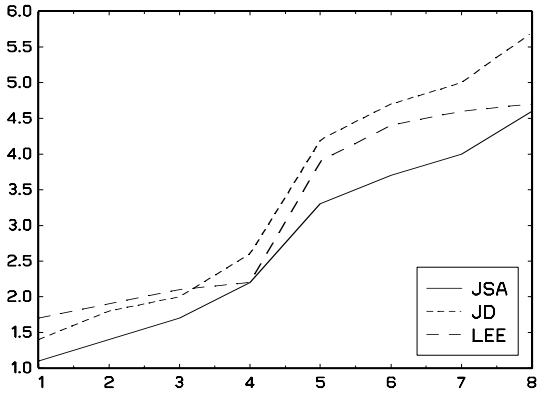
Kunst RM (1993) Seasonal cointegration, common seasonals, and forecasting seasonal series. Empirical Economics 18: 761-776.

Lee HS (1992) Maximum likelihood inference on cointegration and seasonal cointegration. Journal of Econometrics 54: 1-47.

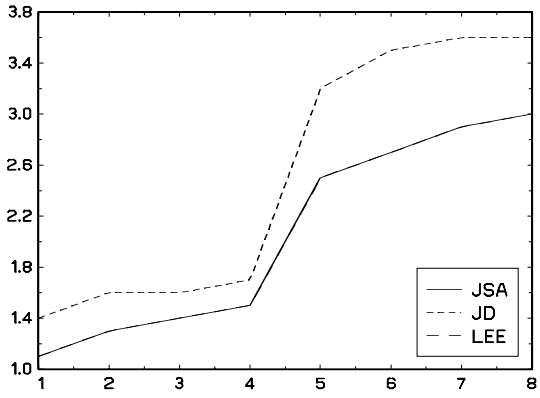
Lee HS and Siklos P (1995) A note on the critical values for the maximum likelihood (seasonal) cointegration tests. Economic Letters 49: 137-145.

Reimers H-E (1997) Forecasting of seasonal cointegrated processes. International Journal of Forecasting 13: 369-380.

A Figures

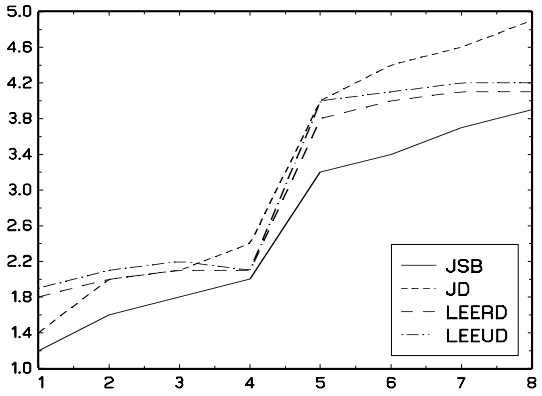


(a) Rank=1 at all frequencies, T=100.

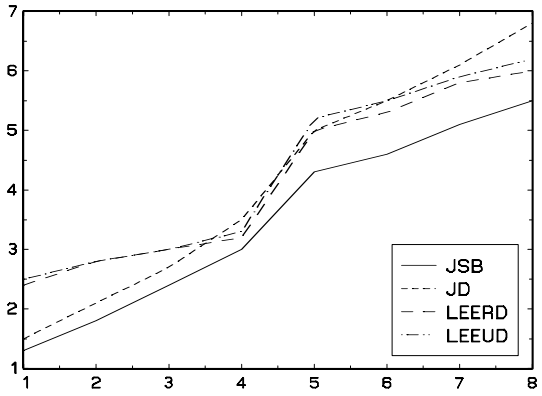


(b) Rank=1 at all frequencies, T=100.

Figure 1: Mean squered errors (in levels) over replicates and variables for k steps ahead forecasts. DGP based on estimates using (a) JSA specification and (b) Lee specification.

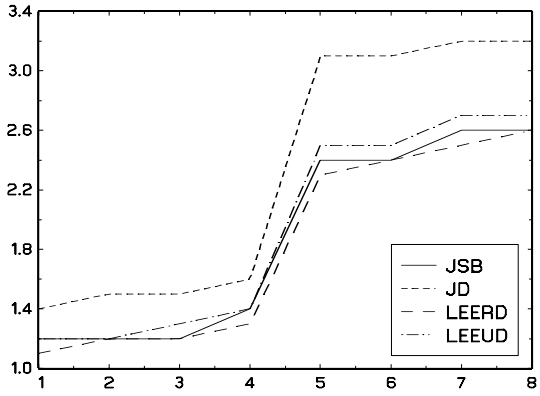


(a) Rank=1 at all frequencies, T=100.

