

Estimating the natural rate of unemployment in euro-area countries with co-integrated systems*

by Sven Schreiber[†]

this version: June 2010

Running title: “The natural rate of unemployment in euro-area countries”

Abstract

Given that for France, Germany, Italy, and the Netherlands the unemployment rates are best classified as $I(1)$, we apply permanent-transitory decompositions based on cointegrated VARs with relevant variables (labor productivity, wages, tax wedges, foreign relative prices) to estimate the time-varying natural unemployment rates. In general all variables seem to matter, and the results are quite different from published OECD Nairus. Our implied unemployment gaps are better than the OECD gaps in predicting unemployment changes and inflation gaps, but they are (except for Italy) as bad as the OECD gaps for forecasting inflation changes.

Keywords: euro-area unemployment, VECM, permanent-transitory decomposition

JEL: C51 (model construction), E24 (employment), J23 (employment determination)

*Work on this project started during a visit of the author at the Deutsche Bundesbank, and the hospitality I enjoyed there is gratefully acknowledged. This paper has benefited from comments by anonymous referees and by seminar participants at the Bundesbank, at Humboldt University Berlin, at the ESEM and VfS meetings, and at DIW Berlin; I wish to thank especially Dieter Nautz, Ragnar Nymoen, Christian Schumacher, and Harald Uhlig for useful suggestions.

[†]Macroeconomic Policy Institute (IMK) at Hans Boeckler Foundation, and Goethe University Frankfurt. E-mail: svetosch@gmx.net, fax: +49(0)211-7778-4332, address: Hans-Böckler-Stiftung/IMK, Hans-Böckler-Str. 39, D-40476 Düsseldorf, Germany.

1 Introduction

In this paper we use co-integrated multivariate systems to estimate the time-varying natural unemployment rates of important euro-area countries. The most common measures of the natural rate of unemployment are Nairu, especially those from the OECD (see Richardson et al., 2000, sometimes in the guise of wage-inflation-based “Nawrus”). Those are built around a single-equation Phillips-curve framework, often with some further exogenous variables in the spirit of the “triangle model” of inflation (see Gordon, 1997). Because of the smooth and rising paths of many unemployment rates, a random walk is a popular and hard-to-beat model for the unobserved Nairu component, even though it is clear that such a model can only serve as an empirical approximation in a given finite sample.¹ Given this widespread empirical $I(1)$ assumption for unemployment it seems obvious that the corresponding permanent components should be estimated in a co-integrated multivariate system, by using the duality between co-integration and common stochastic trends. However, to the best of our knowledge this natural co-integrated-VAR step has not been taken in the literature so far. (For example, Fabiani and Mestre, 2004, use a VAR-based method, but they did not find co-integration, and thus could not extract any non-trivial permanent components from the data.) In this paper we therefore specify co-integrated systems and then extract the time-varying natural unemployment rates as the permanent components based on the estimated common trends (Stock and Watson, 1988), in the form suggested by Proietti (1997). As an alternative measure, we also use the Gonzalo-Granger decomposition (Gonzalo and Granger, 1995). A case for using permanent-transitory decompositions has recently been made by Garratt et al. (2006).

By choosing this approach we take into account that “the determination of the joint steady state of real wages, the real exchange rate, and the rate of unemployment requires a full dynamic model rather than wage and price setting equations alone” (Bårdsen and Nymoén, 2003). It is also interesting because the non-stationary development of unemployment in many European countries is still a puzzle, i.e., we lack “a satisfactory empirical explanation” for it (Nickell, 1998, p. 802). Our method adequately reflects the implicit equilibrium relations between the variables, and in contrast to typical unobserved-components models such as those of the OECD there is no need for ad-hoc smoothness restrictions.²

¹Sometimes even a double random walk specification is used (i.e., random walk with stochastic drift, where the drift itself is again a random walk), see for example Laubach (2001). However, such an $I(2)$ specification would imply that unemployment changes themselves are non-stationary and therefore would seem quite extreme.

²It must be acknowledged, however, that there is also a recent literature using more complex multivariate unobserved-components models to estimate potential output and/or the Nairu where a-priori smoothness restrictions are not always needed; see for example Apel and Jansson (1999, using Sweden for illustration), Proietti et al. (2007, dealing with aggregate euro-area data), and Basistha and Startz (2008, for the US).

