Are Interest Rates Responsible for Unemployment In the Eighties? A Bayesian Analysis of Cointegrated Relationship with a Regime Shift *

David de la Croix[1] Michel Lubrano[2]

February 1, 1996

Abstract

To what extent can the persistence of very high unemployment rates in most of the European countries be attributed to the presence of high real interest rates? This question, essentially addressed by the "customer market" price-setting school, was very much debated in Europe these last years. It is empirically analyzed for four European economies (Belgium, Denmark, France, Germany) and the USA in our paper. We use a bivariate cointegrating VAR model with one endogenous breaking point between unemployment and real interest rate. Within this model and devising a new Bayesian approach, the weak and strong exogeneity of the interest rate is tested. For the four European countries the model is shown to be cointegrating provided a break point is allowed. The four posterior densities of the breaking point are very similar, while the classical estimates give more divergent and counter-intuitive information. For the four countries, the real interest rate is weakly exogenous, providing support to the hypothesis of long run causation of interest rates on unemployment after 1974. Short term causation is verified for only three countries.

JEL Classification Numbers: E24, C11, C22

*Financial support from the SPES program E/89400153/PRO and from the Belgian program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming, are gratefully acknowledged. We are grateful to Frédéric Aprahamian, Luc Bauwens and Jean-Pierre Urbain, to the participants to the SPES program E/89400153/PRO and to the participants to seminars in GREQAM and CEPREMAP for their useful comments.

[1]National Fund for Scientific Research (Belgium) and IRES, University of Louvain, Place Montesquieu 3, B-1348 Louvain la Neuve, Belgium. Phone +32.10.47.34.53. E-mail: delacroix@ires.ucl.ac.be.

[2]GREQAM-CNRS, rue de la Charité 2, F-13002 Marseille, France. Phone: +33.91.14.07.45. E-mail: lubrano@ehess.cnrs-mrs.fr.

Introduction

The very high and persistent unemployment rates in Continental Europe at the end of the eighties remain a challenging question for macro-economists. The hopes generated by the decrease in unemployment observed in 1987-1989 have rapidly vanished and most of European countries are now experiencing the same unemployment rate as ten years before.

To what extent can the persistence of these very high unemployment rates be attributed to the presence of very high real interest rates? This is the key question addressed by this paper. The basic idea is very simple and is illustrated in Table 1. Average unemployment rates and real interest rates are computed by sub-periods. Mainly two competing explanations have been proposed in the macro-economic literature and will be discussed in section 1. Let us decompose the observed unemployment rate u_t into an equilibrium unemployment rate $\bar{u}_t$ and an adjustment toward the equilibrium $(u_t - \bar{u}_t)$.

- For one explanation, $\bar{u}_t$ has only slightly increased over the period; but u_t was driven away from $\bar{u}_t$ by important temporary shocks; the short term adjustment of u_t to $\bar{u}_t$ was impeded by hysteresis factors. From an econometric point of view this says that in an error correction model (ECM), the adjustment parameter is very low, nearly zero, implying an absence of co-integration and a model adjusted on differences of the variables.

- For the other type of explanation, $\bar{u}_t$ has changed over the period either because of permanent changes in economic variables affecting unemployment in the long run or because of structural breaks affecting the relationship between these economic variables and unemployment. In light of Table 1, the rise in real interest rates could be central in explaining $\bar{u}_t$.

To address the key questions generated by these two competing explanations of the rise of unemployment rates in Europe, we shall develop a bivariate co-integrating VAR model allowing for one endogenous breaking point. We shall first jointly test for the presence of co-integration and a regime shift. We shall then test for the weak and strong exogeneity of the interest rate. Inference results on four European countries and the United States will give indications on the validity of the second type of explanation for simultaneously high unemployment and high real interest rates in Europe.

Table 1

Unemployment rates and real interest rates

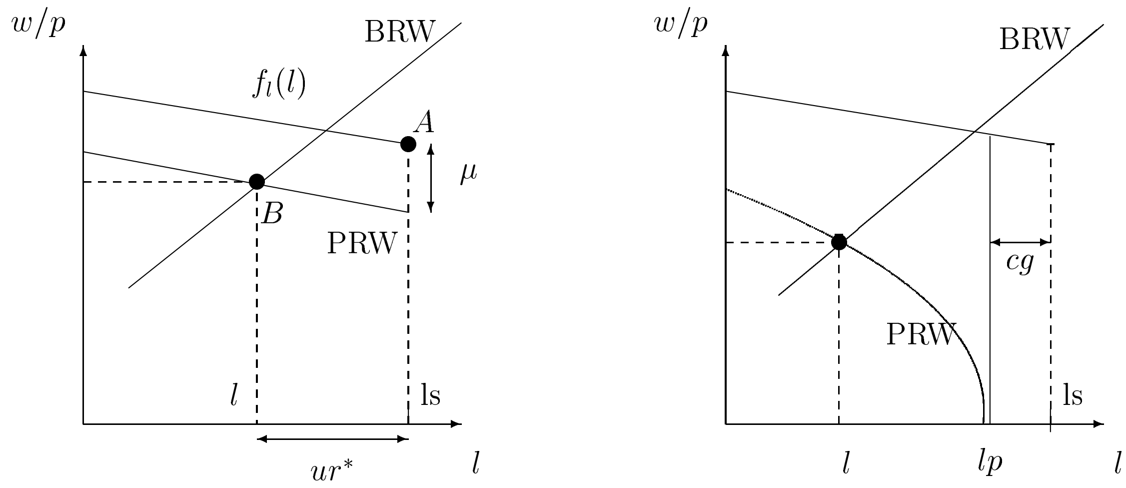
	Unemployment rates			Real interest rates		
	1965-1974	1975-1979	1980-1989	1965-1974	1975-1979	1980- 1989
Germany	1.0 %	3.7 %	6.8 %	4.1 %	3.3 %	4.7 %
France	1.9 %	4.7 %	9.3 %	2.5 %	1.0 %	5.3 %
Belgium	2.2 %	6.4 %	13 %	2.3 %	1.6 %	5.9 %
Denmark	1.4 %	6.4 %	8.9 %	2.6 %	4.6 %	7.4 %
Average	**1.6**%	**5.3** %	**9.5** %	**2.9** %	**2.6** %	**5.8** %
U.S.A.	4.6 %	7.0 %	7.3 %	1.1 %	1.0 %	4.9 %

The mechanisms through which this rise in real interests rates influences unemployment in a permanent way are described and discussed in section 1. Section 2 develops a cointegrated VAR model with a structural break and reports its classical analysis. In section 3, the Bayesian analysis of this model is presented together with a test of exogeneity / causality. The data set upon which the analysis will be carried is presented in section 4. The results of the tests and their interpretation for the link between real interest rates and unemployment are discussed in section 5. Section 6 concludes.

1 Real interest rates and unemployment

In this section, we discuss briefly the economic interpretations that can be behind the link between the interest rate and unemployment. To understand better their implications, Figure 1 displays a synthetic "textbook-like" model of unemployment in a non-competitive framework.[1] In the first figure, we have drawn, in the space of real wage w/p and labor

Figure 1: A synthetic model of unemployment



l, a constant labor supply ls and a negatively sloped demand for labor corresponding to the marginal productivity of labor with decreasing returns ($f_l(l)$). At the competitive equilibrium A the wage is equal to the marginal productivity of labor and unemployment is zero. A simple right-to-manage model of bargaining is introduced through the BRW curve.In this bargaining model, the union's desired real wage is a mark-up over some exogenous reference wage, with the mark-up decreasing in unemployment, so that the bargained real wage (BRW) curve is positively sloped: the higher the employment level,

[1]We do not pretend here to make a short survey of unemployment theories but simply to compare two "schools" between which it is particularly interesting to highlight some links with respect to the role of interest rates. Other ways of treating the link between interest rates and unemployment do exist, see for instance the empirical studies of Bierens and Broersma (1993) and Broersma (1992) who find a Granger causality from nominal interest rates to unemployment. Their results are interpreted in the light of the revenue maximization of Baumol which implies that interest costs (representing fixed costs) are important for the employment decisions of the firms.

the higher the wage claim. A third curve is derived from a price-setting behavior of the firm. In this case, the firm fixes its price as a mark-up μ over the marginal cost of labor, so that the firm's desired real wage falls below marginal productivity. This pricing behavior introduces a PRW curve (price-determined real wage) which is below the competitive demand for labor by a factor μ. At equilibrium B, employment is lower than in the perfect competition case. The equilibrium level of employment depends on the location of the BRW curve, which is mainly determined by union characteristics, and on the location of the PRW relation, which depends on the mark-up rate. The resulting unemployment, i.e. the equilibrium unemployment level ur^*, emerges from a "battle of the markups", meaning that it is the only unemployment rate that makes compatible the mark-up requirements of both agents. The second figure introduces a productive capacity[2] (or potential output) which allows a maximum employment of lp workers. lp is called the full-capacity employment. The difference between lp and ls, if any, is called the "capital gap" which measures the shortage of production capacities relatively to the capacities needed to obtain full-employment. With the presence of potential output, the firm now has to take into account the fact that it can be constrained by its capacity. This modifies the slope and the shape of the PRW curve. The new PRW curve is more negatively sloped because the mark-up rule takes into account the probability of a capacity constraint [see Sneessens (1987)], which is a negative function of employment. Equilibrium employment is lower than in the absence of a capacity constraint [cf. Arnsperger and de la Croix (1993)].

After this brief presentation of the model we now describe how it can be used to explain the rise of unemployment in the eighties.

The most widespread view of high unemployment in the eighties is the one summarized in Layard, Nickell and Jackman (1991): According to Layard et al. (1991), the equilibrium level of unemployment has itself increased a little bit due to an North-West shift of the BRW curve, linked with a too generous social policy. But this increase is not a significant cause of high unemployment. The very long plateau of high unemployment results from the persisting effects of two earlier temporary shocks (the second oil shock and the deflation policies started in 1979) kept alive by hysteresis factors. Therefore, actual unemployment rate lies quite above its equilibrium level. Using multi-country empirical analysis, the absence of adjustment is linked by Layard et al. to the duration of unemployment benefits.

These factors will hardly explain why unemployment still stays at these high levels if the long-run (equilibrium) unemployment rate has not been affected.[3] An alternative view defended by Fitoussi and Phelps (1986) puts more emphasis on the demand for labor side: they claim that the shift in long-term real interests rates has increased *the equilibrium level of unemployment itself* by affecting the optimal price-cost margin of the firms. The idea that high real interest rates are responsible for the slump in Europe in the 1980s has mainly been brought by Fitoussi and Phelps (1986) and (1988). Their view is based

[2]This productive capacity may result from the optimal choice of the capital stock by the firm under a putty-clay technology.

[3]Another criticism addressed to Layard et al. is that they emphasize the role of political and social factors which are important for union behavior rather than the role of economic mechanisms affecting the demand for labor (see e.g. the critical view presented by Phelps (1992): "The authors have taken a huge risk in throwing out nearly the whole corpus of general equilibrium theory in favor of a focus on some social and political parameters.").

on “customer-market” models: in brief, the firm faces a trade-off between today-profit and future profits *when it has to decide about prices*: If the firm chooses a high price-cost margin today (implying a high immediate profit) it will loose progressively its customers, leading to lower profits tomorrow. The extent to which the firm prefer today-profits with respect to future profits is a function of its discount factor, i.e. the expected real interest rate. If this rate increases, the weight of future profits is lowered, leading to increases in mark-up rates. “The sharp elevation of actual and, presumably, of expected real interest rates, we argue, induced firms in Europe to widen their mark-ups since it increased the opportunity cost of “investing” in greater or maintained market share through restraint in present prices at a sacrifice of present cash flow. There being no important demand stimulus to offset it, the result of the price push was a fall of employment in Europe.” say Fitoussi and Phelps (1986, page 498). Using Figure 1, the rise in interest rate increases the mark-up rate μ pushing down the PRW curve, leading to a drop in employment and in wages. Empirical evidence of mark-up rate increases is found in the declining labor share in Europe in the eighties.

