Re-examining the Structural and the Persistence Approach to Unemployment

Tino Berger[*] Gerdie Everaert[†]

February 2006

Abstract

This paper uses an unobserved component model to examine the relative importance of the structural and the persistence approach to unemployment. We derive the NAIRU from a standard imperfect competition model. The price- and wage-setting schedules include a measure for unemployment persistence. Short-run dynamics are introduced through a demand equation which is linked to unemployment via Okun's Law. This multivariate model is then estimated for the US and the euro data using Bayesian techniques and the Kalman filter. The results show that although cyclical shocks are very persistent, most of the increase in European unemployment is driven by supply factors. The degree of persistence is somewhat lower in the US but demand shocks seem to be more important in explaining variation of unemployment.

[*]SHERPPA, Ghent University, Tino.Berger@UGent.be, http://www.sherppa.be
[†]SHERPPA, Ghent University, Gerdie.Everaert@UGent.be, http://www.sherppa.be

1 Introduction

The high and persistent unemployment in Europe is one of the biggest challenges for policymakers and labour economists in recent times. Starting at historically low levels in the post World War II period, the rate of unemployment has increased from the mid 1970s to an average level of 9% in both Europe and the US in the mid 1980s. In the aftermath, the labour markets on the two sides of the Atlantic went into completely different directions. While the high level still persists in Europe, the US managed to recover to a low level of unemployment. The persistence of high unemployment in Europe is still puzzling many observers. Although there is no absolute consensus about its potential causes, two lines of explanations have been developed.

First, according to the structural approach unemployment adjusts quickly to cyclical shocks. As such, it is most of the time at, or close to, its natural level. The idea of a constant natural rate of unemployment has been pioneered by Friedman (1968) and Phelps (1968) who claimed that unemployment is at its natural level when neither inflationary nor deflationary pressure emanates from the labour market. This is called the non-accelerating-inflation-rate of unemployment (NAIRU). The existence of a constant NAIRU has been questioned after the oil price shocks of the 1970s as unemployment remained high in Europe even after inflation has stabilised. More recently, the structuralist school considers the natural rate as a function of labour market institutions (Nickell *et al.*, 2005), real macroeconomic variables such as real interest rates (Blanchard, 2003) and productivity growth (Pissarides, 1990) or interactions between macroeconomic shocks and institutions (Blanchard and Wolfers, 2000).

Second, the persistence approach (see e.g. Blanchard and Summers, 1986, Karanassou and Snower, 1998, Layard *et al.*, 1991) focuses on the dynamic adjustment of unemployment towards the natural rate after a temporary business cycle shock. Cyclical unemployment may translates into medium-run unemployment or even becomes permanent due to labour market rigidities. As such, unemployment can be far away from its equilibrium level for a long period of time. These persistence effects can arise from insider-outsider effects in wage formation (see e.g. Blanchard and Summers, 1986) and/or depreciation of skills and search ineffectiveness of the unemployed (see e.g. Phelps, 1972). A situation where temporary shocks have a permanent impact on unemployment is known as hysteresis.[1]

Understanding to what extent each approach contributes to the explanation of unemployment is important from a policy point of view. If structural protagonists are right, unemployment can only be reduced by changing labour market characteristics. An expansionary policy can only reduce unemployment in the short-run, at the cost of rising inflation. However, in the medium-run there

[1] In the literature the terms hysteresis and persistence are sometimes synonymously used. Here we use term hysteresis only when a transitory shock has a *permanent* impact on unemployment whereas persistence means that unemployment is mean reverting, even when this takes a long time. For an introduction to the concept of hysteresis in the labour market and its potential sources see Røed (1997).

is no effect since the natural rates underlying labour market characteristics remain unchanged. If the persistence protagonists are right, expansionary policy can reduce unemployment in the medium-run as the effect will last much longer. Even when inflation is brought back to its original level unemployment is still lower due to the persistence of shocks in the labour market.[2]

Empirical studies of the structural approach can be divided into two groups. The first group attempts to relate unemployment to various factors that are expected to represent labour market characteristics. Nickell *et al.* (2005), for instance, consider the natural rate to be a function of generous benefits, trade union power, taxes and wage inflexibility. This approach not only aims at disentangling short-run from long-run unemployment but also at measuring the particular impact of every institution under consideration. The main shortcoming of this approach, as highlighted in Daveri (2001), is the omitted variable problem. Economic theory relates equilibrium unemployment to a large variety of factors, some of them being difficult to measure or even unobservable, e.g. the reservation wage which is a function of, among others, the value of leisure. Regressions with missing variables suffer from an omitted variable bias or can even be spurious if the variables under consideration are non-stationary.

The second group overcomes this problem by treating the NAIRU as an unobserved component and using filter technique to estimate it. In this field the dominant approach nowadays is based on the expectation-augmented Phillips-curve. The NAIRU is modelled as a non-stationary process reflecting permanent changes in its underlying determinants. A number of studies (e.g Fabiani and Mestre, 2001, Orlandi and Pichelmann, 2000) use a multivariate approach and include an Okun's-Law relation that links the output gap to cyclical unemployment. However, common to the fast majority of time-varying NAIRU estimates is that they neglect elements of persistence. As a consequence the NAIRU also captures, if present, persistence effects of demand shocks, which makes it hard to conclude whether labour market institutions or persistence effects drive the NAIRU.

Empirical studies of the persistence approach analyse the time series properties of unemployment using (univariate) autoregressive (AR) time series models. Persistence is measured as the sum of the estimated AR coefficients. Most of the studies using this methodology are not able to reject the hypothesis that this sum equals one, i.e. a unit root in unemployment. This suggests the presence of hysteresis in unemployment. Elmeskov and Macfarlan (1993), for instance, show that the hypothesis of a unit root cannot be rejected in any of the 23 OECD countries they consider. However, this method of detecting persistence or hysteresis becomes invalid if it does not account for structural breaks in the mean of unemployment, i.e. which is the case if the NAIRU is time-varying. Even a single break in the mean can induce a unit root in an otherwise stationary process

[2]This further depends on the symmetry of persistence. However, even when the persistence after an adverse shock is higher than after an positive shock, expansionary policy can still lower unemployment in the medium-run.

(Perron, 1990). As shown in Bianchi and Zoega (1998) and in Papell, Murray and Ghiblawi (2000) the degree of persistence in European unemployment is indeed substantially lower if one accounts for a shift in the mean.

The aim of this paper is to combine the two approaches. First, in line with the structural approach, we model the NAIRU as an unobserved non-stationary process reflecting permanent changes in its underlying determinants, as derived from a standard imperfect competition model. Second, we allow for slow adjustment towards this time-varying NAIRU by including elements of persistence in the price- and wage-setting schedules. Furthermore the effects of demand on unemployment are embedded into the model through a demand equation which is linked to unemployment via Okun's Law. This specification enables us to estimates a (long-run) NAIRU which is not affected by persistence effects. The model is estimated using Bayesian technique and applied to the euro area and the US and covers the period 1970Q1-2003Q4.

The paper is structured as follows. Section 2 presents the theoretical model. The estimation methodology is presented in section 3. Section 4 presents the results. Section 5 concludes.

2 Theoretical background

The model outlined in this section is a standard *imperfect competition model* (see e.g. Layard *et al.*, 1991; Bean, 1994). Firms operate in markets with imperfect competition and set prices at the beginning of the period on the basis of expectations of future demand and costs. Output, employment and wages are set during the period. Output is determined by firms simply supplying whatever is demanded at the predetermined prices. Employment is then set to produce output. Wages are set in a non-competitive way, due to e.g. wage bargaining between firms and unions or efficiency wages. The main point is that we consider wages to be influenced by firm specific factors, such as productivity and insider behaviour, and outsider factors, such as wages paid elsewhere and the general state of the labour market. We further assume exogenous determined demand, capital stock and technology.