Since a prerequisite for applying permanent-transitory decompositions is a development of unemployment which is best classified as $I(1)$, all countries had to be excluded where the unemployment rate is stationary, possibly after the removal of some breaks. Information on unit roots in OECD unemployment rates is given in Papell et al. (2000), who find evidence against $I(1)$ -ness in several countries after allowing for mean shifts. The remaining countries where the unit root cannot be rejected are Germany, France, Italy, and the Netherlands (apart from Japan).³ These four European countries that we include in our study represent the largest part of the euro area in terms of output and population.

The structure of this paper is as follows: The next section 2 points to the theory that guides our choice of relevant variables, and clarifies the details of our econometric framework, including an illustrative example in a stripped-down bi-variate system. Afterwards, section 3 provides descriptive evidence of the variables in the four countries and tests for the number of common trends in each country. Section 4 presents the estimated permanent unemployment components as the main results of this study, compares them to the OECD Nairus, and provides an evaluation of their predictive powers. Finally, section 5 offers conclusions.

2 General modeling and specification issues⁴

2.1 Relevant literature and variables

We base the choice of variables on the established literature, of which we only mention a small selection here: First there is the influential book by Layard et al. (1991), mainly based on bargaining theories of the labor market. Another important reference is the discussion of the wage curve and of differences with respect to Phillips curve formulations in Blanchflower and Oswald (1995). Blanchard and Katz (1999) provide a short but very useful discussion of similar issues, and clarify how a wage curve formulation also covers more modern search and matching frameworks. The issue of labor taxes affecting unemployment or not is discussed, e.g., in Pissarides (1998), and a modern formulation of the Phillips curve framework in an open economy is given in Batini et al. (2005). From the

But with these complex models technical problems like non-convergence of the algorithms are common. Our approach does not require iterative numerical algorithms because closed-form algebraic formulas are available.

³See appendix A for additional unit root test results.

⁴The empirical results in this paper have been produced with gretl 1.8.x (Cottrell and Lucchetti, 2009, for tests and estimates of co-integration restrictions and tests on α), gretl scripts and functions written by the authors (for the common-trends tests, tests on α_{\perp} , and the decompositions), and with JMulTi (for some rank and stability tests as well as unit root tests with breaks, see Krätzig, 2004). Some of the custom gretl codes are publicly available as packages from the official gretl function package server; the remaining ones are available from the author.

empirical modeling side we highlight the comprehensive book by Bårdsen et al. (2005) that deals with wage-price-unemployment systems in great detail. Having been guided by these strands of the literature, the variables that enter the multivariate systems in this paper are given by real wages and labor productivity, the domestic inflation rate, some price wedges (especially those between domestic and imported manufactured goods), and a tax wedge between gross and net wages.

2.2 Framework

Our baseline models deal with the following variable vector of dimension $N = 6$, separately for all countries:

$$y_{i,t} = (u_{i,t}, \pi_{i,t}, fx_{i,t}, tw_{i,t}, w_{i,t}, q_{i,t})', \quad (1)$$

where $i \in \{DE, FR, IT, NL\}$ indexes the countries, u is the unemployment rate, π is the annualized quarterly inflation rate, fx are foreign relative prices, and tw is a tax wedge. Details on all variables are given in section 3. The statistical framework is the well-known co-integrated VAR as described by Johansen (1995), which can be represented as the following vector error correction models (VECM) for the different countries:

$$\Delta y_{i,t} = \alpha_i \beta_i' y_{i,t-1} + \sum_{j=1}^{k_i-1} \Gamma_{i,j} \Delta y_{i,t-j} + \mu_i + D_i \delta_{i,t} + \varepsilon_{i,t} \quad t = 1, \dots, T_i \quad (2)$$

In this standard notation β_i is the (country-specific) co-integration matrix of dimension $6 \times r_i$ with full column rank, where $r_i < 6$ is the co-integration rank. The full-rank matrix α_i also has dimension $6 \times r_i$, and thus the long-run matrix $\alpha_i \beta_i'$ is of reduced rank. Note that the individual co-integration vectors need not be identified in any particular way for the purpose of our study, because only the complete co-integration space spanned by the columns of β_i matters, and that space is always identified by the Johansen procedure.