According to the view of Drèze's (1991), the capital gap played an important role in explaining unemployment. If the capital gap increases, the PRW curve of the second figure goes to the South-West, leading to a decrease in employment and in wages. The reasons for this lack of productive capital is found by Lubrano et al. (1991) in the slowdown of world demand. However, de la Croix (1995) shows in a theoretical dynamic bargaining model with investment that the capital gap may depend crucially on real interest rates. The argument is not far from the one of Fitoussi and Phelps: a sharp elevation of expected real interest rates induces firms to reduce their capital stock since it increases the opportunity cost of investing in future capacities through a sacrifice of present cash flow. This reduction in capacities leads to more unemployment. As it is clear now, the capital gap story and the customer market story lead to the same kind of effect of interest rate on wages and employment. They both rely on a trade-off between future and present profits.

The empirical analysis that we will conduct on the link between interest rates and unemployment will not “test” the above models but is aimed at analyzing whether interest rates may play a role in explaining the rise in unemployment and its persistence to very high levels. We want to analyze a bivariate relationship between two variables which will turn out to be nearly integrated of order one or can be assumed to be $I(1)$ on a finite sample. There is a stable relationship between these two variables if they are cointegrated. Once co-integration is achieved, a causal ordering between the two variables may be assumed if one of them is exogenous (see Urbain 1993). It suffices then to make inference in a conditional error correction equation. The potential relationship between equilibrium unemployment and long-run real interest rates

$$ur_t^* = f(r_t^*) \qquad\qquad f' > 0 \tag{1}$$

is conditional on all other variables which determine the position of the PRW and BRW curves in Figure 1. If the long-run relationship (1) between ur and r is not stationary on the sample period, equilibrium (long-run) unemployment is affected by at least one missing variable in addition to the interest rate. As shown in Table 1 of the introduction, the relationship between unemployment and real interest rate has no chance of being stable over the whole period. Moreover, if the role of the interest rates may be prominent in the eighties (for the sole reason that they were very high) their influence may be low or negligible before. This suggests modeling a long-run relationship with a structural break of the form:

$$ur_t^* = f(r_t^*) + ID(\tau)g(r_t^*) \qquad\qquad g' > 0 \tag{2}$$

where τ is the date of the regime shift and $ID(\tau)$ is the indicator function defined by:

$$\begin{aligned} ID(\tau) &= 1 \qquad\qquad \text{if } t > \tau \\ ID(\tau) &= 0 \qquad\qquad \text{otherwise.} \end{aligned} \tag{3}$$

Three questions naturally arise from this way of modeling. The first is whether the allowance made for one regime shift is sufficient to make the relationship cointegrating, τ being determined endogenously. This is tested in the next sections. The second question concerns the singleness of the regime shift. Since our methodology will provide a discrete density of the breaking point, it will give some indications on the number of breaks (e.g. if there is more than one date with the same high probability, we may suspect the presence of multiple regime shifts). The last question is, supposing that (2) cointegrates, what is the interpretation to give to the regime shift? If (2) cointegrates, the unique regime shift together with the movements in the interest rate is sufficient to obtain a stationary long-run relationship, i.e. is sufficient to model the evolution of equilibrium unemployment. Using Figure 1, it corresponds to a shift either in the PRW curve or in the BRW curve. Following the idea of Perron (1989), the oil price shock of 1973 has had a permanent effect on the various macroeconomic variables; in our case, an oil shock may shift down the PRW curve by reducing the productivity of value-added and therefore the one of labor [see e.g. Pfann and Palm (1993)]. Such shock may also shift up the BRW curve by affecting the gap between the consumption price index in which workers are interested in and the producer price index. Linked with the 1973 crisis, the change in the rate of growth of world demand as well as the price of energy may also affect the PRW curve, leading to permanent increase in unemployment. Whether these variables affect the existence of the regime shift and/or the role of interest rate will be discussed at the end of section 5.

2 Co-integration relations with a regime shift

We shall first present a bivariate co-integration VAR model and then show how an exogeneity assumption open the passage between this type of model and a structural conditional error correction model. We shall finally introduce a break in the structural equation and review a classical method to estimate this equation.

2.1 Cointegrating VAR and Structural ECM

Let us consider a bivariate cointegrating VAR model between two I(1) variables y_t and x_t in the notations of Engle and Granger (1987):

$$A(L)\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = -\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}' z_{t-1}(\beta) + \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} \tag{4}$$

where $A(L)$ is a lag matrix polynomial of order p, z_t is the equilibrium vector with:

$$z_t(\beta) = y_t - x_t\beta \tag{5}$$

and v_t is an error term which is serially uncorrelated and normally distributed with zero mean and variance covariance matrix Σ:

$$v_t \sim N(0, \Sigma) \qquad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \tag{6}$$

For the ease of exposition, we shall suppose that $A(L)$ is of degree $p = 1$:

$$A(L) = I_2 - \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} L$$

A structural ECM is obtained by multiplying the model by a matrix C designed as:

$$C = \begin{bmatrix} 1 & -\sigma_{12}\sigma_{22}^{-1} \\ 0 & 1 \end{bmatrix}$$

which leads to a parameterization in term of the conditional distribution of $y_t|x_t$ and the complementary marginal distribution of x_t. If we shorten the notations of the model into:

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} \sim N\left(\begin{bmatrix} m_y \\ m_x \end{bmatrix}, \Sigma \right)$$

the conditional and marginal models are written:

$$\begin{bmatrix} \Delta y_t - \sigma_{12}\sigma_{22}^{-1}\Delta x_t \\ \Delta x_t \end{bmatrix} \sim N\left(\begin{bmatrix} m_y - \sigma_{12}\sigma_{22}^{-1} m_x \\ m_x \end{bmatrix}, \begin{bmatrix} \sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{21} & 0 \\ 0 & \sigma_{22} \end{bmatrix} \right) \tag{7}$$

Defining $c = \sigma_{12}/\sigma_{22}$ the conditional and marginal expectations of y_t and x_t are:

$$\begin{aligned} m_y - c\, m_x &= (a_{11} - c\, a_{21})\Delta y_{t-1} + (a_{12} - c\, a_{22})\Delta x_{t-1} - (\alpha_1 - c\, \alpha_2) z_{t-1}(\beta) \\ m_x &= a_{21}\Delta y_{t-1} + a_{22}\Delta x_{t-1} - \alpha_2 z_{t-1}(\beta) \end{aligned} \tag{8}$$

The variable x_t is exogenous for β if the conditional and the marginal distributions can be analyzed separately which is the case if there is no cross restrictions between the two equations (see Engle, Hendry and Richard (1983) for a more detailed definition of exogeneity). In the bivariate model (7)-(8) there is no cross restriction if $\alpha_2 = 0$ (see for instance Boswijk (1991), Johansen (1992) or Urbain (1992)). The structural error correction model (9) below can be analyzed separately for inference and tests:

$$\Delta y_t = c\,\Delta x_t + (a_{11} - c\, a_{21})\Delta y_{t-1} + (a_{12} - c\, a_{22})\Delta x_{t-1} - \alpha_1 z_{t-1}(\beta) + \epsilon_t \tag{9}$$

with $\epsilon_t = v_{1t} - c\, v_{2t}$. In this model, there is no co-integration if $\alpha_1 = 0$.

In the case of exogeneity, the marginal process of x_t is reduced to:

$$\Delta x_t = a_{21}\Delta y_{t-1} + a_{22}\Delta x_{t-1} + v_{2t} \tag{10}$$

x_t is strongly exogenous if y_t does not help in any way to predict x_t, which means here that $a_{21} = 0$.

When there is exogeneity of x_t it is simpler to analyze (9) than (4) [see Johansen(1992)]. Moreover, in a CVAR model, all the variables play a symmetrical role in the long run solution. With a condition of exogeneity such as $\alpha_2 = 0$, this symmetry is broken and some authors like Urbain (1993) have interpreted this condition as a kind of long run causality. The Granger causality condition $a_{21} = 0$ becomes therefore a condition of short-run causality (see also the discussions in Toda and Phillips (1993) and Mosconi and Giannini (1992)).

2.2 Modeling a break in the long term relation

Using (3) a regime shift in the long run cointegrating relationship is modeled as:

$$z_t(\beta, \tau) = y_t - \beta_1 x_t - \beta_2 x_t ID(\tau) \tag{11}$$

recalling that τ is a discrete parameter indicating the date of the shift. A regime shift must be also allowed for in the short run so the complete[4] error correction model may be written as:

$$\begin{aligned} \Delta y_t = \gamma_0 \Delta x_t &+ \gamma_1 \Delta(x_t ID(\tau)) + (\rho - 1) z_{t-1}(\beta, \tau + 1) \\ &+ \delta_1 \Delta y_{t-1} - \mu_1 \Delta x_{t-1} - \mu_2 \Delta(x_{t-1} ID(\tau + 1)) + \epsilon_t \end{aligned} \tag{12}$$

Conditionally on the exogeneity of x_t, a first fundamental question is if there is co-integration and if modeling a break is necessary to achieve co-integration. We shall first review in the next paragraph a classical approach of the question and then propose a full Bayesian analysis in the next section.

2.3 A classical test

Gregory and Hansen (1992) have proposed a co-integration test along the lines of Engle and Granger (1987) which takes into account the possibility of a structural break. The main question in a classical framework is to derive and tabulate the distribution of the test statistics under the null and eventually give an indication of its power.

The starting point is a static regression as defined in (11) which gives for each different value of τ an estimate of $z_t(\hat{\beta}, \tau)$. Then for each τ an AEG test[5] is computed. The final value of τ is chosen so as to give the minimum value to the AEG test. Critical values for the test are -4.95 at 5% and -4.68 at 10% when there are two variables, a constant and no trend.

3 Bayesian co-integration tests with a structural break

The above described classical approach for testing co-integration in the presence of a structural break has two drawbacks:

- as it is well-known [Kremers et al (1992)], the AEG test imposes a common factor restriction which distorts the size of the test and reduces its power if there is no common factor restriction in the data generating process.
- the procedure of Gregory and Hansen (1992) violates the likelihood principle as τ is not determined by the likelihood function of the model, but by the value of the test statistics. Moreover the value of the AEG test is dependent on the chosen number of lags p which should be adjusted for each value of τ during the grid search.

---

[4]written here with $p = 1$ just for the ease of notations. Note the parameterization $(\rho - 1)$ instead of α_1 which is adopted to make the parallel with the models used for testing for unit roots.

[5]Augmented Engle Granger test, the co-integration test proposed by Engle and Granger (1987) with additional lags introduced to cope with remaining autocorrelation.

A Bayesian analysis of the error correction model (12) should answer these two criticisms.

3.1 Local identification problems

Let us first consider equation (12) which represents the generic form of our error correction model. It is easily seen that conditionally on ρ and τ the model is linear when $p = 1$. It is thus convenient to conduct the analysis conditionally on these two parameters and numerically integrate them out at the end.

The parameter ρ influences greatly the shape of the likelihood function. It takes its values between 0 and 1. As it enters the model in a multiplicative way with $(\rho - 1)\beta$ there is an identification problem at the point $\rho = 1$. Moreover as shown in O'Brien (1970), it becomes difficult to make inference on β as soon as ρ is greater than 0.85. Various approaches are possible to tackle this problem which, we think, are not very satisfactory:

- we can decide to impose a common factor restriction with $\beta = \gamma$ in the generic model $\Delta y_t = \gamma \Delta x_t + (\rho - 1)(y_{t-1} - x_{t-1}\beta) + \epsilon_t$. There is no longer an identification problem at $\rho = 1$, but if the data reject this restriction, the model is misspecified.
- In a Bayesian framework, the natural way of solving an identification problem is to introduce a prior probabilistic information (which is not going to be revised by the sample). Kleiberger and van Dijk (1994) take advantage of the very special shape of their Jeffreys' prior to skip the problem of identification. When the long term regression (11) contains a constant term, the Jeffreys' prior is zero for $\rho = 1$ (see the appendix). But it is easy to show that this property disappears when the constant term is put outside the long term solution (see again the appendix). Moreover this prior favors too much the cointegrated case.