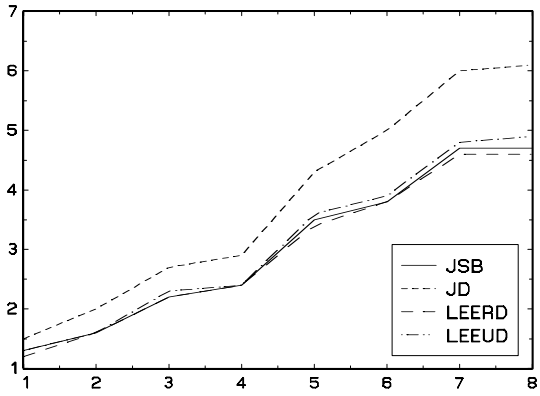


(b) Rank=2 at all frequencies, T=100.

Figure 2: Mean squered errors (in levels) over replicates and variables for k steps ahead forecasts. DGP based on estimates using the JSB specification.

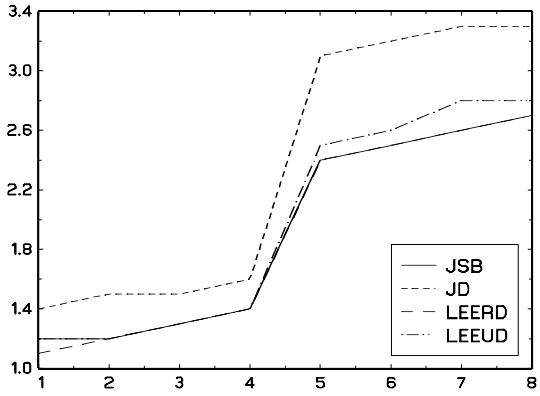


(a) Rank=1 at all frequencies, T=100.

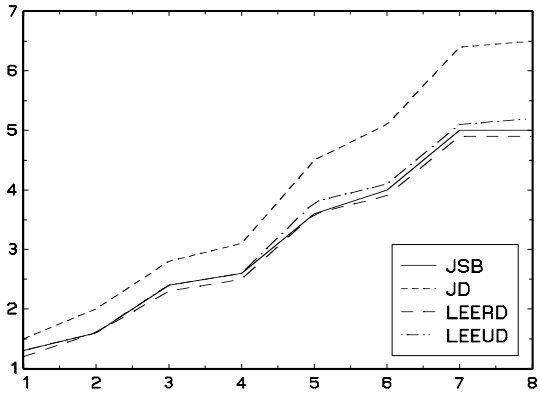


(b) Rank=2 at all frequencies, T=100.

Figure 3: Mean squered errors (in levels) over replicates and variables for k steps ahead forecsats. DGP based on estimates using the LEERD specification.

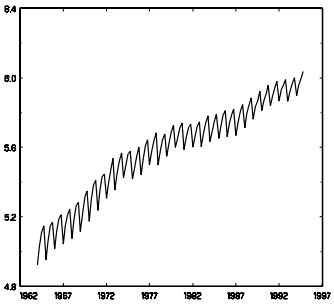


(a) Rank=1 at all frequencies, T=100.

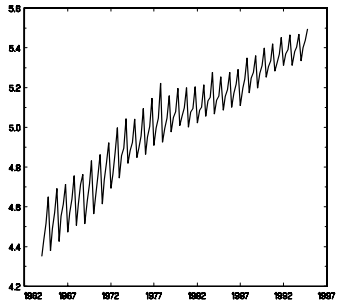


(b) Rank=2 at all frequencies, T=100.

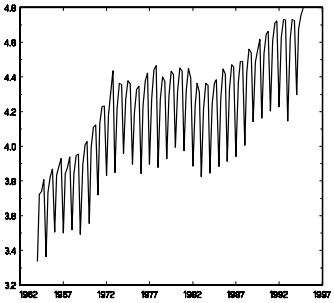
Figure 4: Mean squered errors (in levels) over replicates and variables for k steps ahead forecasts. DGP based on estimates using the LEEUD specification.