The number of $I(1)$ trends driving each system is $6 - r_i = K_i$. The constant term μ_i is left unrestricted, and thus it will accumulate to a linear trend in the non-stationary directions of the process $y_{i,t}$; this trend component would be expected in labor costs, productivity, and possibly other variables such as the external price wedge. The additional regressor $\delta_{i,t}$ (with associated coefficients D_i) contains further deterministic terms such as outlier-removing impulse dummies, and a level shift dummy in the case of Germany.⁵

⁵However, the shift will be restricted to the stationary directions by forcing its coefficient to be of the form $\alpha_i \tau_i$, where only the part τ_i is unrestricted. Put differently, a broken trend in $y_{i,t}$ will be ruled out a priori. Another way of formulating that restriction is to replace $y_{i,t-1}$ and β_i by $y_{i,t-1}^* = (y'_{i,t-1}, s_{i,t-1})'$ and $\beta_i^* = (\beta'_i, \tau_i^*)'$ in the VECM.

However, no linear trend term is included there, because we wish to model the equilibrium relationships between the variables without resorting to unexplained exogenous terms as much as possible. Note that this specification renders a co-integration rank of $r_i = N - 1 = 5$ implausible a priori, because such a result would imply pairwise co-integration between *all* variables, including pairs such as inflation and deterministically trending productivity levels.

A nested specification would be to impose a long-run 1:1-proportionality between productivity and wages. This restriction can be formulated as the following linear hypothesis:

$$\beta_i = \begin{bmatrix} 0_{4 \times 1} & I_4 \\ 1 & \\ -1 & 0_{2 \times 4} \end{bmatrix} \phi_i, \quad (3)$$

where ϕ_i is unrestricted. This restriction means that the co-integration space could be simplified by using the (log) labor share $w_{i,t} - q_{i,t}$ instead of wages and productivity separately. If hypothesis (3) were accepted, one could also consider testing additional short-run homogeneity restrictions, but it turned out that tests reject the unit proportionality restriction (3), see below; therefore wages and productivity have to be considered separately in the models.

Our approach in this paper is to *not* assume many independent stochastic trends a priori, for example because modern macroeconomic theory typically specifies only few common stochastic trends – for the extreme case of only a single trend see King et al. (1991). For the determination of the co-integration rank we therefore complement the popular Johansen test procedure with appropriate common-trends test suggested by Nyblom and Harvey (2000). While the Johansen test starts from the null hypothesis that there are a maximum number of common trends (no co-integration), the Nyblom-Harvey test is a multivariate generalization of the uni-variate KPSS stationarity test (see Kwiatkowski et al., 1992); as such it tests the hypothesis of *few* stochastic trends against the alternative of a higher number, i.e., the null and alternative hypotheses are reversed.

2.3 Extracting permanent components

Given an estimated number of common stochastic trends in each system, the main goal of this study is to estimate the permanent component of each unemployment rate. We use two established permanent-transitory decompositions for that, first the multivariate generalization of the Beveridge-Nelson decomposition based on the estimated common trends, and secondly the one by proposed by Gonzalo and Granger (1995). Usually the common trends are represented in the moving-average form used by Stock and Watson (1988), but for reasons to be explained below we will use a different approach suggested

by Proietti (1997).

Starting with the Gonzalo-Granger (GG) decomposition for country i , it is defined as

$$y_{i,t} = \beta_{i\perp}(\alpha'_{i\perp}\beta_{i\perp})^{-1}\alpha'_{i\perp}y_{i,t} + \alpha_i(\beta'_i\alpha_i)^{-1}\beta'_iy_{i,t}, \quad (4)$$

where the orthogonal complements $\alpha_{i\perp}$ and $\beta_{i\perp}$ are $N \times (N - r_i)$ -matrices such that $\alpha'_{i\perp}\alpha_i = 0 = \beta'_{i\perp}\beta_i$. Note that although the choice of the orthogonal complements is not unique, the concrete choice is irrelevant for the decomposition in (4). Given that $\beta'_iy_{i,t}$ are the stationary error-correction terms, the transitory component of $y_{i,t}$ is obviously given by $\alpha_i(\beta'_i\alpha_i)^{-1}\beta'_iy_{i,t}$, whereas the permanent part consists of the common factors $\alpha'_{i\perp}y_{i,t}$ that drive the system and are loaded into the variables by the coefficients $\beta_{i\perp}(\alpha'_{i\perp}\beta_{i\perp})^{-1}$.