In this paper, we prefer to devise another route. The identification problem comes from the simultaneous inference made on ρ and β. If we manage to separate the inference on these two parameters, the identification problem disappears. In the classical literature, there is a method which uses a similar approach: the two step method of Engle and Granger (1987). We shall show below how a simple linearization of (12) helps build a Bayesian equivalent of the two step approach of Engle and Granger[6].

In the simple model $\Delta y_t = \gamma \Delta x_t + (\rho - 1)(y_{t-1} - x_{t-1}\beta) + \epsilon_t$, let us consider the product $(\rho - 1)\beta$ and its first order Taylor expansion around the, for the while unspecified, values $\hat{\rho}$ and $\hat{\beta}$:

$$(\rho - 1)\beta \simeq (\hat{\rho} - 1)\hat{\beta} + (\rho - \hat{\rho})\hat{\beta} + (\beta - \hat{\beta})(\hat{\rho} - 1)$$

For $\hat{\rho} = 1$ (hypothesis of no co-integration) the linearization reduces to $(\rho - 1)\beta \simeq (\rho - 1)\hat{\beta}$ and the model becomes $\Delta y_t = \gamma \Delta x_t + (\rho - 1)(y_{t-1} - x_{t-1}\hat{\beta}) + \epsilon_t$. This suggests a "long term" model $y_t = x_t\beta + u_t$ to make inference on β and a "short term" model from which

[6]We can think of other cases in the literature where linearizing solves a problem of identification: for instance testing for linearity in the non-linear LSTAR model $y_t = \beta' z_t + F(z_t, \gamma)\theta' z_t + u_t$. The parameter θ is not identified under the null hypothesis $\gamma = 0$, but is identified under the alternative $\gamma \neq 0$. An approach adopted in Luukkonen *et al* (1988) consists in linearizing $F(z_t, \gamma)$ around $\gamma = 0$ to solve the identification problem. Of course, we cannot deny the fact that linearizing violates a bit the likelihood principle. But let us try not to be too dogmatic.

inference on ρ (but also on γ) can be drawn. This last model will be mainly used for testing for co-integration (it does not impose any common factor restriction). Our case is a bit more complex due to the presence of the endogenous break-point. The "long term" model is non- linear in τ:

$$y_t = \beta_1 x_t + \beta_2 x_t ID(\tau) + u_t \tag{13}$$

and will be used for inference both on β and τ. It is convenient to define:

$$x_t'(\tau) = [x_t', x_t' ID(\tau)] \tag{14}$$

and write the likelihood function (assuming that u_t is Normally distributed with zero mean and variance ω^2):

$$L(y; \beta, \tau, \omega^2) \propto \omega^{-T} \exp -\frac{1}{2\,\omega^2} \sum_{t=1}^{T} (y_t - x_t'(\tau)\beta)^2 \tag{15}$$

Under a diffuse prior for β and ω^2, the conditional posterior density of β is Student:

$$D''(\beta|\tau, y) = f_t(\beta|\beta_*(\tau), M_*(\tau), s_*(\tau), T) \tag{16}$$

together with:

$$\begin{array}{rcl} M_*(\tau) & = & \sum x_t(\tau)\, x_t'(\tau) \\ \beta_*(\tau) & = & M_*^{-1}(\tau) \sum x_t(\tau)\, y_t \\ s_*(\tau) & = & \sum y_t^2 - \beta_*'(\tau)\, M_*(\tau)\, \beta_*(\tau) \end{array} \tag{17}$$

The posterior density of τ obtains as one over the integrating constant of the above Student density times the prior density of τ:

$$D''(\tau|y) \propto |s_*(\tau)|^{-T/2}\, |M_*(\tau)|^{-1/2}\, D'(\tau) \tag{18}$$

By numerical integration of this posterior density we get the posterior expectation τ noted in the sequel $\hat{\tau}$. The marginal posterior in τ is discrete. Each point of this marginal can be interpreted as the probability that τ takes a particular value. So a model with break point will dominate a model with no break if a point has a probability greater than 0.5.

The posterior moments of β are obtained by numerical integration using the traditional formulae for conditional expectations and variances derived from (16). So:

$$\mathrm{E}(\beta|y) = \int \mathrm{E}(\beta|y, \tau)\, D''(\tau|y)\, d\tau \tag{19}$$

In the sequel, we shall note this posterior expectation $\hat{\beta}$. The marginal variance of β is a bit more complicated to compute as:

$$\mathrm{Var}(\beta|y) = \mathrm{E}_\tau[\mathrm{Var}(\beta|y, \tau)] + \mathrm{Var}_\tau[\mathrm{E}(\beta|y, \tau)] \tag{20}$$

These computations are decomposed into:

$$\mathrm{E}_\tau[\mathrm{Var}(\beta|y, \tau)] = \int \mathrm{Var}(\beta|y, \tau)\, D''(\tau|y)\, d\tau \tag{21}$$

$$\mathrm{Var}_\tau[\mathrm{E}(\beta|y, \tau)] = \int \mathrm{E}(\beta|y, \tau)\, \mathrm{E}(\beta|y, \tau)'\, D''(\tau|y)\, d\tau - \mathrm{E}(\beta|y)\, \mathrm{E}(\beta|y)' \tag{22}$$

Let us now consider the "short term" model (first for $p = 0$) which is (12) where β and τ have been replaced by their expectation:

$$\Delta y_t = \gamma_0 \Delta x_t + \gamma_1 \Delta(x_t ID(\hat{\tau})) + (\rho - 1) z_{t-1}(\hat{\beta}, \hat{\tau} + 1) + \epsilon_t \tag{23}$$

The main use of this model will be testing for co-integration. It is linear in ρ, but we want to be free to use any type of prior density for this parameter. So we shall consider this model as non-linear in ρ. When $p > 0$ the model has $p(2\,k+1)$ extra regressors which are here to cope with remaining autocorrelation in the residuals. The p extra lags in (12) can be written in the following form (for $p = 1$):

$$\delta_1 \Delta(y_{t-1} - \lambda_1 x_{t-1} - \lambda_2 x_{t-1} ID(\hat{\tau} + 1))$$

with $\lambda_1 = \mu_1/\delta_1$, $\lambda_2 = \mu_2/\delta_1$. Provided we accept to impose the restriction $\lambda = \beta$, we can replace the p extra lags by p lags of $\Delta z_t(\hat{\beta}, \hat{\tau})$.

Let us define the following notations:

$$\begin{aligned} y_t(\rho) &= \Delta y_t - (\rho - 1)\, z_{t-1}(\hat{\beta}, \hat{\tau} + 1) &(24)\\ \tilde{x}'_t &= [\Delta x_t, \Delta(x_t ID(\hat{\tau})), \Delta z_{t-1}(\hat{\beta}, \hat{\tau} + 1), \Delta z_{t-2}(\hat{\beta}, \hat{\tau} + 2), \cdots] &(25)\\ \theta' &= [\gamma_0, \gamma_1, \delta_1, \delta_2, \cdots] &(26) \end{aligned}$$

The likelihood function of the "short term" model becomes:

$$L(\Delta y; \hat{z}_{t-1}, \theta, \rho, \sigma^2) \propto \sigma^{-T} \exp \frac{-1}{2\,\sigma^2} \sum_{t=p+2}^{T} (y_t(\rho) - \tilde{x}'_t \theta)^2 \tag{27}$$

The conditional posterior density of θ has also a Student shape and is:

$$D''(\theta \mid y, \rho) = f_t(\theta \mid \theta_*(\rho), N_*, ss_*(\rho), T - p - 2) \tag{28}$$

with:

$$\begin{aligned} N_* &= \sum \tilde{x}_t \tilde{x}'_t \\ \theta_*(\rho) &= N_*^{-1} \sum \tilde{x}_t y_t(\rho) \\ ss_*(\rho) &= \sum y_t^2(\rho) - \theta'_*(\rho) N_* \theta_*(\rho) \end{aligned} \tag{29}$$

The posterior density of ρ obtains as before:

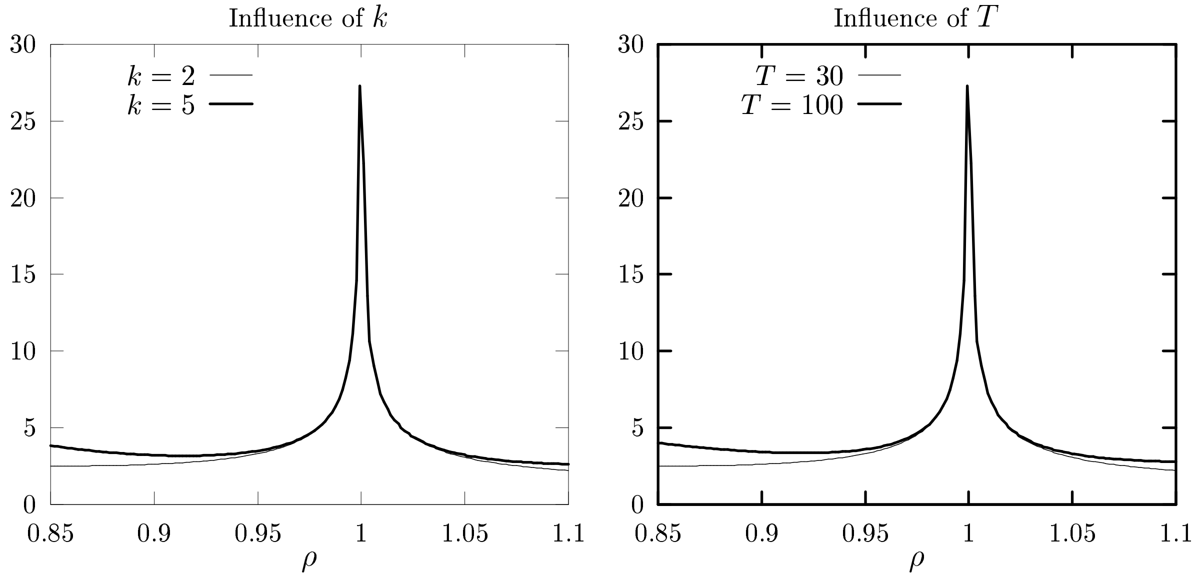
$$D''(\rho \mid y) \propto [ss_*(\rho)]^{-(T+p-1)/2} . D'(\rho) \tag{30}$$

This marginal distribution is continuous. To decide if there is co-integration or not, we shall compute a posterior confidence interval for ρ and say that there is no co-integration if the value 1 belongs to this interval. We can alternatively compute $\Pr(\rho \geq 1 \mid y)$ and compare this value to a pre-assigned probability value like 5%.