2.1 Static model

Price-setting

A profit maximising firm sets prices as a mark-up on expected wages. The aggregate price equation can then be written as

$$p_t - w_t^e = \beta_0 + \beta_1 u_t - \beta_2 (p_t - p_t^e) - q_t + z_p, \tag{1}$$

where p_t is the price level at time t and w_t^e is the expected wage level. All variables are expressed in logarithms. The price mark-up depends on unemployment u_t, on price surprises $(p_t - p_t^e)$, on trend labour productivity q_t, and on other price push variables, z_t, such as oil prices or import prices. The constant term in the price setting schedule β_0, is a function of the market structure, including the degree of product market competition and the price elasticity of demand. The effect of demand on prices is proxied by the level of unemployment. Falling unemployment is associated with rising demand which generates an upward pressure on prices. The price surprise term reflects nominal inertia which may result from price-adjustment costs and staggered price setting.

Wage-setting

Aggregate wages are set as a mark-up on expected prices as

$$w_t - p_t^e = \alpha_0 - \alpha_1 u_t - \alpha_2 \left(w_t - w_t^e\right) + q_t + z_w, \tag{2}$$

with the mark-up decreasing in unemployment, wage surprises $(w_t - w_t^e)$, trend labour productivity q_t, and a wage pressure variable z_w. As in the price-setting schedule the wage surprise term captures potential nominal inertia effects in wage setting which may arise as the result of staggered wage contracts. The wage pressure variable is a function of labour market institutions such as union power or the generosity of unemployment benefits.

Equilibrium Rate of unemployment

The long-run equilibrium rate of unemployment u_t^*, is defined as the situation where expectations are fulfilled, i.e. $w_t = w_t^e$ and $p_t = p_t^e$. Combining equation (1) and (2) gives

$$u_t^* = \frac{\beta_0 + \alpha_0 + z_w + z_p}{\beta_1 + \alpha_1} \tag{3}$$

The equilibrium rate of unemployment is a function of the constant terms in the wage and price equations, α_0 and β_0, wage and price push variables, z_w and z_p, and real wage and price flexibility, α_1 and β_1.

The unemployment-inflation trade-off (Phillips curve) and the NAIRU

If actual wages w_t and actual prices p_t are not at their expected values w_t^e and p_t^e, we have

$$u_t = u_t^* - \frac{\left(1 + \alpha_2\right)\left(w_t - w_t^e\right) + \left(1 + \beta_2\right)\left(p_t - p_t^e\right)}{\beta_1 + \alpha_1} \tag{4}$$

Now assume that wage and price surprises are equal, i.e. $(w_t - w_t^e) = (p_t - p_t^e)$ and that inflation $\triangle p_t$ follows a random walk, i.e.[3]

$$\triangle p_t = \triangle p_{t-1} + \upsilon_t,$$

where υ_t is a white noise process. Then the rational forecast of inflation is

$$E_{t-1}\left(\triangle p_t\right) = \triangle p_{t-1}$$

[3]The assumption of a unit root in inflation is consistent with empirical studies and has become standard in the literature (Staiger *et al.*, 1997).

Therefore

$$p_t^e = p_{t-1} + \Delta p_{t-1}$$

and

$$p_t - p_t^e = \Delta p_t - \Delta p_{t-1} = \Delta^2 p_t \tag{5}$$

Combining equations (4) and (5) yields

$$\Delta^2 p_t = -\theta_1 \left(u_t - u_t^*\right), \tag{6}$$

where $\theta_1 = (\beta_1 + \alpha_1) / (2 + \beta_2 + \alpha_2)$ is a measure of real wage and price flexibility. This equation represents the unemployment-inflation trade-off known as the Phillips-curve. When unemployment is lower than u_t^*, inflation is increasing, and vice versa. Thus u_t^* can be thought of as the NAIRU.

2.2 Medium-run unemployment dynamics

In order to analyse the potential role of persistence or hysteresis effects we consider dynamic versions of the price and wage setting equations (1) and (2). The particular choice of functional forms is similar to the approach used by the OECD (OECD, 1999). The dynamic price-setting schedule is

$$p_t - w_t^e = \beta_0 + \beta_1 u_t + \beta_{11} \Delta u_t - \beta_2 \left(p_t - p_t^e\right) - q_t + z_p. \tag{7}$$

Unemployment now also enters the price equation in its first difference to capture labour adjustment costs. If these costs delay employment adjustment, and hence unemployment adjustment, marginal costs are higher in the short-run than in the long-run (where employment is at its optimal level). Thus the effect of prices arising from changes in demand are greater in the short-run, $(\beta_1 + \beta_{11})$, than in the long-run, (β_1). The dynamic wage-setting schedule is

$$w_t - p_t^e = \alpha_0 - \alpha_1 u_t - \alpha_{11} \Delta u_t - \alpha_2 \left(w_t - w_t^e\right) + q_t + z_w. \tag{8}$$

Unemployment now also enters the wage equation in its first difference. This should capture hysteresis effects, caused by insider-outsider behaviour and/or duration composition effects. In the former, a transitory shock reduces the number of insiders and thus puts upward pressure on wages. This results in a positive effect of lagged unemployment which together with the standard negative effect of contemporaneous unemployment gives the change term of unemployment. In the latter, the duration of unemployment matters for aggregate wages as long-term unemployed are less strong competitors for jobs and therefore put less pressure on wages than short-run unemployed. The change term Δu_t now captures the idea that wage pressure is lower when unemployment has recently risen $(\Delta u_t > 0)$ as people that became recently unemployed are stronger competitors for jobs. The Phillips-curve is now

$$\Delta^2 p_t = -\theta_1 \left(u_t - u_t^*\right) - \theta_{11} \Delta u_t, \tag{9}$$

where $\theta_{11} = (\beta_{11} + \gamma_{11}) / (2 + \beta_2 + \alpha_2)$. The impact of persistence and hysteresis effects in wage and price setting becomes more clear if we rewrite equation (9) as

$$u_t = \frac{\theta_1}{\theta_1 + \theta_{11}} u_t^* + \frac{\theta_{11}}{\theta_1 + \theta_{11}} u_{t-1} - \frac{1}{\theta_1 + \theta_{11}} \Delta^2 p_t.$$

Unemployment is affected by its long-run equilibrium level, its own past, and by short-run cyclical unemployment captured by $\Delta^2 p_t$. This specification shows that even when inflation is stable unemployment can be far away from its natural rate due to persistence effects. The higher θ_{11} relative to $\theta_1$the more persistent unemployment is. The unemployment rate which stabilises inflation u_t^n is given by

$$u_t^n = \kappa u_t^* + (1 - \kappa) u_{t-1},$$

where $\kappa = \theta_1 / (\theta_1 + \theta_{11})$. Layard *et al.* (1991) refer to this as the short-run NAIRU. It is a weighted average of u_t^* and u_{t-1}. We can distinguish three cases of interest: (i) $\kappa = 0$ means that cyclical shocks have a permanent impact on unemployment. The long-run NAIRU u_t^* is not an attractor anymore since unemployment is only affected by its own history with no tendency to revert to an equilibrium. This is known as hysteresis; (ii) if $\kappa = 1$ unemployment is not affected by persistence effects, i.e. the short-run NAIRU u_t^n equals the long-run NAIRU u_t^*; (iii) if $0 < \kappa < 1$ unemployment converges to u_t^* after a business cycle shock. However, the speed of adjustment depends on κ. In terms of policy, persistence means that, once unemployment has risen, it cannot be brought back at once without a permanent increase in inflation. But it can be reduced gradually without inflation rising.

2.3 Short-run unemployment fluctuations

In the long-run unemployment is determined by long-run supply factors and equals u_t^*. In the short-run unemployment is determined by the interaction of aggregate supply, given by the Phillips curve in equation (9), and aggregate demand y_t^d given by

$$y_t^d = \frac{1}{\lambda_1}(m_t - p_t) + \frac{1}{\lambda_2} x_t, \tag{10}$$

where m_t is the nominal money stock and x_t captures all exogenous real factors driving demand, e.g. fiscal policy. This equation is simply the reduced form of an IS-LM system. The link between aggregate demand and unemployment is given by Okun's Law[4]

$$y_t^d - \overline{y}_t = -\omega(u_t - u_t^*), \tag{11}$$

where $\overline{y}_t$ is potential output. Taking (10) and (11) together it follows

$$u_t - u_t^* = -\frac{1}{\lambda_1 \omega}(m_t - p_t) - \frac{1}{\lambda_2 \omega} x_t + \frac{1}{\omega} \overline{y}_t \tag{12}$$

[4]This relation can be derived from a production function as shown in Appendix 1.