Given the additive nature of the identity (4) and the fact that no deterministic and/or exogenous terms appear there, it is clear that its transitory part is only stationary around the path of any restricted terms of the model (2), and that the absolute level of the permanent part does not automatically correspond to the levels of $y_{i,t}$. For our purposes this is inconvenient, and therefore we have adjusted the components by the following two steps. First of all, we correct for the mean of the stationary directions, which is known to be (Johansen, 1995)

$$(\alpha'_i\alpha_i)^{-1}\alpha'_i(\Psi_i C_i - I)\mu_i, \quad (5)$$

where $\Psi_i = I - \sum_{j=1}^{k-1}\Gamma_{i,j}$, and $C_i = \beta_{i\perp}(\alpha'_{i\perp}\Psi_i\beta_{i\perp})^{-1}\alpha'_{i\perp}$ is the long-run impact matrix of reduced rank $N - r_i$. Thus in the first step we subtract the estimate of (5) pre-multiplied by the relevant loadings $\alpha_i(\beta'_i\alpha_i)^{-1}$ from the transitory component and add that to the permanent component to preserve the additivity of the decomposition. Secondly, to account for any exogenous terms restricted to the co-integration space –i.e., the step dummy for Germany– we compare the raw error correction terms $\beta'_iy_{i,t}$ to the augmented ones including the restricted terms $\beta_i^*y_{i,t}^*$, and again subtract that difference times the relevant loadings from the transitory component and add it to the permanent component.

Altogether, the resulting permanent GG-component of $y_{i,t}$ looks as follows:

$$y_{i,t}^{GG} = \beta_{i\perp}(\alpha'_{i\perp}\beta_{i\perp})^{-1}\alpha'_{i\perp}y_{i,t} + \alpha_i(\beta'_i\alpha_i)^{-1}((\alpha'_i\alpha_i)^{-1}\alpha'_i(\Psi_i C_i - I)\mu_i + \beta'_iy_{i,t} - \beta_i^*y_{i,t}^*), \quad (6)$$

where in our application the natural unemployment rate is simply the first element of the vector.

The main identifying feature of the GG decomposition within the class of permanent-transitory decompositions is that the permanent component is a linear combination of the contemporaneous observables, which renders the interpretation of the permanent factors easier. However, this implies that in general the GG permanent component is auto-

correlated in differences, which means that the permanent component will change even if no new shocks hit the system. This property makes it appear less useful as an equilibrium measure.

Another popular approach is therefore to find the multivariate random-walk component of the variable set, based on the Stock-Watson (SW) common trends. The random-walk property of this multivariate Beveridge-Nelson decomposition means that changes of its permanent component are unpredictable. Thus it provides the long-run forecast for the included variables, given the information of the current period. Of course, the GG and SW permanent components only differ by stationary terms and must be co-integrated, therefore they share the same long-run features. Note that the GG and SW decompositions actually coincide if there are common cycles in the data, and if their number is equal to the co-integration rank (Proietti, 1997).

Given the coefficients of the VECM (2), a fixed initial value, and indicating estimated magnitudes with hats, the permanent SW components are given by

$$\widehat{y}_{i,t}^{SW} = y_{i,0} + \widehat{C}_i \widehat{\mu}_{it} + \widehat{C}_i \sum_{s=1}^t \widehat{\varepsilon}_{i,s}. \quad (7)$$

However, we do not apply the formula (7) directly, because the common trends must be estimated from the residuals of the model, and they are therefore heavily influenced by outliers and other features that the empirical model may not have fully captured. We therefore follow Proietti (1997) and instead determine the common-trends based permanent components as a distributed lag of the observable variables:

$$y_{i,t}^{SWP} = (I - P_i)(\Gamma_i^*(1) + \alpha_i \beta_i')^{-1} \Gamma_i^*(L) y_{i,t} \quad (8)$$

where $P_i = (\Gamma_i^*(1) + \alpha_i \beta_i')^{-1} \alpha_i [\beta_i' (\Gamma_i^*(1) + \alpha_i \beta_i')^{-1} \alpha_i]^{-1} \beta_i'$, with $\Gamma_{i,j}^* = \Gamma_{i,j} + \alpha_i \beta_i'$, and $\Gamma_i^*(L) = I - \sum_{j=1}^{k_i-1} \Gamma_{i,j}^* L^j$. For any further terms such as the German level shift dummy, we apply the same adjustment as for the GG decomposition.