3.2 Choosing a prior for testing

One of the primary reason for analyzing model (12) is to test for co-integration. The choice of an appropriate prior on ρ is then important. We know that in the generic

Figure 2: Beta-Jeffreys' prior



autoregressive process $y_t = \rho y_{t-1} + \epsilon_t$ the sampling properties are different when ρ is lower than one, equal to one and greater than one. The Jeffreys' principle (which makes use of the information matrix) can be used to define a prior on ρ which may reflect these properties. We show in the appendix that many options exist to define such a prior. We have selected a prior which is inspired and adapted from Berger and Yang (1994). It is based on the information matrix of the regression model with autocorrelated errors and k I(1) regressors. The information matrix is computed for the stationary case ($\rho < 1$) and extended to the nonstationnary case by symetrization via the transformation $\tilde{\rho} = 1/\rho$. From the appendix we get:

$$D'(\sigma^2) \propto \frac{1}{\sigma^2} \tag{31}$$

$$D'(\beta|\sigma^2) \propto \frac{1}{\sigma} \tag{32}$$

$$D'(\rho|\beta,\sigma^2) \propto \left[\frac{1}{1-\rho^2}[1+\frac{1}{2}(T-1)(1-\rho)^2]^k\right]^{1/2} \qquad \text{if } |\rho| < 1 \tag{33}$$

$$D'(\rho|\beta,\sigma^2) \propto \frac{1}{|\rho|^{k+1}}\left[\frac{1}{\rho^2-1}\left[\rho^2+\frac{1}{2}(T-1)(1-\rho)^2\right]^k\right]^{1/2} \qquad \text{if } |\rho| > 1 \tag{34}$$

This prior displays quite interesting properties as shown in figure 2. It reckons the fact that something special happens at $\rho = 1$, but may seem rather flat outside that value. However the tails are inflated by an increase of k or T.

The prior on τ has to be derived in a different way as the likelihood function is not differentiable in this parameter. It is chosen to be a discrete uniform probability distribution over a finite interval of dates:

$$\begin{cases} D'(\tau) \propto \dfrac{1}{\mathrm{Card}[a,b]} & \text{if } \tau \in [a,b] \\ D'(\tau) = 0 & \text{otherwise} \end{cases} \tag{35}$$

where Card$[a,b]$ means the number of elements in the discrete interval $[a,b]$. Any informative variation over this probability distribution can be accommodated of course.

3.3 Testing for exogeneity and causality

Exogeneity properties are analyzed with the marginal model while co-integration was tested with the conditional model. Introducing a break in the marginal model gives (for $p=1$):

$$\Delta x_t = \varphi_0 \Delta x_{t-1} + \varphi_1 \Delta(x_{t-1} ID(\hat{\tau}+1)) + \varphi_2 \Delta y_{t-1} - \alpha_2 z_{t-1}(\hat{\beta}, \hat{\tau}+1) + v_t \tag{36}$$

Weak exogeneity for β[7] (and consequently long term causality as interpreted in Urbain (1993)) needs $\alpha_2 = 0$. Granger (or short term) non-causality is achieved for $\varphi_2 = 0$. Both tests are easily obtained as the model is linear. They are based on the posterior density of these two parameters computed under a non informative prior. With the Normality assumption made in (6) the posterior density of α_2 is Student and a Student table can be used to test for $\alpha_2 = 0$. When $p > 1$ testing for short term causality involves a multivariate Student posterior density as φ_2 is now of dimension p. As shown in Zellner (1971, p. 385), $(\varphi_2 - \mathrm{E}(\varphi_2|y))'\mathrm{Var}(\varphi_2|y)^{-1}(\varphi_2 - \mathrm{E}(\varphi_2|y))/p$ is distributed as a Fisher with p and $(T-3p-1)$ degrees of freedom. Consequently a Fisher table can be used to test for short term causality.

4 The data set and its sample characteristics

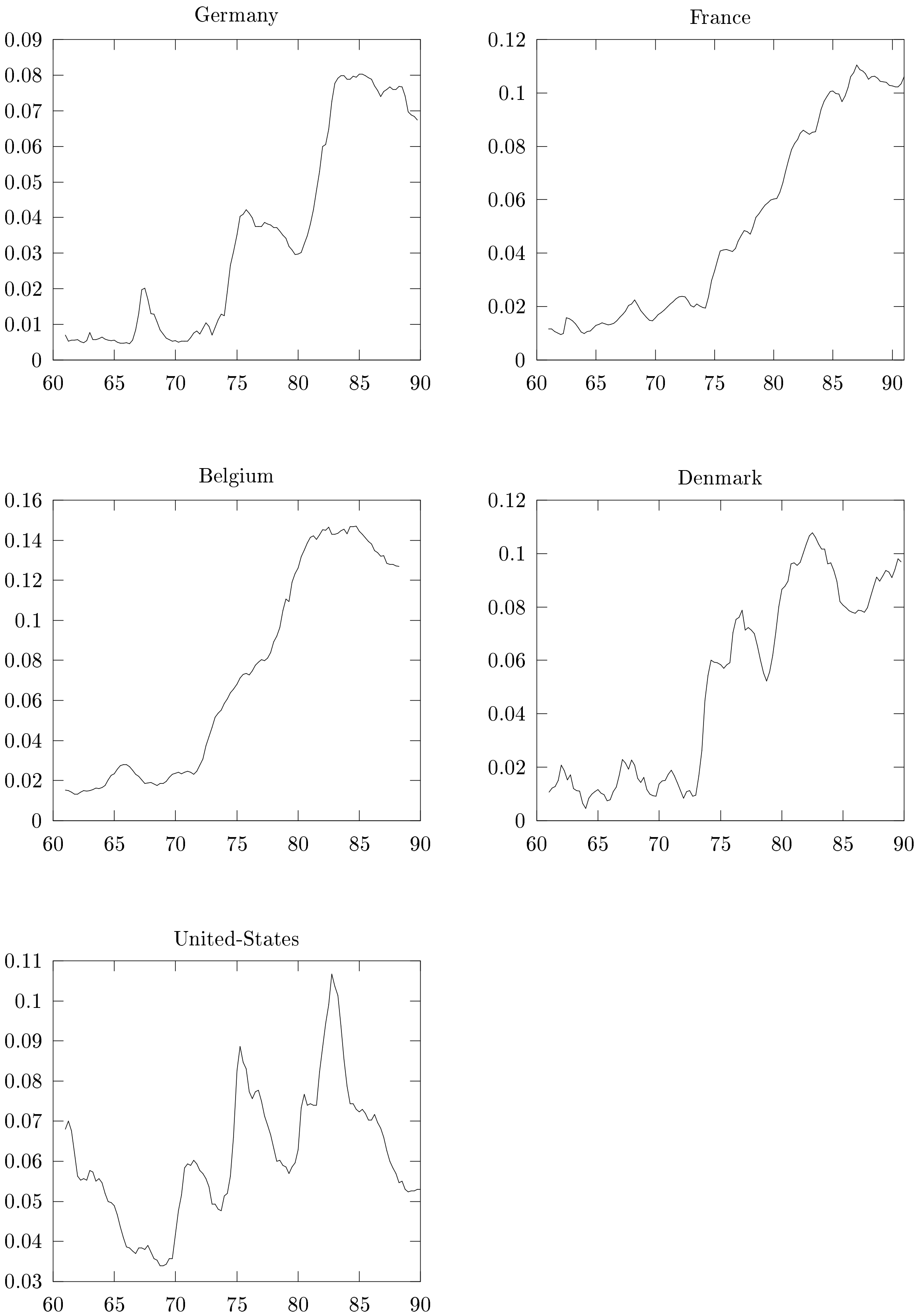
Before testing for the long run relationship (2), it is necessary to justify the reason for testing for co-integration between bounded variables like the unemployment rate and to a certain extent the rate of interest. If a variable is bounded, it cannot be $I(1)$ since an integrated variable has no bounded mean or variance. However this is true only if we consider a sample of an infinite length. In a finite sample a bounded series may look very much like an $I(1)$ series, provided its auto-regressive coefficient is greater than say 0.95. In this case, the series is better predicted if the forecasting model imposes the unit root (see Campbell and Perron (1991)). Consequently if a test does not reject the null of a unit root for a bounded series, we shall consider it as a locally $I(1)$ series.

4.1 The data

Our quarterly data concern four European countries covered by the SPES program "Wage Formation and Unemployment Persistence" plus the USA. The Netherlands have been

[7] conditionally on τ!

Figure 3: Unemployment Rates



excluded because their unemployment quarterly data cover too small a period. We have the following sample sizes:

Table 2
Data sets

country	period
Belgium	1963:3 1990:4
Denmark	1962:1 1990:4
France	1961:1 1991:4
Germany	1961:1 1989:4
USA	1961:1 1990:2

Some of the series were seasonally adjusted and are available only under this form. Thus we had to adjust the remaining ones. The real interest rate is defined as $r = i - \Delta_4 \log Pc$, where i in the long-term nominal interest rate and Pc the price index of consumption goods and $\Delta_4 \log Pc$ the annual rate of inflation.

The relationship in de la Croix (1995) linking unemployment and interest rates through the capital gap is highly non-linear. That is an argument for using variables taken in logarithm so as to log-linearize the relationship (2)[8]. A second reason for using logarithms is related to the above discussion on bounded variables. As our unemployment series takes value between 0 and 1 their logarithm is unbounded at least at their lower end[9]. A graph of each series is given in Figures 3 and 4.[10]

4.2 Unit root tests

We shall perform two types of unit root tests: a classical ADF test and a Bayesian unit root test along the lines developed in Lubrano (1995). The asymptotic critical values for the ADF test and for the AEG co-integration test for two variables are:

Table 3
Critical values for ADF and AEG tests

size	constant	trend	co-integration for $n = 2$
1%	-3.434	-3.964	-3.900
5%	-2.862	-3.413	-3.338
10%	-2.567	-3.128	-3.046

These asymptotic values are drawn from MacKinnon (1991). The Dickey-Fuller regression is:

$$\Delta y_t = \rho y_{t-1} + \beta_0 + \beta_1 t + \sum_{i=1}^{p-1} \alpha_i \Delta y_{t-i} + \epsilon_t$$

We selected the number of extra lags by the Schwarz criterion. The results are presented in Tables 4 and 5.

[8]For the interest rate, it is necessary to take the logarithm of $1+r$ because of the negative values of r.
[9]For more on bounded variables, see Wallis (1987).
[10]Belgian data have been adjusted to take into account the old unemployed workers which have been subtracted from official statistics after 1985.