Combining this equation with the short-run aggregate supply equation, the Phillips curve, we have

$$u_t = \rho u_t^* + (1-\rho)u_{t-1} + \frac{\rho\omega\lambda_1}{\theta_1}\Delta u_t^* - \frac{\rho}{\theta_1}\left(\Delta\sigma_t^d - \lambda_1\Delta\overline{y}_t - \Delta p_{t-1}\right) \tag{13}$$

where $\rho = \frac{\theta_1}{\theta_1+\theta_{11}+\omega\lambda_1}$ and $\sigma_t^d = m_t + \frac{\lambda_1}{\lambda_2}x_t$.

Equation (13) explains unemployment by supply (u_t^*) and demand (σ_t^d relative to $\overline{y}_t$ and p_{t-1}) factors. If $\rho = 0$ (the hysteresis case), both supply and demand shocks have a permanent impact on u_t. If $\rho > 0$ the long-run level of u_t is entirely determined by supply-side factors, i.e. u_t is attracted by u_t^*. In the short-run, demand shocks push unemployment away from its long-run value. An expansionary fiscal or monetary policy for instance will reduce unemployment. The fall in unemployment is only temporary as rising inflation will drive unemployment back to its equilibrium level. The speed at which unemployment adjusts depends on the size of persistence effects. If there are no persistence effects ($\rho = 1$), the speed of adjustment depends on real wage and price flexibility, as captured by θ_1, only. In this case, unemployment is at its long-run equilibrium level u_t^* when inflation has stabilised, i.e. the short-run NAIRU u_t^n equals the long-run NAIRU u_t^*. Note that inflation is stable when $\Delta p_{t-1} = \Delta\sigma_t^d - \lambda_1\Delta\overline{y}_t$. So $\Delta\sigma_t^d - \lambda_1\Delta\overline{y}_t$ is the long-run equilibrium level of inflation. If there are persistence effects ($0 < \rho < 1$), unemployment converges to the long-run NAIRU at a speed given by ρ. As such, u_t deviates from u_t^* even when inflation has stabilised.

3 Estimation methodology

3.1 State space representation of the model

In this section we cast the model outlined in the previous section into a state space representation.[5] In a state space model, the development over time of the system under study is determined by an unobserved series of vectors $\alpha_1, \ldots, \alpha_n$, which are associated with a series of observed vectors $y_1, \ldots, y_n$. A general linear Gaussian state space model can be written in the following form

$$y_t = Z\alpha_t + Ax_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, H), \tag{14}$$

$$\alpha_{t+1} = T\alpha_t + R\eta_t, \qquad \eta_t \sim N(0, Q), \qquad t = 1, \ldots, n, \tag{15}$$

where y_t is a $p \times 1$ vector of observed endogenous variables, modelled in the observation equation (14), x_t is a $k \times 1$ vector of observed exogenous or predetermined variables and α_t is a $m \times 1$ vector of unobserved states, modelled in the state equation (15). The disturbances ε_t and η_t are assumed to be independent sequences of independent normal vectors. The matrices Z, A, T, R, H, and Q are parameter matrices.

The model outlined in the previous section includes the observed endogenous variables y_t, u_t and p_t and the unobserved states $\overline{y}_t$, u_t^* and σ_t^d. Writing this model in the general state space

[5] See e.g. Durbin and Koopman (2001) for an extensive overview of state space models.

representation in equations (14)-(15) requires two steps.[6] First, to derive the reduced form for the observed endogenous variables as a function of the unobserved states and the lagged (i.e. predetermined) observed endogenous variables by solving equations (9), (10) and (11) for $y_t, u_t,$ and p_t as

$$y_t^d = \overline{y}_t + \frac{\omega}{\theta_1}\rho\left(\sigma_t^d - \lambda_1\overline{y}_t - 2p_{t-1} + p_{t-2}\right) + \frac{\omega\theta_{11}}{\theta_1}\rho\left(u_t^* - u_{t-1}\right) + \varepsilon_{1t} \tag{16}$$

$$u_t = u_t^* - \frac{\rho}{\theta_1}\left(\sigma_t^d - \lambda_1\overline{y}_t - 2p_{t-1} + p_{t-2}\right) - \frac{\theta_{11}\rho}{\theta_1}\left(u_t^* - u_{t-1}\right) + \varepsilon_{2t} \tag{17}$$

$$p_t = \sigma_t^d - \lambda_1\overline{y}_t - \frac{\lambda_1\omega}{\theta_1}\rho\left(\sigma_t^d - \lambda_1\overline{y}_t - 2p_{t-1} + p_{t-2}\right) - \frac{\lambda_1\omega\theta_{11}}{\theta_1}\rho(u_t^* - u_{t-1}) + \varepsilon_{3t} \tag{18}$$

Second, the dynamics of the unobserved states are assumed to be given by

$$\overline{y}_{t+1} = \overline{y}_t + \psi_t + \eta_{1t} - \omega\eta_{3t} \tag{19}$$

$$\psi_{t+1} = \psi_t + \eta_{2t} \tag{20}$$

$$u_{t+1}^* = (1+\delta)u_t^* - \delta u_{t-1}^* + \eta_{3t} \tag{21}$$

$$\sigma_{t+1}^d = \phi_t + \eta_{4t} \tag{22}$$

$$\phi_{t+1} = \phi_t + \tau_t + \eta_{5t} \tag{23}$$

$$\tau_{t+1} = \tau_t + \eta_{6t}, \tag{24}$$

where the error terms ε_{it} with $i = 1, ..., 3$ and η_{it} with $i = 1, ..., 6$ are mutually independent zero mean white noise processes. The ε's are interpreted as measurement errors whereas the η's represent structural shocks.

Following Harvey (1985) and Stock and Watson (1998), among others, equations (19)-(20) model potential output as a random walk with drift, with the drift term ψ_t varying over time according to a random walk process. The time-variation in ψ_t allows for the possibility of permanent changes in the trend growth of real output, e.g. the productivity slowdown of the early 1970s.[7] Potential output is further affected by u_t^* through the term $-\omega\eta_{3t}$. This negative relationship between the equilibrium levels of output and unemployment is intuitively clear and can be seen in the derivation of Okun's Law in Appendix 1.

Equation (21) specifies the long-run NAIRU u_t^* as a non-stationary process, i.e. shifts in its underlying determinants are assumed to be permanent. As a pure random walk process would result in a non-smooth series that is hard to reconcile with the expected smooth evolution of the structural characteristics driving the long-run NAIRU, the AR(2) specification in equation (21)

[6] The exact specification of the vectors y_t, x_t and α_t and the matrices Z, A, T, R, H and Q which cast the model in equations (16)-(24) in the general state space representation in equations (14)-(15) is provided in Appendix 2.

[7] Note that the random walk in equation (20) implies that $\overline{y}_t$, and therefore also y_t, is an I(2) process. This seems inconsistent with the empirical evidence from Dickey-Fuller (DF) unit root tests that real output is I(1). Stock and Watson (1998) argue, though, that when the variance of η_{2t} is small relative to the variance of η_{1t}, $\Delta\overline{y}_t$ has a moving average (MA) root close to unity. Schwert (1989) and Pantula (1991) show that the size of the standard DF unit root test is severely upward biased in the presence of a large MA root. In this case, the standard DF unit root test is inappropriate to pick up a possible I(2) component in real output.

allows for a smooth evolution of u_t^* over time, i.e. the closer δ is to one the smoother u_t^* is. If $\delta = 0$, u_t^* is a pure random walk process. Note that in order to induce this smoothness, the NAIRU is nowadays often modelled as an I(2) series, i.e. δ is set to one (see e.g. Orlandi and Pichelmann 2000), in particular when euro area NAIRU's are estimated. We do not restrict δ to be equal to one in equation (21) as in this case u_t^* exhibits a (time-varying) drift, which would be hard to justify from an economic perspective.

Equations (22)-(24) model the demand factor σ_t^d as the sum of three components: (i) an erratic component driven by η_{4t}, (ii) a level component driven by η_{5t}, and (iii) a drift component driven by η_{6t}. The erratic component is included to capture temporary shifts in demand, like e.g. a temporary increase in government spending, while the level component captures permanent shifts in demand. The drift component captures permanent changes in monetary policy, i.e. a permanent change in the growth rate of the money stock m_t which induces a permanent change in the level of inflation Δp_t.