2.4 A bi-variate illustration of the approach

Before dealing with the full six-dimensional systems we demonstrate our approach in a small bi-variate system. For this purely illustrative purpose consider the Dutch unemployment rate and the tax wedge with $k = 2$ lags, such that in addition to the co-integration and adjustment coefficients only one matrix of short-run coefficients needs to be estimated. For the sake of the illustration we will ignore the remaining residual auto-correlation, and also simply impose the co-integration rank of one here. The resulting system estimate

with $\widehat{\beta}$ normalized on the unemployment rate is as follows:

$$\Delta \begin{bmatrix} u_{NL,t} \\ tw_{NL,t} \end{bmatrix} = \begin{bmatrix} -0.022 \\ 0.114 \end{bmatrix} \begin{bmatrix} 1 & -0.111 & 3.736 \end{bmatrix} \begin{bmatrix} u_{NL,t-1} \\ tw_{NL,t-1} \\ 1 \end{bmatrix} \\ + \begin{bmatrix} 0.958 & 0.001 \\ 1.072 & 0.166 \end{bmatrix} \Delta \begin{bmatrix} u_{NL,t-1} \\ tw_{NL,t-1} \end{bmatrix} + \widehat{\varepsilon}_t$$

Note that here the constant is restricted to the co-integration relations, in contrast to the specification of the larger systems below. The needed quantities for the SWP components are

$$\widehat{\Gamma}^*(L) = I - \begin{bmatrix} 0.936 & 0.003 \\ 1.186 & 0.153 \end{bmatrix} L, \\ (I - \widehat{P}) \left(\widehat{\Gamma}^*(1) + \widehat{\alpha} \widehat{\beta}' \right)^{-1} = \left(I - \begin{bmatrix} 1.129 & -0.125 \\ 1.164 & -0.129 \end{bmatrix} \right) \times \begin{bmatrix} 24.561 & 0.029 \\ 31.570 & 1.237 \end{bmatrix} \\ = \begin{bmatrix} 0.784 & 0.151 \\ 7.060 & 1.362 \end{bmatrix},$$

such that the SWP random-walk component for the unemployment rate would be given by $\widehat{u}_{NL,t}^{SWP} = 0.784u_{NL,t} + 0.151tw_{NL,t} - 0.913u_{NL,t-1} - 0.026tw_{NL,t-1}$, if we abstract from the restricted constant term for simplicity here. This could alternatively be expressed as $\widehat{u}_{NL,t}^{SWP} = -0.129u_{NL,t} + 0.125tw_{NL,t} + 0.913\Delta u_{NL,t-1} + 0.026\Delta tw_{NL,t-1}$. Taking into account the different scales and variations of the two variables (see section 3) we see that the bulk of the long-run variations in the unemployment rate is attributed to the tax wedge in this (merely illustrative) case.

Indeed, the adjustment coefficient in the tw_{NL} -equation is insignificant with a t-ratio of 1.12, which suggests that the tax wedge here is weakly exogenous. Imposing this zero-restriction on the second row of α would imply that $\alpha_{\perp} = (0; 1)'$, and hence the permanent shocks would be given by the innovations in the tax wedge equation alone. Equivalently, the permanent factor in the Gonzalo-Granger sense would be given by the tax wedge, since here $\alpha'_{\perp}(u_{NL,t}, tw_{NL,t})' = tw_{NL,t}$. Furthermore, in a bi-variate system with $r = N - r = 1$ there is an algebraic symmetry between α and α_{\perp} which implies that the χ^2_1 tests of weak exogeneity mirror the tests of exclusion from the Gonzalo-Granger factors. Thus the p-values of the weak exogeneity test on the tax wedge would be identical here to the ones of the factor exclusion test on unemployment. However, in a higher-dimensional system there are in general more possibilities.