Figure 4: Real Interest Rates

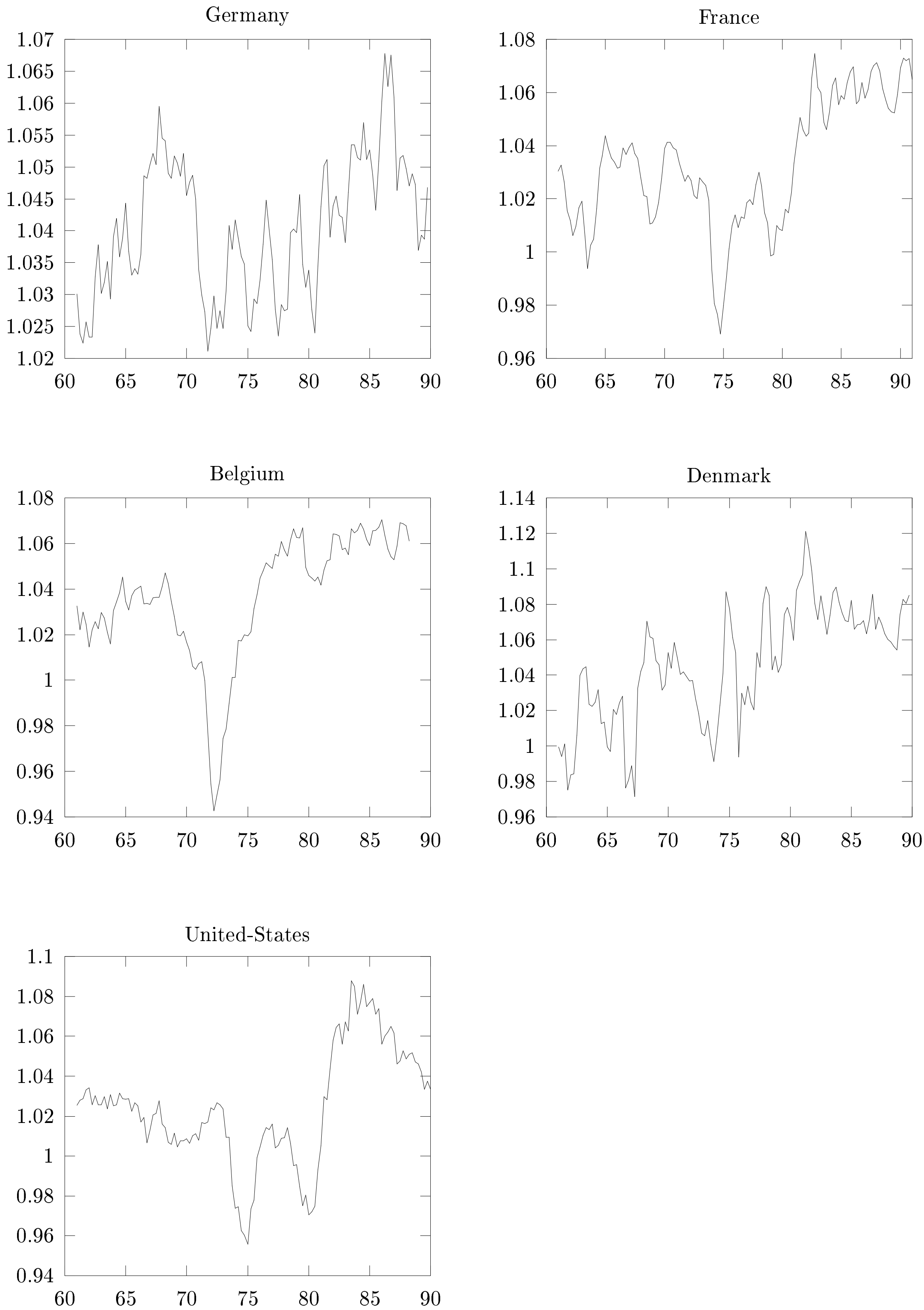


Table 4
Unit root tests for $\ln ur$

country	n lags	trend	ADF	$E(\rho\|y)$	$Pr(\rho < 1\|y)$
BE	1	yes	-1.535	0.988	0.929
DK	1	yes	-2.966	0.932	0.965
FR	1	yes	-3.070	0.949	0.957
GY	5	yes	-2.844	0.941	0.966
US	1	no	-2.349	0.975	0.969

If we take a posterior probability level of 97%[11], the unemployment rate series taken in logarithms all present a unit root according either to the classical or to the Bayesian tests.

Table 5
Unit root tests for $\ln(1 + r)$

country	n lags	trend	ADF	$E(\rho\|y)$	$Pr(\rho < 1\|y)$
BE	1	no	-1.789	0.968	0.924
DK	4	no	-2.502	0.954	0.787
FR	4	no	-1.393	0.983	0.756
GY	4	no	-2.641	0.919	0.922
US	2	no	-2.106	0.953	0.971

The picture is not clear cut with real interest rates. At the 10% level, the series for Denmark and Germany have no unit root. The Bayesian criterion to the contrary indicates that these series have a unit root. But it questions the presence of a unit root for the US series.

Besides the unit root tests, we may also look at the value of ρ implied by the regressions. If this value is very close to unity, following the advise of Campbell and Perron (1991), it may be preferable to assimilate near $I(1)$ variables to $I(1)$ ones as their asymptotic behavior is more adequately described by that of a unit root process than by that of a stationary process. Therefore, even if for reasons of logical consistency our variables cannot contain a unit root, given the results of the tests and the values of ρ, it seems reasonable to approximate the near-unit root by a unit root and to pursue the analysis by searching for a co-integration relationship between unemployment rate and real interest rate.

5 The results

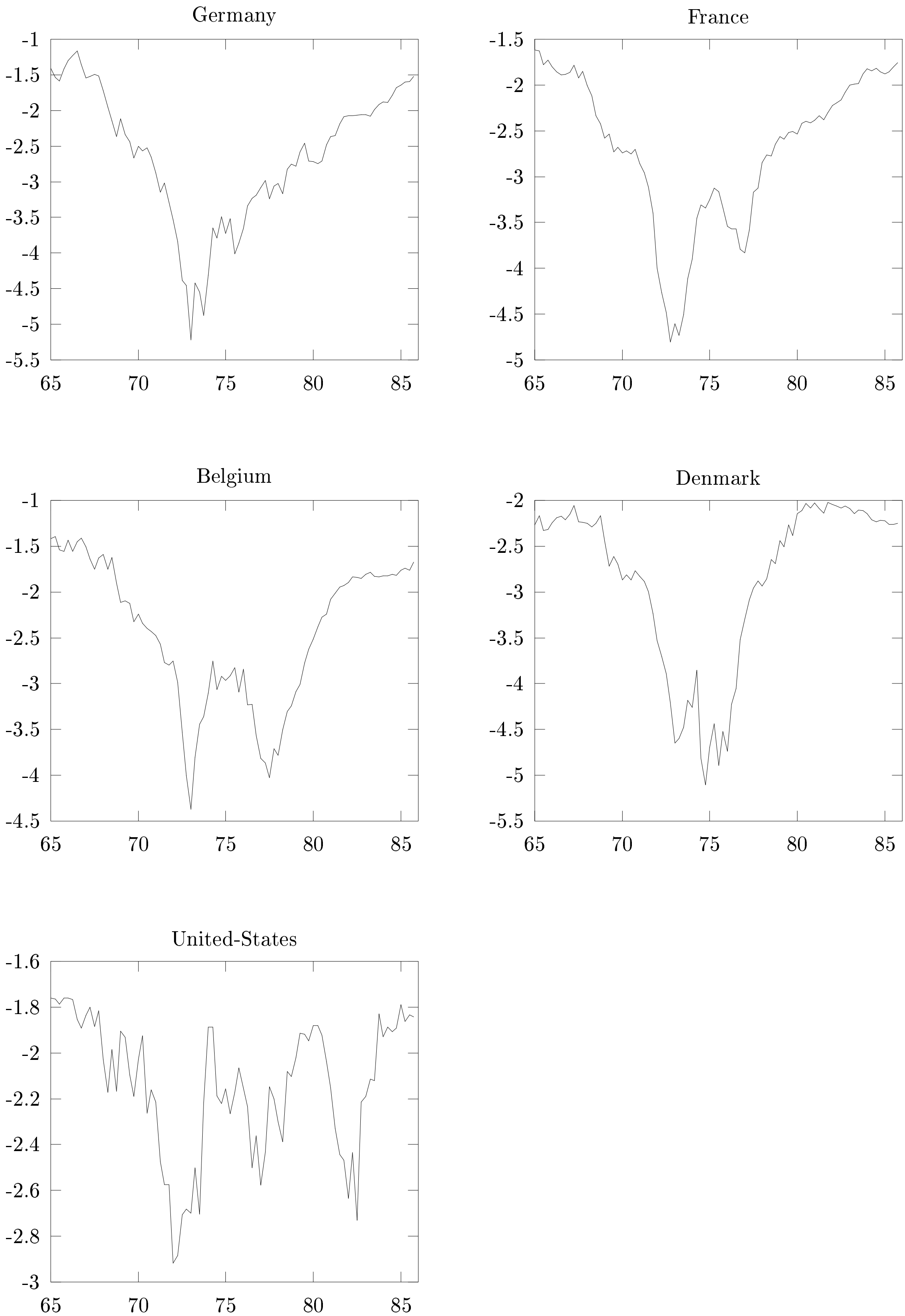
5.1 Classical co-integration tests

The functional form we have retained for estimating functions f and g of (2) is a log-linear one:

$$\ln ur_t = \beta_0 + \beta_1 \ln(1 + r_t) + ID(\tau)\,[\gamma_0 + \gamma_1 \ln(1 + r_t)] \tag{37}$$

[11]To cope with the Lindley's paradox, one has to adjust the level of a test to the sample size. See Berger (1985) for a numerical illustration.

Figure 5: Regime shifts with the classical procedure: ADF(4)



In a first step, we have estimated a long run static regression without break by ordinary-least-squares between $\ln ur_t$ and $\ln(1 + r_t)$ and we have applied the standard Engle-Granger co-integration test to see to what extent a break is necessary for co-integration. The "no break" column of Table 6 provides the AEG statistic for each country. It appears clearly that co-integration is rejected in all cases. This result is different from the one of Reichlin and Guillemineau (1989) who find co-integration for France and the US between unemployment and real interest rates over the years 1900-1985 using annual data.

Table 6
Classical co-integration tests

	no break	with break	
country	AEG(4)	break in	AEG(4)
BE	-1.41	73:1	-4.37
DK	-2.24	74:4	-5.11
FR	-1.71	72:4	-4.80
GY	-1.14	73:1	-5.22
US	-1.75	72:1	-2.91

In a second step, we apply the Gregory and Hansen (1992) classical method for testing co-integration in the presence of regime shift. (37) is estimated by OLS over a grid search for different values of τ ranging from 1965:1 to 1985:4. An AEG test with four lags is then computed for each possible break. The values of these AEG tests are plotted in Figure 5 for each country. The dates corresponding to the lower values of these tests are reported in the third column of Table 6, the corresponding AEG values being given in the fourth column. It appears that co-integration with regime shift is not rejected at 5% for Germany and Denmark. Co-integration with regime shift is not rejected at 10% for France. Co-integration with regime shift is rejected at 10% for Belgium and the United-States (the critical values are -4.95 at 5% and -4.68 at 10%). Concerning the date of the shift, the classical method does not give the same date for the three countries where co-integration is found. They range between 1972:4 and 1974:4. The break point for Germany corresponds to the date retained by Perron (1989) and occurs 3 quarters before the oil shock. The break point for France is even one quarter earlier. We cannot give too much importance to the exact date of the regime shift because it is not selected according to a likelihood criterion.

5.2 The Bayesian results

The Bayesian tests are presented in Table 7. The degree p of $A(L)$ in the complete VAR model (4) was determined with a Schwarz criterion. We give the posterior expectation of the break point $E(\tau|y)$, its mode and probability, the posterior expectation of ρ and the posterior probability of no co-integration.