3.2 Identification of the unobserved states

Assuming that Z, A, T, R, H, and Q are known, the purpose of state space analysis is to infer the relevant properties of the α_t's from the observations $y_1, \ldots, y_n$ and $x_1, \ldots, x_n$.

Intuitively, potential output $\overline{y}_t$ is identified as the equilibrium level of output as shown in equation (16) of the reduced form. Deviation of output from its potential level, i.e. the output gap, result from demand and/or persistence effects captured by the last term of this equation. Similarly, the long-run NAIRU u_t^* is the long-run level of unemployment after filtering out demand and persistence effects from the unemployment rate. The permanent component ϕ_t in σ_t^d is identified through inflation since it is the moving average of the inflation rate.

Formally, the unobserved states are identified through the subsequent use of two recursions, i.e. the Kalman filter and the Kalman smoother. The objective of filtering is to obtain the distribution of α_t, for $t = 1, \ldots, n$, conditional on Y_t and X_t, where $Y_t = \{y_1, \ldots, y_t\}$ and $X_t = \{x_1, \ldots, x_t\}$. In a linear Gaussian state space model, the distribution of α_t is entirely determined by the filtered state vector $a_t = E(\alpha_t \mid Y_t, X_t)$ and the filtered state variance matrix $P_t = Var(\alpha_t \mid Y_t, X_t)$. The (contemporaneous) Kalman filter algorithm (see e.g. Hamilton, 1994, or Durbin and Koopman, 2001) estimates a_t and P_t by updating, at time t, a_{t-1} and P_{t-1} using the new information contained in y_t and x_t. The Kalman filter recursion can be initialised by the assumption that $\alpha_1 \sim N(a_1, P_1)$. In practice, a_1 and P_1 are generally not known though. Therefore, we assume that the distribution of the initial state vector α_1 is

$$\alpha_1 = V\Gamma + R_0\eta_0, \qquad \eta_0 \sim N(0, Q_0), \qquad \Gamma \sim N(0, \kappa I_r), \tag{25}$$

where the $m \times r$ matrix V and the $m \times (m - r)$ matrix R_0 are selection matrices composed of columns of the identity matrix I_m. They are defined so that, when taken together, their columns constitute all the columns of I_m and $V'R_0 = 0$. The matrix Q_0 is assumed to be positive definite and known. The $r \times 1$ vector Γ is a vector of unknown random quantities, referred to as the diffuse vector as we let $\kappa \to \infty$. This leads to

$$\alpha_1 \sim N(0, P_1), \qquad P_1 = \kappa P_\infty + P_*, \tag{26}$$

where $P_\infty = VV'$ and $P_* = R_0 Q_0 R_0'$. The Kalman filter is modified to account for this diffuse initialisation implied by letting $\kappa \to \infty$ by using the exact initial Kalman filter introduced by Ansley and Kohn (1985) and further developed by Koopman (1997) and Koopman and Durbin (2003).

Subsequently, the Kalman smoother algorithm is used to estimate the distribution of α_t, for $t = 1, \ldots, n$, conditional on Y_n and X_n, where $Y_n = \{y_1, \ldots, y_n\}$ and $X_n = \{x_1, \ldots, x_n\}$. Thus, the smoothed state vector $\widehat{a}_t = E(\alpha_t \mid Y_n, X_n)$ and the smoothed state variance matrix $\widehat{P}_t = Var(\alpha_t \mid Y_n, X_n)$ are estimated using all the observations for $t = 1, \ldots, n$. In order to account for the diffuse initialisation of α_1, we use the exact initial state smoothing algorithm suggested by Koopman and Durbin (2003).

3.3 Parameter estimation

Bayesian analysis

The filtering and smoothing algorithms both require that Z, A, T, R, H, and Q are known. In practice these matrices generally depend on elements of an unknown parameter vector ψ. One possible approach is to derive, from the exact Kalman filter, the diffuse loglikelihood function for the model under study (see e.g. de Jong 1991, Koopman and Durbin 2000, Durbin and Koopman 2001) and replace the unknown parameter vector ψ by its maximum likelihood (ML) estimate. This is not the approach pursued in this paper. Given the fairly large number of unknown parameters (and unobserved states) the numerical optimisation of the sample loglikelihood function is quite tedious. Therefore, we analyse the state space model from a Bayesian point of view, i.e. we use prior information to down-weight the likelihood function in regions of the parameter space that are inconsistent with out-of-sample information and/or in which the structural model is not interpretable (Schorfheide, 2006). As such, the Bayesian approach is based on the same likelihood function as classical ML estimation but adds information that may help to discriminate between alternative parameterisations of the model. More formally, we treat ψ as a random parameter vector with a known prior density $p(\psi)$ and estimate the posterior densities $p(\psi \mid y, x)$ for the parameter vector ψ and $p(\widehat{\alpha}_t \mid y, x)$ for the smoothed state vector $\widehat{\alpha}_t$, where y and x denote the stacked vectors $(y_1', \ldots, y_n')'$ and $(x_1', \ldots, x_n')'$ respectively, by combining information contained in

$p(\psi)$ and the sample data. This boils down to calculating the posterior mean $\overline{g}$

$$\overline{g} = E\left[g(\psi) \mid y, x\right] = \int g(\psi)\, p(\psi \mid y, x)\, d\psi \tag{27}$$

where g is a function which expresses the moments of the posterior densities $p(\psi \mid y, x)$ and $p(\widehat{\alpha}_t \mid y, x)$ in terms of the parameter vector ψ.

Importance sampling

In principle, the integral in equation (27) can be evaluated numerically by drawing a sample of n random draws of ψ, denoted $\psi^{(i)}$ with $i = 1, \ldots, n$, from $p(\psi \mid y, x)$ and then estimating $\overline{g}$ by the sample mean of $g(\psi)$. As $p(\psi \mid y, x)$ is not a density with known analytical properties, such a direct sampling method is not feasible, though. Therefore, we switch to importance sampling. The idea is to use an importance density $g(\psi \mid y, x)$ as a proxy for $p(\psi \mid y, x)$, where $g(\psi \mid y, x)$ should be chosen as a distribution that can be simulated directly and is as close to $p(\psi \mid y, x)$ as possible. By Bayes' theorem and after some manipulations, equation (27) can be rewritten as

$$\overline{g} = \frac{\int g(\psi)\, z^g(\psi, y, x)\, g(\psi \mid y, x)\, d\psi}{\int z^g(\psi, y, x)\, g(\psi \mid y, x)\, d\psi} \tag{28}$$

with

$$z^g(\psi, y, x) = \frac{p(\psi)\, p(y \mid \psi)}{g(\psi \mid y, x)} \tag{29}$$

Using a sample of n random draws $\psi^{(i)}$ from $g(\psi \mid y, x)$, an estimate $\overline{g}_n$ of $\overline{g}$ can then be obtained as

$$\bar{g}_n = \frac{\sum\limits_{i=1}^{n} g\left(\psi^{(i)}\right) z^g\left(\psi^{(i)}, y, x\right)}{\sum\limits_{i=1}^{n} z^g\left(\psi^{(i)}, y, x\right)} = \sum_{i=1}^{n} w_i g\left(\psi^{(i)}\right) \tag{30}$$

with w_i

$$w_i = \frac{z^g\left(\psi^{(i)}, y, x\right)}{\sum\limits_{i=1}^{n} z^g\left(\psi^{(i)}, y, x\right)} \tag{31}$$

The weighting function w_i reflects the importance of the sampled value $\psi^{(i)}$ relative to other sampled values.

Geweke (1989) shows that if $g(\psi \mid y, x)$ is proportional to $p(\psi \mid y, x)$, and under a number of weak regularity conditions, $\overline{g}_n$ will be a consistent estimate of $\overline{g}$ for $n \to \infty$.