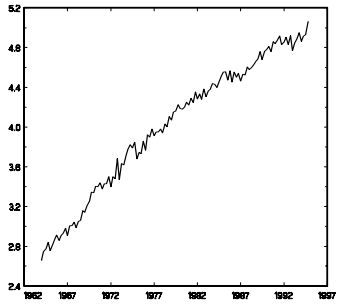
(a) Gross domestic product



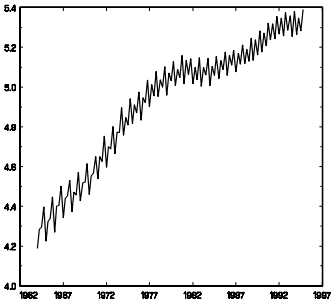
(b) Private consumption



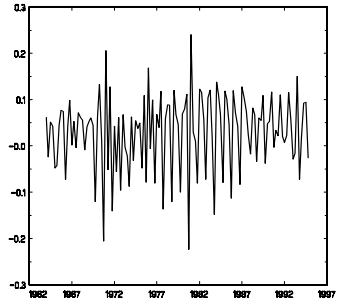
(c) Gross fixed investment



(d) Goods exports

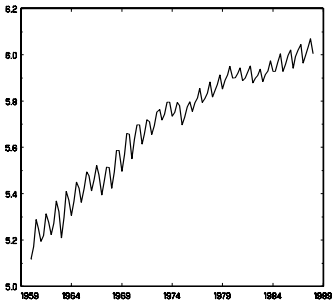


(e) Wages

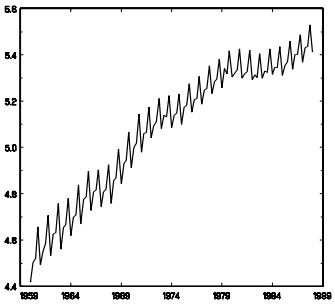


(f) Real interest rate

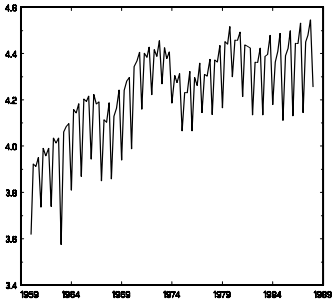
Figure 5: Austrian data, 1964:1-1994:4.



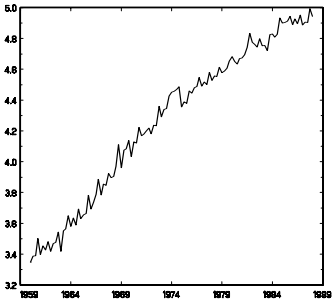
(a) Gross domestic product



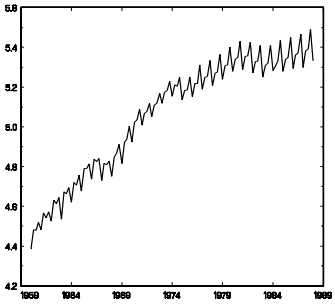
(b) Private consumption



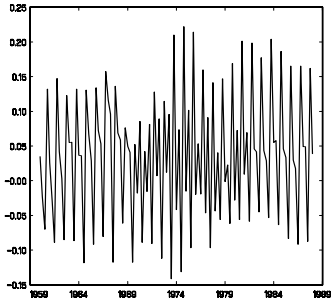
(c) Gross fixed investment



(d) Goods exports

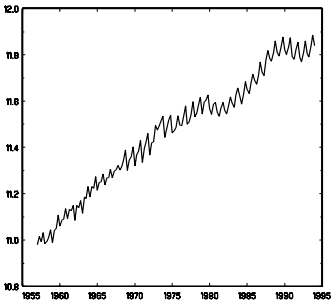


(e) Wages

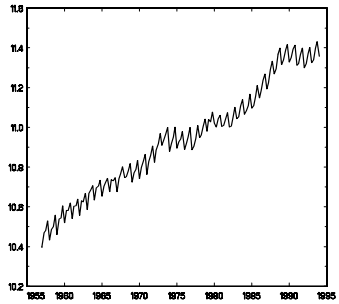


(f) Real interest rate

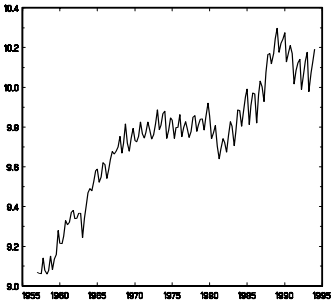
Figure 6: German data, 1960:1-1988:1.



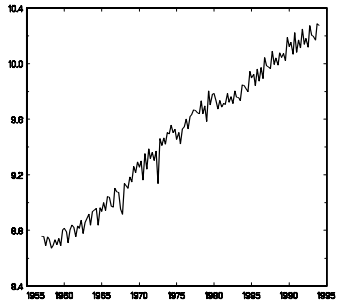
(a) Gross domestic product



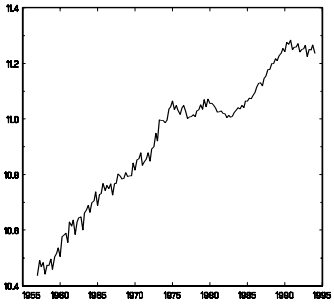
(b) Private consumption



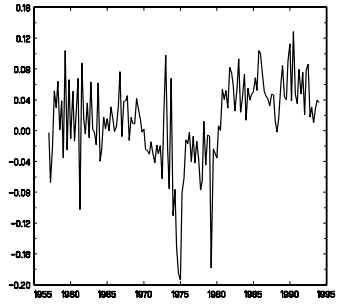
(c) Gross fixed investment



(d) Goods exports



(e) Wages



(f) Real interest rate

Figure 7: UK data, 1957:1-1994:1.