Figure 6: Regime shifts with the Bayesian procedure: $D''(\tau|y)$

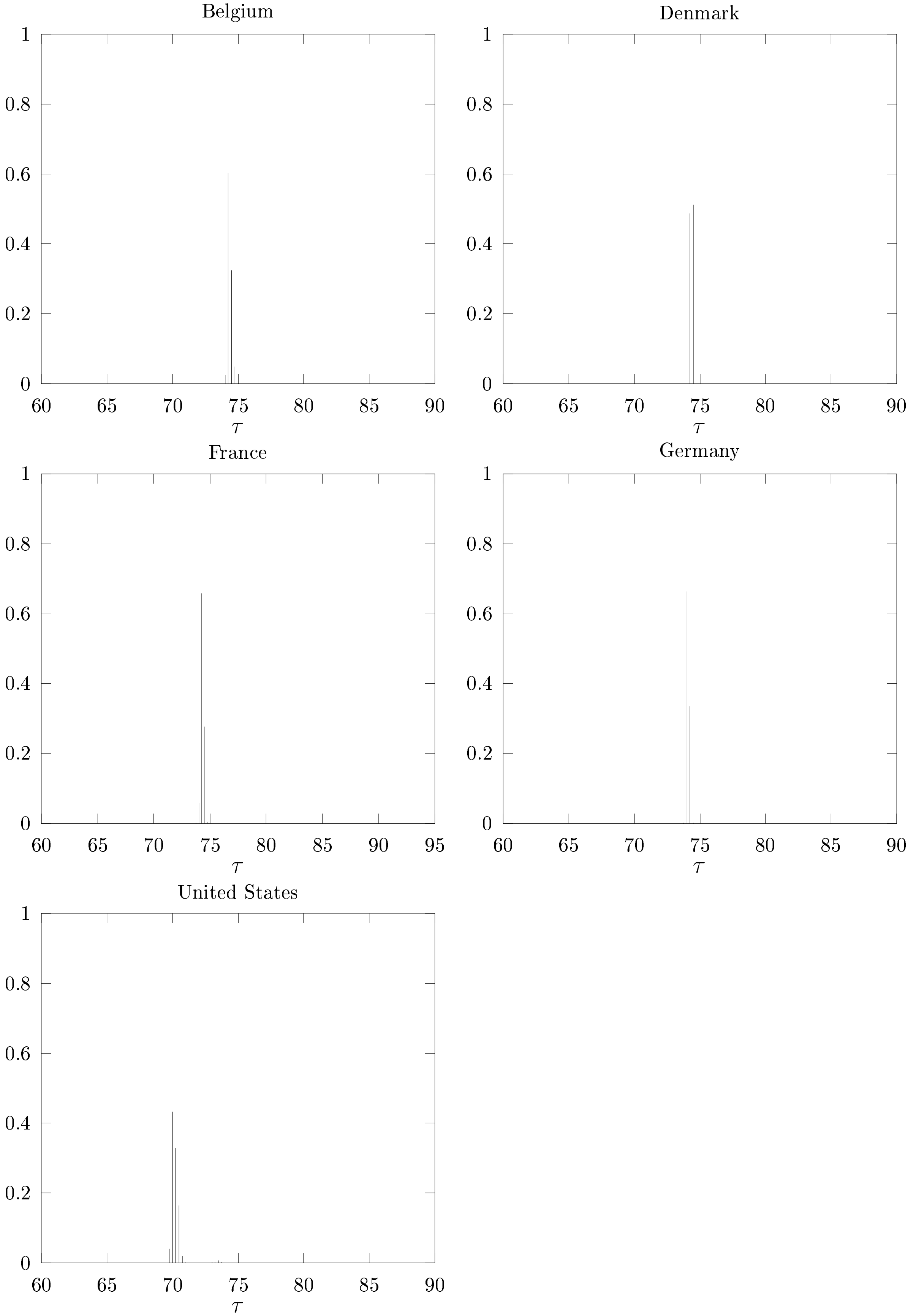


Table 7

Bayesian co-integration tests with regime shift

country	p	$E(\tau\|y)$	$\text{Mode}(\tau\|y)$	$\Pr(\tau = \text{Mode}\|y)$	$E(\rho\|y)$	$\Pr(\rho > 1\|y)$
BE	3	74:2	74:2	0.60	0.925	0.002
DK	3	74:3	74:3	0.51	0.642	0.000
FR	1	74:2	74:2	0.66	0.930	0.012
GY	4	74:1	74:1	0.66	0.820	0.000
US	3	70:2	70:1	0.43	0.957	0.111

The posterior density of τ is provided in Figure 6 for the five countries. For the United-States, we reject co-integration for any breaking point even at a 10% probability level. The posterior probability of the mode of τ is lower than 0.50 indicating that there is no clear regime shift. For the European countries, a clear pattern emerges. Co-integration is never rejected, contradicting the classical test for the Belgian case. For every country, a break point emerges with a posterior probability greater than 0.50. The date of the regime shift is 1974:1 for Germany. It is 1974:2/3 for the three other European countries. These dates are very similar, while the differences in the dates given by the Gregory-Hansen method are quite large.

We give in figure 7 the graph of the posterior density of ρ. In most cases (when there is co-integration), the Jeffreys' prior is dominated by the sample. For the US case the graph displays the important role of the prior for testing.

The posterior expectations of the coefficients of the long-run equation (33) and their standard errors are given in Table 8.

Table 8

The long term model

	β_0	β_1	τ	γ_0	γ_1	ω^2
BE	-3.92 (0.03)	0	74:2	1.11 (0.06)	12.04 (0.88)	0.0454
DK	-4.37 (0.04)	0	74:3	1.46 (0.10)	5.82 (1.48)	0.0743
FR	-4.14 (0.03)	0	74:2	1.00 (0.05)	13.78 (0.99)	0.0448
GY	-5.48 (0.18)	14.60 (4.56)	74:1	1.41 (0.25)	12.17 (6.10)	0.108

It is interesting to note that interest rates play no role in the first regime (before 1974) except for Germany. The shift in the constant term is very important and should be interpreted under the light of the discussion in section 1. Note also that the country where the elasticity of unemployment to interest rates is the highest is Germany.

5.3 The validity of the approach

We have estimated a conditional model of unemployment by implicitly assuming that the real interest rate is weakly exogenous. This procedure is correct if the real interest rate is actually exogenous. As a by-product of inference, we have estimated the marginal model (36) with the same number of lags as in the conditional model. As we have noted (36) is a simple linear model with a Student posterior density. We have computed the Student and Fisher "statistics" which can be used together with the appropriate table to judge if

$\Pr(|\alpha_2| > 0|y)$ and $\Pr(|\varphi_2| > 0|y)$ are greater than a pre-assigned probability level. These values are presented in table 9:

Table 9
Exogeneity and causality tests

Country	Weak exogeneity	Two tails t probability	Short term non-causality	Upper tail F probability
BE	t(93) = -0.48	0.632	F(3,93) = 1.45	0.233
DK	t(99) = 1.71	0.904	F(3,99) = 4.11	0.009
FR	t(117) = 0.07	0.944	F(1,117) = 3.48	0.065
GY	t(94) = 1.16	0.249	F(4,94) = 0.11	0.979

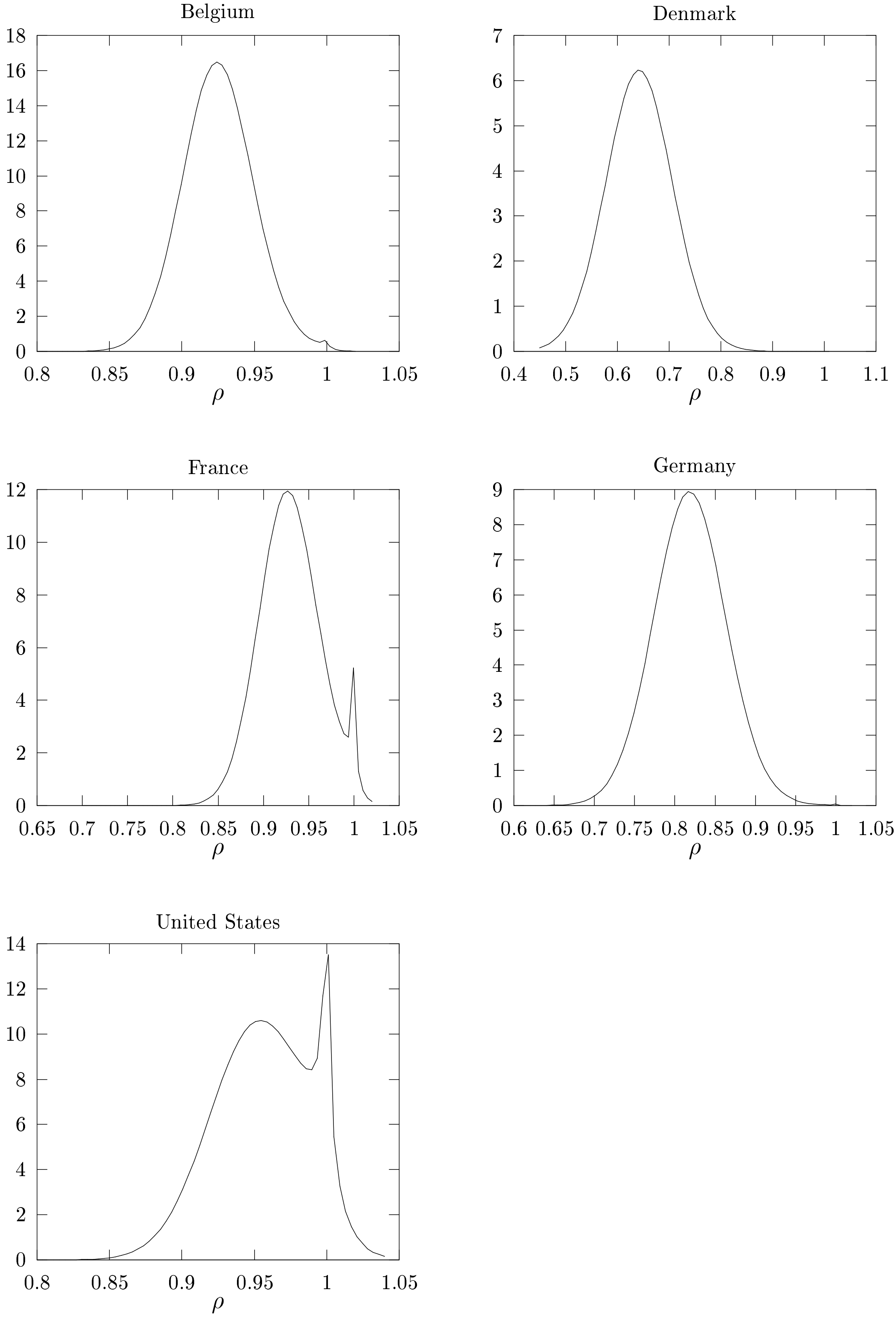
At the 5 % probability level, weak exogeneity is never rejected. So our inference procedure was correct. Moreover Urbain (1993) has interpreted weak exogeneity in bivariate cointegrated models with one vector of co-integration as being long-term causality. In this sense, we have here verified that the direction of long-term causality is from interest rates to unemployment. In the short term the story is not very different. Unemployment is of no use to predict interest rates, except possibly in Denmark.

Table 10
Decomposition of the change in unemployment

period	Δur	$\Delta ur - \Delta ur^*$	Δur^*	
		short-run factors	regime shift	interest rates
$72:4 \rightarrow 79:4$	5.6	-3.8	6.3	3.1
$79:4 \rightarrow 85:4$	6.3	6.3	0	0
$85:4 \rightarrow 89:4$	-1.5	-2	0	0.5
Total for Belgium	10.4	0.5	6.3	3.6
$72:4 \rightarrow 79:4$	4.1	-1.3	4.2	1.2
$79:4 \rightarrow 85:4$	3.0	2.7	0	1.3
$85:4 \rightarrow 89:4$	1.1	0.9	0	0.2
Total for Denmark	8.2	2.3	4.2	2.7
$72:4 \rightarrow 79:4$	3.8	0.7	4.7	-1.6
$79:4 \rightarrow 85:4$	3.7	-2.3	0	6
$85:4 \rightarrow 89:4$	0.6	1.8	0	-1.2
Total for France	8.1	0.2	4.7	3.2
$72:4 \rightarrow 79:4$	2.1	-0.6	2.6	0.1
$79:4 \rightarrow 85:4$	4.9	2	0	2.9
$85:4 \rightarrow 89:4$	-1.2	-0.4	0	-0.8
Total for Germany	5.8	1	2.6	2.2

Using these inference results it is possible to precisely quantify the effect of the rise of interest rates on the rise in equilibrium unemployment. The decomposition of the change in unemployment per subperiods detailed in Table 10 show that the rise in unemployment during the first subperiod 1973–1979 is essentially linked with the regime shift. The role of interest rate is dominant in the second subperiod 1979–1985 except in the Belgian case where it played already during 1973–1979. In Denmark, the observed rate of unemployment lies quite above its equilibrium level at the end of the eighties. The prominent role

Figure 7: Posterior density of ρ



of the interest rate in the eighties can be interpreted in the light of Fitoussi and Phelps (1988) argument or with the help of quantity rationing models. In this last case, the second wave of unemployment is due to an increased capital gap caused itself by high real interest rates.

It is always possible to claim that other variables are affecting the level of unemployment equilibrium, the most obvious ones being the world demand, as exemplified in Lubrano *et al* (1991), and the real price of energy. We have tried to introduce these variables in our equations to see whether the structural break is "explained" by these new variables and whether the interest rate loose its explanatory power. A break is still necessary to get co-integration and the date of the break is not much changed. The elasticity of unemployment to the interest rate is only slightly lowered.