Computational aspects of importance sampling

As a first step importance density $g(\psi \mid y, x)$, we take a large sample normal approximation to $p(\psi \mid y, x)$, i.e.

$$g(\psi \mid y, x) = N\left(\widehat{\psi}, \widehat{\Omega}\right), \tag{32}$$

where $\widehat{\psi}$ is the mode of $p(\psi \mid y, x)$ obtained from maximising

$$\log p(\psi \mid y, x) = \log p(y \mid \psi) + \log p(\psi) - \log p(y) \tag{33}$$

with respect to $\widehat{\psi}$ and where $\widehat{\Omega}$ denotes the covariance matrix of $\widehat{\psi}$. Note that $p(y \mid \psi)$ is given by the likelihood function derived from the exact Kalman filter and we do not need to calculate $p(y)$ as it does not depend on ψ. In drawing from $g(\psi \mid y, x)$, efficiency was improved by the use of antithetic variables, i.e. for each $\psi^{(i)}$ we take another value $\widetilde{\psi}^{(i)} = 2\widehat{\psi} - \psi^{(i)}$, which is equiprobable with $\psi^{(i)}$. This results in a simulation sample that is balanced for location (Durbin and Koopman 2001).

As any numerical integration method delivers only an approximation to the integrals in equation (28), we monitor the quality of the approximation by estimating the probabilistic error bound for the importance sampling estimator $\overline{g}_n$ (Bauwens, Lubrano and Richard 1999, chap. 3, eq. 3.34). This error bound represents a 95% confidence interval for the percentage deviation of $\overline{g}_n$ from $\overline{g}$. It should not exceed 10%. In practice this can be achieved by increasing n, except when the coefficient of variation of the weights w_i is unstable as n increases. An unstable coefficient of variation of w_i signals poor quality of the importance density.

Note that the normal approximation in equation (32) selects $g(\psi \mid y, x)$ in order to match the location and covariance structure of $p(\psi \mid y, x)$ as good as possible. One problem is that the normality assumption might imply that $g(\psi \mid y, x)$ does not match the tail behaviour of $p(\psi \mid y, x)$. If $p(\psi \mid y, x)$ has thicker tails than $g(\psi \mid y, x)$, a draw $\psi^{(i)}$ from the tails of $g(\psi \mid y, x)$ can imply an explosion of $z^g\left(\psi^{(i)}, y, x\right)$. This is due to a very small value for $g(\psi \mid y, x)$ being associated with a relatively large value for $p(\psi)\, p(y \mid \psi)$, as the latter is proportional to $p(\psi \mid y, x)$. Importance sampling is inaccurate in this case as this would lead to a weight w_i close to one, i.e. $\overline{g}_n$ is determined by a single draw $\psi^{(i)}$. This is signalled by instability of the weights and a probabilistic error bound that does not decrease in n. In order to help prevent explosion of the weights, we change the construction of the importance density in two respects (Bauwens *et al.* 1999, chap. 3). First, we inflate the approximate covariance matrix $\widehat{\Omega}$ a little. This reduces the probability that $p(\psi \mid y, x)$ has thicker tails than $g(\psi \mid y, x)$. Second, we use a sequential updating algorithm for the importance density. This algorithm starts from the importance density defined by (32), with inflation of $\widehat{\Omega}$, estimates posterior moments for $p(\psi \mid y, x)$ and then defines a new importance density from these estimated moments. This improves the estimates for $\widehat{\psi}$ and $\widehat{\Omega}$. We continue updating the importance density until the weights stabilise. The number of importance samples n was chosen to make sure that the probabilistic error bound for the importance sampling estimator $\overline{g}_n$ does not exceed 10%.

4 Estimation Results[8]

4.1 Data

We use quarterly data for the US and the euro area from 1970Q1 to 2003Q4. US data are taken from the OECD Economic Outlook. Euro area data, which are aggregate series for 12 countries[9], are taken from the area-wide model (AWM) of Fagan *et al.* (2005). The unemployment rate, u_t, is the quarterly unemployment rate. For prices, P_t, we use the seasonally adjusted quarterly GDP deflator. Output, y_t^d, is the log of seasonally adjusted quarterly GDP in constant prices.

4.2 Prior distribution of the parameters

Prior information on the unknown parameter vector ψ is included in the analysis through the prior density $p(\psi)$. Detailed information on $p(\psi)$ can be found in Table 1. As stated above, the main motivation for setting these priors is to down-weight the likelihood function in regions of the parameter space that are inconsistent with out-of-sample information and/or in which the structural model is not interpretable. Previous estimates as well as economic theory give us an idea about the approximate values of ω. This parameter is known as Okun's Law coefficient and measures the percentage rise of the output gap when the unemployment gap falls by one percent. Okun stated that this relation is linear and ω is roughly three (Okun, 1970). Here we set ω to 2.5 since more recent empirical studies found Okun's Law coefficient somewhat lower than 3 (see e.g Orlandi and Pichelmann, 2000). λ_1^{-1} measures the impact of real balances on aggregate demand. The prior mean and variance for λ_1 are chosen so that its 95% confidence interval ranges from 0.82 to 0.98, implying a roughly unit impact of real balances on aggregate demand. Setting priors on θ_1 and θ_{11} is more difficult as the fast majority of Phillips-curve estimates does not include a persistence measure such as θ_{11} and thus cannot be used here. Moreover, we do not want to make these priors too informative since measuring the degree of persistence is of particular interest in our analysis. Therefore we leave a considerable amount of uncertainty around these two parameters by choosing a high prior variance. The same argument is true for δ which is only included to allow for smoothness in u_t^*. As we do not want to be too informative in this direction, we have chosen a flat prior for δ.

Priors on state variances are found by filtering out the unobserved states using the parameter's prior mean instead of their posterior values. The state variances are then set so that the resulting output gap matches with the commonly accepted timing of the business cycle with respect to shape and frequency of the output gap. Again, we leave a considerable amount of uncertainty around these prior variances. The output gap's obtained from this exercise are shown in Figure 3

[8]The GAUSS code to obtain the results presented in this section is available from the authors on request.

[9]Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, and Spain.

and 4 in the Appendix.

Table 1: Prior distribution

	Euro Area			US		
Parameter	5 p.c.	Mean	95 p.c.	5 p.c.	Mean	95 p.c.
θ_1	0.13	0.65	1.17	0.13	0.65	0.17
θ_{11}	-0.31	0.35	1.01	-0.31	0.35	1.01
ω	2.13	2.50	2.87	2.13	2.50	2.87
λ_1	0.82	0.90	0.98	0.82	0.90	0.98
δ	-0.69	0.95	2.59	-0.69	0.95	2.59
$\sigma^2_{\varepsilon_1}$	0.03	0.04	0.06	0.02	0.03	0.04
$\sigma^2_{\varepsilon_2}$	0.03	0.04	0.06	0.02	0.03	0.04
$\sigma^2_{\varepsilon_3}$	0.03	0.04	0.06	0.02	0.03	0.04
$\sigma^2_{\eta_1}$	$4.92{*}10^{-4}$	$8.00{*}10^{-4}$	$1.17{*}10^{-3}$	$3.10{*}10^{-3}$	$5.00{*}10^{-3}$	$7.28{*}10^{-3}$
$\sigma^2_{\eta_2}$	$6.23{*}10^{-6}$	$1.00{*}10^{-5}$	$1.47{*}10^{-5}$	$3.07{*}10^{-5}$	$5.00{*}10^{-5}$	$7.29{*}10^{-5}$
$\sigma^2_{\eta_3}$	$3.10{*}10^{-6}$	$5.00{*}10^{-6}$	$7.34{*}10^{-6}$	$3.12{*}10^{-5}$	$5.00{*}10^{-5}$	$7.30{*}10^{-5}$
$\sigma^2_{\eta_4}$	0.09	0.15	0.21	0.03	0.05	0.07
$\sigma^2_{\eta_5}$	0.05	0.08	0.11	$4.60{*}10^{-3}$	0.08	0.01
$\sigma^2_{\eta_6}$	0.02	0.03	0.04	$1.55{*}10^{-3}$	0.03	$3.67{*}10^{-3}$

The prior distribution is assumed to be Gaussian for all elements in ψ, except for the variance parameters which are assumed to be gamma distributed.