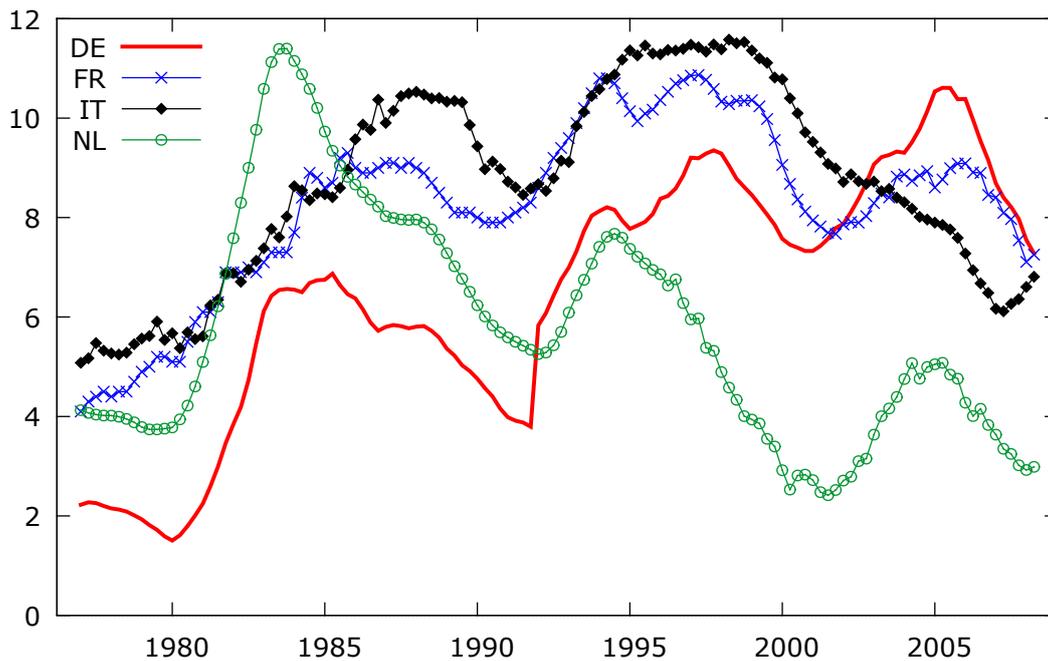


Figure 1: Unemployment rates

3 Descriptive evidence and rank tests

3.1 Time series developments

Because we chose to exclude the ongoing financial and economic crisis from our sample, the analyzed time period spans 1977q1 through 2008q2 ($T = 126$). See the appendix A for unit root tests of the included variables.

Unemployment: The developments of unemployment in France and Italy are quite similar in the long run (see figure 1). Due to steep increases after unification and during the recession of 2000-2005 Germany has recently caught up with France and Italy. The Dutch unemployment rate presents a completely different picture, showing a downward trend for more than two decades now. Therefore the Dutch case will be interesting as a cross-check for other countries' findings.

Wage and productivity dynamics: Labor productivity and real hourly labor costs themselves are not shown because they are in general smoothly growing and therefore they are difficult to analyze visually. As a closely related substitute graph consider the labor shares in the four countries in figure 2, which have been calculated as the gap between (log) labor costs and (log) labor productivity. From the downward developments we see that on average labor productivity has grown faster than real labor costs. The Italian share displays somewhat different behavior in that it has been especially low since the mid 1990s, which is partially due to reforms that lowered the social security contributions.