6 Conclusion

Can the persistence of very high unemployment rates in most of European countries be attributed to the presence of high real interest rates? Fitoussi and Phelps (1988) would answer yes to this question by arguing that the sharp elevation of expected real interest rates induced firms in Europe to widen their mark-up rates, since it increased the opportunity cost of investing in greater market share through restraint in present prices at a sacrifice of present profit.

In this paper, this question was empirically analyzed for four European economies and the USA. Since, obviously, interest rates are not the only explanatory element of unemployment, we have tried to model a long-run relationship between unemployment and interest rate allowing for one regime shift. This regime shift is supposed to capture the change in non-modeled variables implied by the first oil shock. The question is whether the unique regime shift together with the movements in the interest rate is sufficient to model the evolution of equilibrium (long-run) unemployment.

For our four European countries (Belgium, Denmark, France and Germany) the relation with a shift around 1974 is shown to be cointegrating. For the United-States, co-integration is rejected and no clear break point does appear. In these relations, the first wave of unemployment occurring in the seventies is essentially explained by the regime shift of 1973/1974 while the second wave which took place at the beginning of the eighties is linked with the rise in real interest rates. This suggests that a reduction in real interest rates to their normal historical level could lower unemployment rates in the above mentioned European countries significantly.

In this empirical analysis we use a fully Bayesian approach, either for unit root tests or for co-integration tests when there is a shift in regime. We have thus proved the feasibility of the approach. Moreover, concerning the date of the regime shifts, the results obtained with the Bayesian approach are more in accordance with our intuition than the classical ones.

References

[1] Arnsperger C. and D. de la Croix (1993), "Bargaining and Equilibrium Unemploy-

ment: Narrowing the Gap Between New Keynesian and "Disequilibrium" Theories", *European Journal of Political Economy*, 9, 163–190.

[2] Berger J.O. (1985), *Statistical Decision Theory and Bayesian Analysis*, Springer Verlag, New-York.

[3] Berger J.O. and R.Y. Yang (1994), "Noninformative Priors and Bayesian Testing for the AR(1) Model", *Econometric Theory*, 10, 461-482.

[4] Bierens J. and L. Broersma (1993), "The Relation Between Unemployment and Interest Rate: Some International Evidence", *Econometric Reviews*, 12, 217–256.

[5] Boswijk H.P. (1991), "The LM Test for Weak Exogeneity in Error Correction Models", Report AE 13/91, Institute of Actuarial Science and Econometrics, University of Amsterdam.

[6] Broersma L. (1992) *The Relation Between Unemployment and Interest Rate: Empirical Evidence and Theoretical Justification*, Thesis/Tinbergen Institute, Amsterdam.

[7] Campbell, J. and P. Perron (1991), "Pitfalls and Opportunities : What Macroeconomists should know About Unit Roots", in O.J. Blanchard and S. Fisher (eds) *NBER Macroeconomic Annual*, Cambridge: MIT Press.

[8] de la Croix D. (1995), "Unemployment,and Capacity Utilization in a Tobin's q Model",*Recherches Economiques de Louvain* 61, 403-431.

[9] Drèze J. (1991), *Underemployment Equilibria*, Cambridge University Press.

[10] Engle R.F. and C.W.J. Granger (1987), "Co-Integration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55, 251-276.

[11] Engle R.F., D.F. Hendry and J.F. Richard (1983) "Exogeneity", *Econometrica*, 51, 277-304.

[12] Fitoussi J.-P. and E. Phelps (1986), "Causes of the 1980's Slump in Europe", *Brookings Papers on Economic Activity*, 2, 487–521.

[13] Fitoussi J.-P. and E. Phelps (1988), *The Slump in Europe*, Oxford, Basil Blackwell.

[14] Gregory A. and B. Hansen (1992), "Residual-Based Tests for Co-integration in Models with Regime Shifts", Queen's University Working Paper.

[15] Johansen S. (1992), "Co-integration in Partial Systems and Efficiency of Single Equation Analysis", *Journal of Econometrics*, 52, 389–402.

[16] Kleiberger F. and H.K. van Dijk (1994), "On the Shape of the Likelihood/Posterior in Co-integration Models", *Econometric Theory*, 10, 514-551.

[17] Kremers J., N. Ericsson and J. Dolado (1992), "The Power of Co-integration Tests", *Oxford Bulletin of Economics and Statistics*, 54, 325–348.

[18] Layard R., Nickell S. and R. Jackman (1991), *Unemployment*, Oxford University Press.

[19] Lubrano M. (1995), "Testing for Unit Roots in a Bayesian Framework," *Journal of Econometrics* 69, 81-109.

[20] Lubrano M., F. Mehta and H. Sneessens (1991) "Estimation of a Wage Equation for Belgium within a Disequilibrium Framework," Paper presented at the European Unemployment Problem Conference, Louvain la Neuve, 25–26 Mars 1991.

[21] Luukkonen R., P. Saikkonen and T. Terasvirta (1988), "Testing linearity against smooth transition autoregressive models", *Biometrika*, 75, 491–499.

[22] MacKinnon J. (1991), "Critical Values for Co-integration Tests" in *Long-Run Economic Relationships*, R.F. Engle and C.W.J. Granger eds., Oxford University Press.

[23] Mosconi R. and C. Giannini (1992) "Non-Causality in Cointegrated Systems: Representation, Estimation and Testing", *Oxford Bulletin of Economics and Statistics*, 54, 399–417.

[24] O'Brien, R.J., (1970), Serial Correlation in Econometric Models in: K. Hilton and D.F. Heathfield eds., *The Econometric Study of the United Kingdom*, (MacMillan, London) 375–437.

[25] Perron P. (1989), "The Great Crash, the Oil Price Shock and Unit Root Hypothesis", *Econometrica*, 57, 1361–1401.

[26] Pfann G. and F. Palm (1993), "Asymmetric Adjustment Costs in Non-linear Labor Demand Models For the Netherlands and UK Manufacturing Sectors", *Review of Economic Studies*, 60, 397–412.

[27] Phelps E. (1992), "A Review of Unemployment", *Journal of Economic Literature*, 30, 1476–1490.

[28] Phillips P.C.B. (1991), "To Criticize the Critics: an Objective Bayesian Analysis of Stochastic Trends," *Journal of Applied Econometrics*, 6, 333–364.

[29] Reichlin L. and C. Guillemineau (1989), "Chômage et Croissance en France et aux Etats-Unis", *Observations et Diagnostics Economiques*, 29, 161–185.

[30] Schotman P.C. and H.K. Van Dijk (1991), "Bayesian Analysis of the Unit Root in Real Exchange Rates". *Journal of Econometrics*, 49, 195–238.

[31] Sneessens H. (1987), "Investment and the Inflation-Unemployment Trade off in a Macroeconomic Rationing Model with Monopolistic Competition," *European Economic Review*, 31, 781–815.

[32] Thornber H. (1967), "Finite Sample Monte Carlo Studies: an Autoregressive Illustration," *Journal of the American Statistical Association*, 801–818.

[33] Toda H. Y. and P.C.B. Phillips (1993), "Vector Autoregressions and Causality", *Econometrica*, 61, 1367–1393.

[34] Uhlig H. (1994), "What Macroeconomists should Know About Unit Roots as Well: The Bayesian Perspective", *Econometric Theory*, 10, 645-671.

[35] Urbain J.P. (1992), "On Weak Exogeneity in Error Correction Models", *Oxford Bulletin of Economics and Statistics*, 54, 187–208.

[36] Urbain J.P. (1993), *Exogeneity in Error Correction Models*, Lectures Notes in Economics and Mathematical System, Springer Verlag.

[37] Wallis K.F. (1987), "Time Series Analysis of Bounded Economic Variables", *Journal of Time Series Analysis*, 8, 115–123.

[38] Zellner A. (1971) *An Introduction to Bayesian Inference in Econometrics*, New-York: John Wiley.

Appendix: The Jeffreys' Prior in Error Correction Models for Co-integration

Let us consider the following model:

$$y_t = x_t'\beta + u_t \tag{38}$$

$$u_t = \frac{\epsilon_t}{(1-\rho L)} \tag{39}$$

$$\epsilon_t \sim N(0,\sigma^2) \tag{40}$$

Conditionally on the first observation ($t = 0$), the log likelihood function of the model is:

$$L(y;\beta,\rho,\sigma^2) \propto -\frac{T}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}((y_t-\rho y_{t-1})-(x_t'-\rho x_{t-1}')\beta)^2 \tag{41}$$

The Jeffreys' prior for this model is obtained from the information matrix:

$$D_J'(\sigma^2,\rho,\beta) \propto \left| \begin{pmatrix} -\mathrm{E}\frac{\partial^2 L}{\partial\sigma^4} & -\mathrm{E}\frac{\partial^2 L}{\partial\rho\partial\sigma^2} & 0 \\ \mathrm{E}\frac{\partial^2 L}{\partial\sigma^2\partial\rho} & -\mathrm{E}\frac{\partial^2 L}{\partial\rho^2} & 0 \\ 0 & 0 & -\mathrm{E}\frac{\partial^2 L}{\partial\beta\partial\beta'} \end{pmatrix} \right|^{1/2} \tag{42}$$

The presence of the off diagonal elements is due to the dependence of the variance of the error term on ρ. We shall neglect these elements and retain only the diagonal elements of the information matrix in order to keep our results in symmetry with those of Phillips (1991) and Schotman and van Dijk (1991). Let us first compute the second order derivatives of the likelihood function:

$$\frac{\partial^2 L}{\partial\sigma^4} = \frac{T}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{t=1}^{T}(y_t-\rho y_{t-1}-(x_t'-\rho x_{t-1}')\beta)^2 \tag{43}$$

$$\frac{\partial^2 L}{\partial\rho^2} = -\frac{1}{\sigma^2}\sum_{t=1}^{T}(y_{t-1}-x_{t-1}'\beta)^2 \tag{44}$$

$$\frac{\partial^2 L}{\partial\beta\partial\beta'} = -\frac{1}{2\sigma^2}\sum_{t=1}^{T}(x_t-\rho x_{t-1})(x_t-\rho x_{t-1})' \tag{45}$$

We have then to compute minus the expectation of the matrix of the second order derivatives.