4.3 Posterior distribution

In this section we present estimates of the posterior mean $\overline{\psi} = E[\psi \mid y, x]$ of the parameter vector ψ and the posterior mean $\overline{\alpha}_t = E[\widehat{\alpha}_t \mid y, x]$ of the smoothed state vector $\widehat{\alpha}_t$. An estimate $\widetilde{\psi}$ of $\overline{\psi}$ is obtained by setting $g(\psi^{(i)}) = \psi^{(i)}$ in equation (30) and taking $\widetilde{\psi} = \overline{g}_n$. An estimate $\widetilde{\alpha}_t$ of $\overline{\alpha}_t$ is obtained by setting $g(\psi^{(i)}) = \widehat{\alpha}_t^{(i)}$ in equation (30) and taking $\widetilde{\alpha}_t = \overline{g}_n$, where $\widehat{\alpha}_t^{(i)}$ is the smoothed state vector obtained from the Kalman smoother using the parameter vector $\psi^{(i)}$.

We also present the 5th and 95th percentiles of the posterior densities $p(\psi \mid y, x)$ and $p(\widehat{\alpha}_t, x)$. Let $F(\psi_j \mid y, x) = \Pr\left(\psi_j^{(i)} \leq \psi_j\right)$ with ψ_j denoting the j-th element in ψ. An estimate $\widetilde{F}(\psi_j \mid y, x)$ of $F(\psi_j \mid y, x)$ is obtained by setting $g(\psi^{(i)}) = I_j\left(\psi_j^{(i)}\right)$ in equation (30) and taking $\widetilde{F}(\psi_j \mid y, x) = \overline{g}_n$, where $I_j\left(\psi_j^{(i)}\right)$ is an indicator function which equals one if $\psi_j^{(i)} \leq \psi_j$ and zero otherwise. An estimate $\widetilde{\psi}_j^{5\%}$ of the 5th percentile of the posterior density $p(\psi \mid y, x)$ is chosen such that $\widetilde{F}\left(\psi_j^{5\%} \mid y, x\right) = 0.05$. An estimate $\widetilde{\alpha}_{j,t}^{5\%}$ of the 5th percentile of the jth element of the posterior density $p(\widehat{\alpha}_t \mid y, x)$ is obtained by setting $g(\psi^{(i)}) = \widehat{\alpha}_{j,t}^{(i)} - 1.645\sqrt{\widehat{P}_{j,t}^{(i)}}$ in equation (30) and taking $\widetilde{\alpha}_{j,t}^{5\%} = \overline{g}_n$, where $\widehat{\alpha}_{j,t}^{(i)}$ denotes the j-th element in $\widehat{\alpha}_t^{(i)}$ and $\widehat{P}_{j,t}^{(i)}$ is the (j,j)th element of the smoothed state variance matrix $\widehat{P}_t^{(i)}$ obtained using the parameter vector $\psi^{(i)}$. The 95th percentiles are constructed in a similar way.

Posterior distribution of the parameters

Table 2 presents the posterior mean and the 5% and 95% percentile of the posterior distribution for the euro area and the US estimates. The hypothesis of hysteresis in unemployment must be rejected for the US and for the euro area. However the degree of unemployment persistence is fairly high in Europe. As explained earlier the quotient $\kappa = \frac{\theta_1}{\theta_1+\theta_{11}}$ determines the speed at which the equilibrium rate u_t^*, adjusts to the short-run NAIRU u_t^n. In the euro area we find that $\kappa = 0.08$, implying that this adjustment is rather slow and thus unemployment is very persistent. The results for the US show that $\kappa = 0.36$ and therefore US unemployment is adjusting somewhat faster.[10]

Table 2: Posterior distribution

	Euro Area			US		
Parameter	5 p.c.	Mean	95 p.c.	5 p.c.	Mean	95 p.c.
θ_1	0.09	0.16	0.23	0.07	0.08	0.1
θ_{11}	1.42	1.77	2.15	0.04	0.14	0.1
ω	2.15	2.30	2.45	2.16	2.29	2.45
λ_1	0.88	0.96	1.04	0.82	0.89	0.94
δ	0.97	0.98	0.98	0.92	0.97	0.99
$\sigma^2_{\varepsilon_1}$	0.12	0.15	0.17	0.07	0.08	0.1
$\sigma^2_{\varepsilon_2}$	0.02	0.03	0.04	0.05	0.05	0.06
$\sigma^2_{\varepsilon_3}$	0.08	0.09	0.11	0.06	0.07	0.8
$\sigma^2_{\eta_1}$	$5.24*10^{-4}$	$8.62*10^{-4}$	$1.32*10^{-3}$	$5.40*10^{-3}$	$7.21*10^{-3}$	$9.89*10^{-3}$
$\sigma^2_{\eta_2}$	$9.95*10^{-6}$	$1.58*10^{-5}$	$2.36*10^{-5}$	$3.07*10^{-5}$	$5.11*10^{-5}$	$7.77*10^{-5}$
$\sigma^2_{\eta_3}$	$7.76*10^{-6}$	$1.13*10^{-5}$	$1.54*10^{-5}$	$5.55*10^{-5}$	$7.46*10^{-5}$	$1.01*10^{-4}$
$\sigma^2_{\eta_4}$	0.04	0.06	0.10	0.07	0.10	0.12
$\sigma^2_{\eta_5}$	0.04	0.07	0.10	$5.80*10^{-3}$	$7.78*10^{-3}$	0.01
$\sigma^2_{\eta_6}$	0.02	0.03	0.04	0.05	0.06	0.07

Note that the approximate covariance matrix $\hat{\Omega}$ is inflated with a factor 1.5. The coefficient of variation of the weights stabilised after 5 updates of the importance function for both the euro area and the United States. With $n = 10000$, the probabilistic error bound for the importance sampling estimator $\overline{g}_n$ is well below 10 % for all coefficients.

Posterior distribution of the states

The posterior distribution of u_t^* are shown in Figures 1 and 2. The estimated equilibrium rates of unemployment are both very smooth and do not follow the data closely. The reason for this might be that we have controlled for persistence.[11] This allows us to interpret the graph as long-run equilibrium determined entirely by supply side factors. The graph for the euro area shows a clear upward trend form the beginning of the 1970s up to the middle of the 1990s while from that time on its downward sloping. The equilibrium rate of the US seems to be rather stable throughout the sample period with a decrease of 2% in the 1980s. The demand variable σ^d was decomposed into a permanent and a transitory component. Figure 5 and 6 show the permanent component of σ^d (adjusted for real trend growth) as the long-run trend of inflation. Demand effects explain most of unemployment variation in the US whereas the upward drift in euro area unemployment is supply side driven.

[10]Worth mentioning is that the estimates of θ_1 and θ_{11} are for both data sets robust in the sense that changing priors does not affect the outcome of these two parameters much.

[11]The resulting inflation stabilising rate of unemployment, u_t^n follows the actual rate of unemployment rather close.

Figure 1: Equilibrium unemployment in the euro area

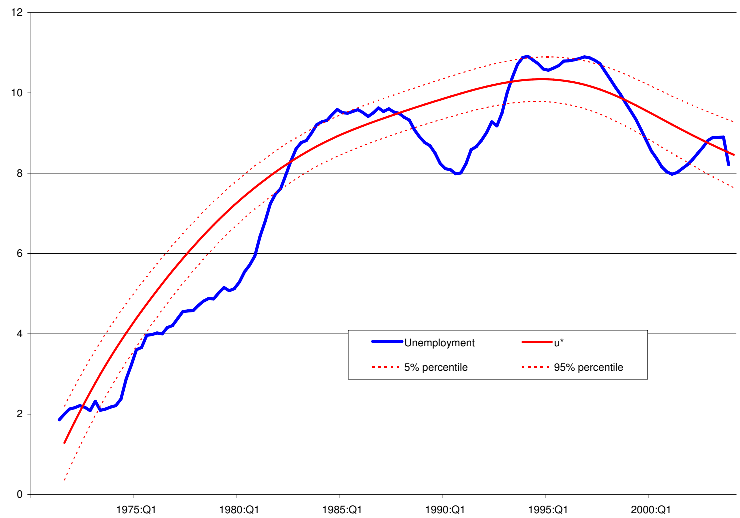
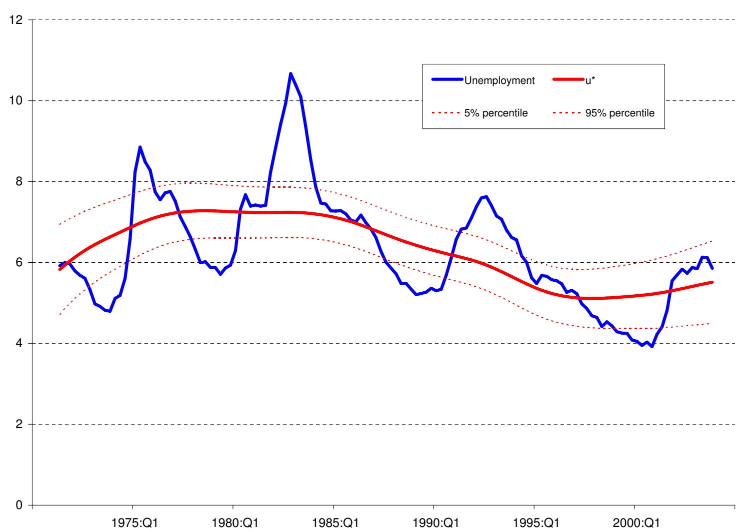