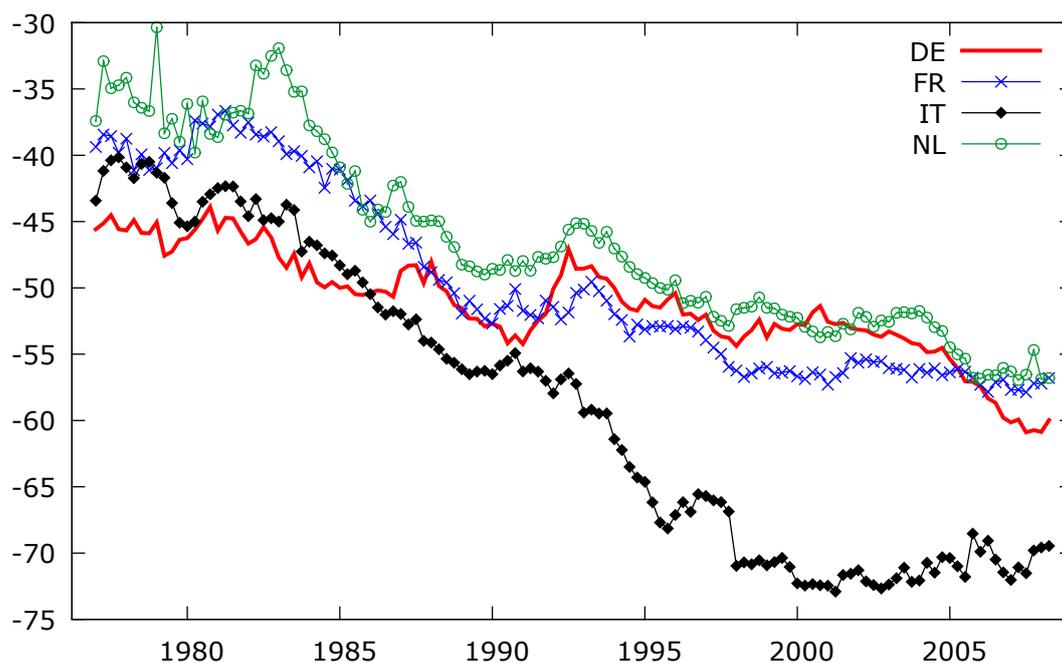


Figure 2: Labor shares. Calculated as log hourly real labor costs minus log hourly productivity, times 100. (The individual components labor costs and productivity are not restricted to form the labor share in the empirical model, see below.)

Inflation dynamics: All inflation series except perhaps the Dutch data exhibit the typical gradual disinflation process that implies non-stationarity of inflation within the sample (see figure 3). The degree of disinflation is highest for Italy.⁶

Relative external prices: Interestingly, the external price wedges as measured by the ratios of domestic prices and prices of imports (excluding raw materials) display some common features (see figure 4): in all countries domestic prices have risen much more in the last 30 years than the prices of those imports that are likely to compete with home goods. This would indicate a considerable fall in competitiveness of domestic producers, although part of it could reflect Balassa-Samuelson-type effects.

Tax and social security wedges: The wedge between product and consumer wages⁷ is often discussed as a determinant of real wage resistance and/or unemployment. Unfortunately, a conceptual problem of using any kind of wedges to explain labor market slack is the fact that most of the government revenue is not simply dead-weight loss. Apart from

⁶The nominal (hourly) wage inflation displays roughly the analogous behavior, albeit with more short-term volatility. The inflation of imported raw materials (in local currency) is virtually identical for all four countries, since the price volatility dwarfs any exchange rate discrepancies. Also, it is clearly stationary. The inflation series of non-commodity imports is much more volatile than the domestic inflation series and contains a less pronounced downward tendency over the longer run. (All those series and results are not shown to save space.)

⁷Or more precisely but also more involved: the wedge between total labor costs deflated by some producer price index and the after-tax wage income deflated by the consumer price index.