- There is no special problem to compute the first expectation as we have:

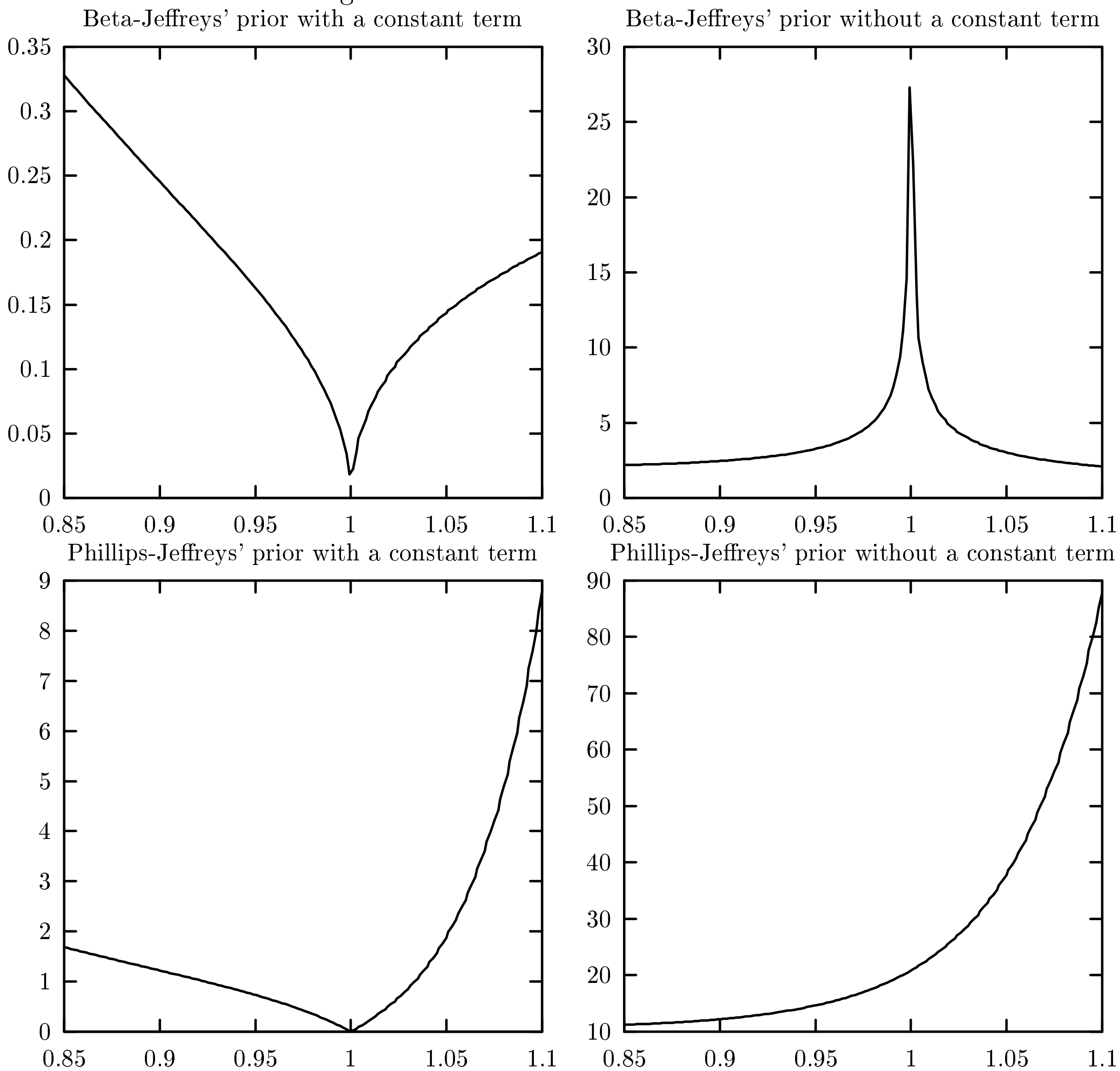
$$-\mathrm{E}\frac{\partial^2 L}{\partial\sigma^4} = \frac{T}{2\sigma^4} \tag{46}$$

- To compute the second block of the information matrix, we have to consider two cases: $\rho < 1$ and $\rho = 1$.

- For $\rho < 1$ we are in the stationary case:

$$-\mathrm{E}\frac{\partial^2 L}{\partial\rho^2} = \frac{1}{\sigma^2}\sum_{t=0}^{T-1}\mathrm{E}(y_t - x_t'\beta)^2$$

Figure 8: Priors for the Error Correction Model



As $(y_t - x_t'\beta) = \sum_{i=0}^{\infty} \rho^i \epsilon_{t-i}$ has a zero expectation,

$$-\,\mathrm{E}\frac{\partial^2 L}{\partial \rho^2} = \frac{T}{\sigma^2}\frac{\sigma^2}{1-\rho^2} = \frac{T}{1-\rho^2} \tag{47}$$

- For $\rho = 1$ we can no longer compute the expectation in this way and have to use a decomposition of $y_t - x_t'\beta = \epsilon_t/(1-\rho L)$ which is conditional on the initial condition $z_0 = (y_0 - x_0'\beta)$. We have:

$$y_t - x_t'\beta = \sum_{i=0}^{t-1} \rho^i \epsilon_{t-i} + \rho^t z_0$$

The expectation to compute is:

$$-\mathrm{E}\frac{\partial^2 L}{\partial \rho^2} = \frac{1}{\sigma^2}\sum_{t=0}^{T-1} \mathrm{E}(\sum_{i=0}^{t-1} \rho^i \epsilon_{t-i} + \rho^t z_0)^2$$

As:

$$\mathrm{E}(\sum_{i=0}^{t-1}\rho^i\epsilon_{t-i}+\rho^t z_0)^2=\sigma^2\sum_{i=0}^{t-1}\rho^{2i}+\rho^{2t}z_0^2=\sigma^2\frac{1-\rho^{2t}}{1-\rho^2}+\rho^{2t}z_0^2$$

we get:

$$\begin{aligned}-\mathrm{E}\frac{\partial^2 L}{\partial\rho^2} &= \frac{1}{\sigma^2}\sum_{t=0}^{T-1}\left[\sigma^2\frac{1-\rho^{2t}}{1-\rho^2}+\rho^{2t}z_0^2\right]\\ &= \sum_{t=0}^{T-1}\left[\frac{1-\rho^{2t}}{1-\rho^2}+\rho^{2t}\frac{z_0^2}{\sigma^2}\right]\\ &= \frac{1}{1-\rho^2}\left[T-\frac{1-\rho^{2T}}{1-\rho^2}+(1-\rho^{2T})\frac{z_0^2}{\sigma^2}\right] \end{aligned} \tag{48}$$

If we suppose that $z_0=0$, this expression reduces to:

$$\frac{1}{1-\rho^2}\left[T-\frac{1-\rho^{2T}}{1-\rho^2}\right] \tag{49}$$

Whether to use (47) or (49) to build a prior on ρ has been debated for long [see Uhlig (1994)]. Thornber (1967) uses (47). He supposes that ρ is defined over $]-1,1[$ and notices that the resulting Jeffreys' prior is the limit of a Beta prior defined by $D'(\rho)\propto(1+\rho)^{\nu_1}(1-\rho)^{\nu_2}$ where the degrees of freedom ν_1 and ν_2 are taken equal to as $-\frac{1}{2}$. Phillips (1991), Schotman and van Dijk (1991) use (49). For $\rho=1$, (49) is equal to $T(T-1)/2$. For $\rho>1$, (49) tends to infinity.

- The last block of the information matrix is the most interesting to compute as the presence or absence of a constant term in the matrix x_t changes dramatically the shape of the final Jeffreys' prior for the model. Let us define x_t as:

$$x_t'=[1,S_t] \qquad S_t=\sum_{i=0}^{t-1}v_{t-i} \qquad v_t\sim N(0,\tau^2) \tag{50}$$

in order to take into account the presence of I(1) regressors. Then:

$$-\mathrm{E}\frac{\partial^2 L}{\partial\beta\partial\beta'}=\frac{1}{2\sigma^2}\sum_{t=1}^{T}\mathrm{E}(x_t-\rho x_{t-1})(x_t-\rho x_{t-1})'$$

Let us compute

$$\begin{aligned}&\mathrm{E}(x_t-\rho x_{t-1})(x_t-2\rho x_{t-1})'\\ &= \mathrm{E}\begin{bmatrix}1 & S_t\\ S_t & S_t^2\end{bmatrix}-\rho\,\mathrm{E}\begin{bmatrix}1 & S_{t-1}\\ S_t & S_tS_{t-1}\end{bmatrix}+\rho^2\mathrm{E}\begin{bmatrix}1 & S_{t-1}\\ S_{t-1} & S_{t-1}^2\end{bmatrix}\\ &= \begin{bmatrix}1 & 0\\ 0 & t\tau^2\end{bmatrix}-\rho\begin{bmatrix}1 & 0\\ 0 & (t-1)\tau^2\end{bmatrix}+\rho^2\begin{bmatrix}1 & 0\\ 0 & (t-1)\tau^2\end{bmatrix}\end{aligned}$$

Consequently:

$$-\mathrm{E}\frac{\partial^2 L}{\partial\beta\partial\beta'}=\frac{T}{2\sigma^2}\begin{bmatrix}(1-\rho)^2 & 0\\ 0 & \tau^2(\frac{1}{2}(T-1)(1-\rho)^2+1)\end{bmatrix} \tag{51}$$

τ^2 can be supposed equal to one without a loss of generality.

The Jeffreys' prior is found by taking the square root of the product of the diagonal elements of the information matrix. We shall consider two main cases, corresponding to (47) and (49), and in each case the presence or absence of a constant term in x_t. k is the number of I(1) variables in x_t.

- For the stationary case we have:

$$D'(\sigma^2,\rho,\beta) \propto \frac{1}{\sigma^3}\left[\frac{(1-\rho)^2}{1-\rho^2}[\frac{1}{2}(T-1)(1-\rho)^2+1]^k\right]^{\frac{1}{2}} \tag{52}$$

  This prior is equal to zero for $\rho = 1$. If we put the constant term out of x_t, we get:

$$D'(\sigma^2,\rho,\beta) \propto \frac{1}{\sigma^3}\left[\frac{1}{1-\rho^2}[\frac{1}{2}(T-1)(1-\rho)^2+1]^k\right]^{\frac{1}{2}} \tag{53}$$

  This prior has a singularity at $\rho = 1$ which is of order $\rho^{1/2}$ and hence does not cause integrability problems. We can extend (52) or (53) to cover the explosive case by considering the change of variable[12] $\tilde{\rho} = 1/\rho$ of Jacobian $1/\tilde{\rho}^2$. This gives for (52):

$$D'(\sigma^2,\rho,\beta) \propto \frac{1}{\sigma^3}\frac{1}{|\rho|^{k+2}}\left[\frac{(\rho-1)^2}{\rho^2-1}[\rho^2+\frac{1}{2}(T-1)(\rho-1)^2]^k\right]^{\frac{1}{2}} \tag{54}$$

  which is singular at $\rho = 1$ and tends to 0 for $\rho \to \infty$, but at a slow pace. We can do the same extension for (53) to get:

$$D'(\sigma^2,\rho,\beta) \propto \frac{1}{\sigma^3}\frac{1}{|\rho|^{k+1}}\left[\frac{1}{\rho^2-1}[\rho^2+\frac{1}{2}(T-1)(\rho-1)^2]^k\right]^{\frac{1}{2}} \tag{55}$$

  This prior also tends to 0 for $\rho \to \infty$, but at a quicker pace.

- For the unit root case, the previous result is a bit altered as we have:

$$D'(\sigma^2,\rho,\beta) \propto \frac{1}{\sigma^3}\left[\frac{1-\rho}{1+\rho}[T-\frac{1-\rho^{2T}}{1-\rho^2}][\frac{1}{2}(T-1)(1-\rho)^2+1]^k\right]^{\frac{1}{2}} \tag{56}$$

  It is still equal to zero for $\rho = 1$. If we put the constant term out of x_t, we get:

$$D'(\sigma^2,\rho,\beta) \propto \frac{1}{\sigma^3}\left[\frac{1}{1-\rho^2}[T-\frac{1-\rho^{2T}}{1-\rho^2}][\frac{1}{2}(T-1)(1-\rho)^2+1]^k\right]^{\frac{1}{2}} \tag{57}$$

  This prior tends to ∞ for $\rho \to \infty$.

Let us call (52)-(54) and (53)-(55) Beta-Jeffreys' prior for reasons that are evident from above. We shall call (56) and (57) Phillips-Jeffreys' priors as Peter Phillips was the first to propose to compute the Jeffreys' prior using (49) for the unit root model. We can reach the following conclusions with the help of the four figures given above. The Phillips-Jeffreys' prior (56) has a very special shape at $\rho = 1$ when there is a constant term in x_t, putting a zero weight on the null hypothesis of no co-integration. This is very undesirable. When the constant term is put outside x_t, the Phillips- Jeffreys' prior recovers its usual shape. But (57) has very undesirable properties has it favors too much explosive values of ρ and consequently the null hypothesis of no co-integration. The Beta-Jeffreys' prior (52)-(54) is not very appealing. Consequently, the sole prior which reckons that there is something unusual that happens at $\rho = 1$ and that give a very small weight to strictly explosive values of ρ is (53)-(55).

[12]this suggestion was originally made by Berger in a seminar at Yale (1992) for the unit root model.