Figure 2: Equilibrium unemployment in the US



5 Conclusion

This paper examines the relative importance of the structural and the persistence approach to unemployment. It estimates a time-varying equilibrium rate of unemployment as a measure of the structural component of unemployment. Particular attention is paid to unemployment persistence, i.e. the slow response of unemployment after business cycle shocks. In the literature unemployment persistence is often found to behave like a random walk, implying that any shock has a permanent impact on unemployment. It is measured as the sum of its AR components, neglecting the possibility of structural breaks. Studies that allow for a moving natural rate of unemployment usually reject the hysteresis hypothesis. They, however, estimate a rather short-run NAIRU instead of a long-run equilibrium rate since elements of persistence are usually not considered. This study differs from existing measures of time-varying equilibrium rates in two points: (i) we measure the persistence of shocks and (ii) we derive the equilibrium rate from a theoretical model which explains unemployment dynamics by demand and supply factors as well as by persistence mechanism. Persistence effects are introduced into wage and price setting. The effect of demand on unemployment is not imposed but embedded in the model. The multivariate model is then estimated using the Kalman filter and Bayesian econometrics. Our results show that the hysteresis hypothesis must be rejected for both data sets. We found a fairly high degree of persistence in Europe while unemployment is much less persistent in the US. Nevertheless, the increase of euro area unemployment until the late 80s is driven by supply side factors. In contrast most of unemployment variation in the US since the beginning of the 70ies is driven by demand shocks. Given these results we conclude that both, the structural and the persistence approach, are needed to explain variation in unemployment.

References

[1] Ansley, C. and R. Kohn, 1985, "Estimation, filtering and smoothing in state space models with incompletely specified initial conditions", *Annals of Statistics*, 13, 1286-1316.

[2] Bauwens, L., Lubrano, M. and J.F. Richard, 1999, *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, Oxford.

[3] Bean, C. R. 1994, "European Unemployment: A Survey", *Journal of Economic Literature*, 32, 573- 619.

[4] Blanchard, O., 2003 "Monetary policy and unemployment", remarks at the conference Monetary policy and the labor market", in honor of James Tobin, held at the New School University, NY.

[5] Blanchard, O. and L. Summers, 1986, "Hysteresis and the European Unemployment Problem", NBER Macroeconomics Annual 1986, 15-77.

[6] Blanchard, O. and J. Wolfers, 2000 "The Role of Shocks and institutions in the rise of European Unemployment: the Aggregate Evidence", *Economic Journal*, 110.

[7] Bianchi, M. and G. Zoega, 1998 "Unemployment persistence: Does the size of the shock matter", *Journal of Applied Econometrics*, 13, 283-304.

[8] de Jong, P., 1991, "The diffuse Kalmn filter,", *Annals of Statistics*, 19, 1073-1083

[9] Daveri, F., 2001, "Labor Taxes and Unemployment, A Survey of the Aggregate Evidence", Conference paper. 2nd Annual CERP Conference on Pension Policy Harmonization in an Integrating Europe, Moncalieri, Turin, 223 June 2001.

[10] Durbin, J. and S. Koopman, 2001, *Time Series Analysis by State Space Methods*, Oxford Statistical Science Series, Oxford University Press, Oxford.

[11] Elmeskov, J. and M. MacFarlan, 1993 "Unemployment Persistence", *OECD Economic Studies*, No. 21, p. 59 88

[12] Fabiani, S. and R. Mestre, 2001, "A System Approach for Measuring the euro Area NAIRU," European Central Bank Working Paper, No. 65.

[13] Fagan, G., Henry J. and R. Mestre, 2005, "An area-wide model (AWM) for the euro area", *Economic Modelling*, 22, 39-59.

[14] Friedmann, M., "The role of monetary policy", *The American Economic Review*, 58, 1-17.

[15] Geweke, J., 1989, "Bayesian inference in econometric models using Monte Carlo integration", *Econometrica*, 57, 1317-1339.

[16] Hamilton, J., 1994, *Time Series Analysis*, Princeton University Press, Princeton.

[17] Harvey, A.C., 1985, "Trends and Cycles in Macroeconomic Time Series", *Journal of Business and Economic Statistics*, 3, 216-227.

[18] Harvey, A., 1989, *Forecasting structural time series models and the Kalman filter*, Cambridge, Melbourne and New York: Cambridge University Press.

[19] Karanassou, M. and D.J. Snower, 1998 "How Labour Market Flexibility Affects Unemployment: Long Term Implication of the Chain Reaction Theory", *Economic Theory*, 108, 832-849.

[20] Koopman, S., 1997, "Exact initial Kalman filtering and smoothing for non-stationary time series models", *Journal of the American Statistical Association*, 92, 1630-1638.

[21] Koopman, S. and J. Durbin, 2000, "Fast Filtering and Smoothing for Multivariate State Space Models", *Journal of Time Series Analysis*, 21, 281-296.

[22] Koopman, S. and J. Durbin, 2003, "Filtering and Smoothing of State Vector for Diffuse State-Space Models", *Journal of Time Series Analysis*, 24, 85-98.

[23] Layard, R., S. Nickell and R. Jackmann, 1991, *Unemployment, Macroeconomic Performance and the Labour Market*,Oxford University Press, Oxford.

[24] Lindbeck, A. and D.J. Snower, 1987, "Efficiency Wages versus Insider and Outsiders", *European Economic Review*, 31, 407-416.

[25] Orlandi, F. and Pichelmann, K. 2000, "Disentangling Trend and Cycle in the EUR-11 Unemployment Series: An Unobserved Component Modelling Approach", European Commission Working Papers Series No. 140.

[26] OECD 1999, "Is the NAIRU a reliable concept in the EU context? Methodological lessons from Member States experience and the novelties of the EMU set-up" J. Elmeskov and Scarpetta, S., Seminar at the Commission of the European Communities, Brussels September 1999

[27] Okun 1970, "The Political Economy of Prosperity", Washington, 1970

[28] Pantula, S.G., 1991, "Asymptotic distributions of the unit-root tests when the process is nearly stationary", *Journal of Business and Economic Statistics*, 9, 325-353.

[29] Papell. D.H., C.J. Murray and H. Ghiblawi, 2000, "The structure of unemployment", *Review of Economics and Statistics*, 82, 309-315.

[30] Phelps, E., 1968, "Money-Wage Dynamics and Labor Market Equilibrium", *Journal of Political Economy*, 76, 678-711.

[31] Perron, P., 1990, "Testing for a unit root in a time series with a changing mean", *Journal of Business and Economic Statistics*, 8, 153-162.

[32] Pissarides, C., 1990, *Equilibrium Unemployment Theory*, Basil Blackwell, Oxford.

[33] Roed, K., 1997, "Hysteresis in Unemployment", *Journal of Economic Surveys* , 11, 389-418.

[34] Schorfheide, F., 2006, "Bayesian Methods in Macroeconomics", in *The New Palgrave Dictionary of Economics*, ed. by S.N. Durlauf and L.E. Blume, forthcoming, Palgrave MacMillan.

[35] Schwert, G.W., 1989, "Tests for unit roots: a Monte Carlo investigation," *Journal of Business and Economic Statistics*, 7, 147-160.

[36] Staiger, D., Stock, J. and M. Watson, 1997, "The NAIRU, Unemployment and Monetary Policy", *Journal of Economic Perspectives*, 11, 33-49.

[37] Stock, J.H. and M.W. Watson, 1998, "Median Unbiased Estimation of Coefficient Variance in a Time-Varying Parameter Model", *Journal of the American Statistical Association*, 93, 349-358.

[38] Young, P.C., Lane, K., Ng, C.N. and D. Palmer, 1991, "Recursive forecasting, smoothing and seasonal adjustment of nonstationary environmental data", *Journal of Forecasting*, 10, 57-89.

Appendix 1: Okun's Law

Assume firms (i) supply whatever is demanded and (ii) have a constant-returns technology of the form

$$y - k = \omega(n - k)$$

where k is capital stock and n is employment. Now, we define the potential output $\overline{y}$ as the level of output that corresponds to equilibrium level of unemployment:

$$\overline{y} - k = \omega(l - k - u^*)$$

where l is the labour force. Taking these two equations together and using the definition $u = l - n$ it follows

$$y - \overline{y} = -\omega(u - u^*) + \epsilon$$

The link between $\overline{y}$ and u^* is given by

$$\frac{\partial \bar{y}}{\partial u^*} = -\omega$$

Appendix 2: General State Space representation

$$y_t = \begin{bmatrix} y_t & u_t & p_t \end{bmatrix}'; \; x_t = \begin{bmatrix} p_{t-1} & p_{t-2} & u_{t-1} \end{bmatrix}';$$

$$\alpha_t = \begin{bmatrix} \overline{y}_t & \psi_t & u_t^* & \sigma_t^d & \phi_t & \tau_t & u_{t-1}^* \end{bmatrix}';$$

$$A = \begin{bmatrix} -2\frac{\omega\rho}{\theta_1} & \frac{\omega\rho}{\theta_1} & -\frac{\omega\theta_{11}\rho}{\theta_1} \\ 2\frac{\rho}{\theta_1} & -\frac{\rho}{\theta_1} & \frac{\theta_{11}\rho}{\theta_1} \\ 2\frac{\lambda_1\omega\rho}{\theta_1} & -\frac{\lambda_1\omega\rho}{\theta_1} & \frac{\lambda_1\rho\omega\theta_{11}}{\theta_1} \end{bmatrix};$$

$$z = \begin{bmatrix} \left(1-\frac{\lambda_1\omega\rho}{\theta_1}\right) & 0 & \frac{\omega\theta_{11}\rho}{\theta_1} & \frac{\omega\rho}{\theta_1} & 0 & 0 & 0 \\ \frac{\lambda_1\rho}{\theta_1} & 0 & \left(1-\frac{\rho\theta_{11}}{\theta_1}\right) & -\frac{\rho}{\theta_1} & 0 & 0 & 0 \\ -\lambda_1\left(1-\frac{\lambda_1\omega\rho}{\theta_1}\right) & 0 & -\frac{\lambda_1\omega\theta_{11}}{\theta_1} & \left(1-\frac{\lambda_1\omega\rho}{\theta_1}\right) & 0 & 0 & 0 \end{bmatrix};$$

$$T = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & (1+\delta) & 0 & 0 & 0 & -\delta \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}; \quad R = \begin{bmatrix} 1 & 0 & \omega & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} & \varepsilon_{2t} & \varepsilon_{3t} \end{bmatrix}'; \; \eta_t = \begin{bmatrix} \eta_{1t} & \eta_{2t} & \eta_{3t} & \eta_{4t} & \eta_{5t} & \eta_{6t} \end{bmatrix}';$$

$$H_t = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & 0 & 0 \\ 0 & \sigma_{\varepsilon_2}^2 & 0 \\ 0 & 0 & \sigma_{\varepsilon_3}^2 \end{bmatrix}; \; Q_t = \begin{bmatrix} \sigma_{\eta_1}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\eta_2}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\eta_3}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\eta_4}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\eta_5}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{\eta_6}^2 \end{bmatrix}$$

Figure 3: Simulated Output Gap in the euro area

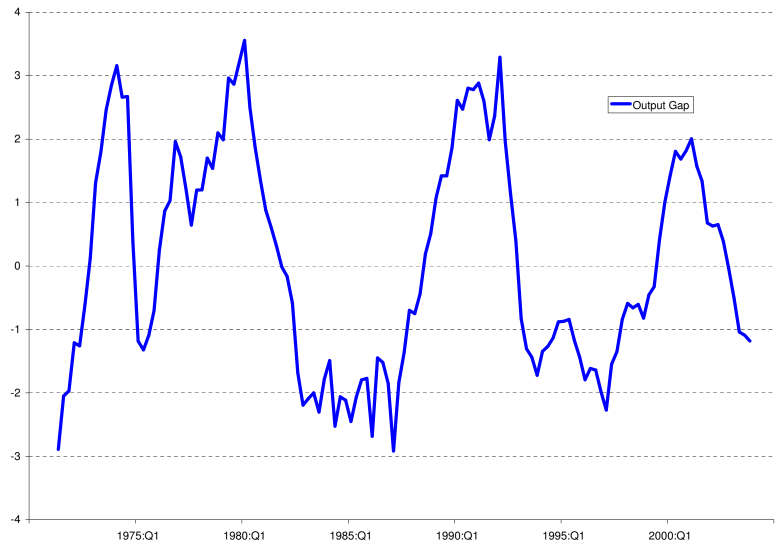


Figure 4: Simulated Output Gap in the US

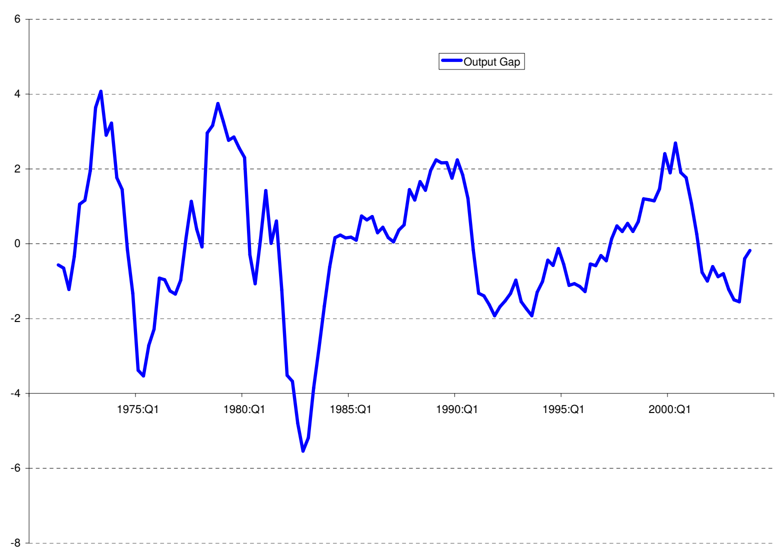


Figure 5: Inflation and Demand in the euro area

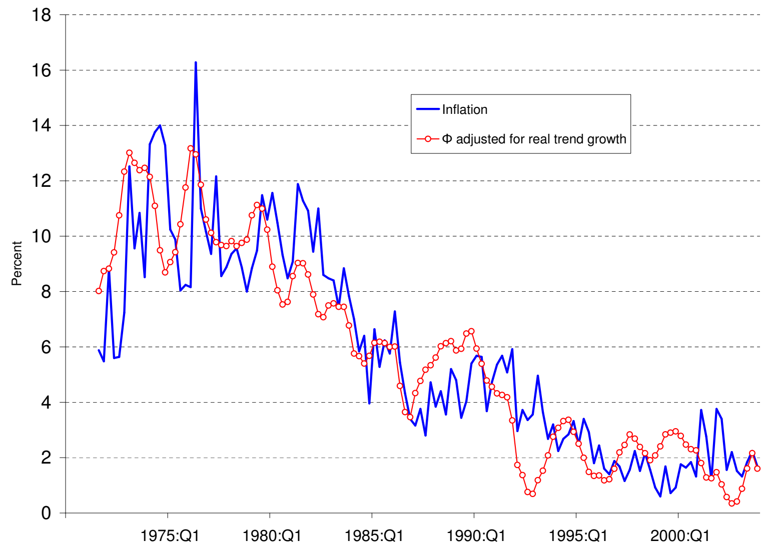


Figure 6: Inflation and Demand in the US