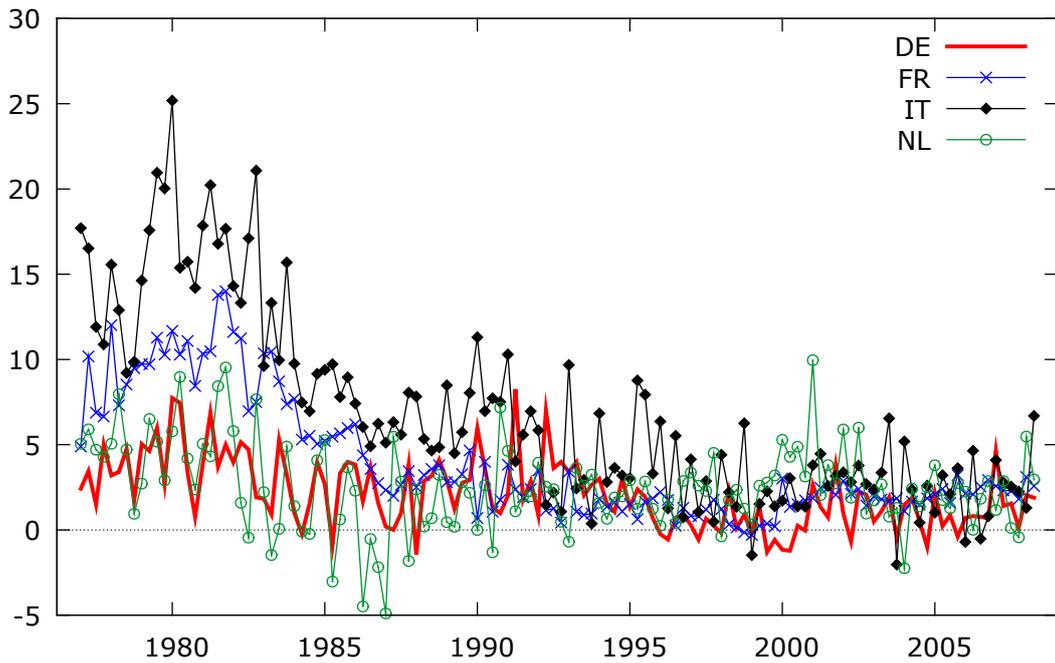


Figure 3: Domestic inflation. Annualized growth rate of GDP deflator.

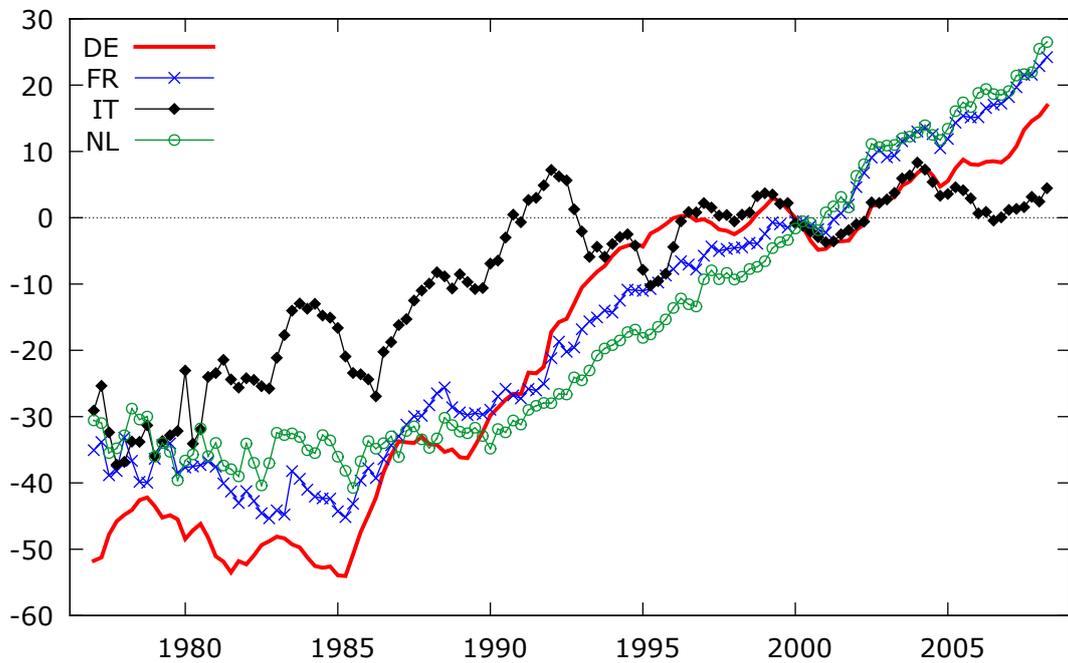


Figure 4: External price wedge. GDP deflator relative to price index of imports excluding raw materials (partially following Bårdsen et al., 2005), in logs and times 100 to (roughly) read off percentage changes; base year 2000. Import prices are measured in local currency and thus account for nominal exchange rate movements.

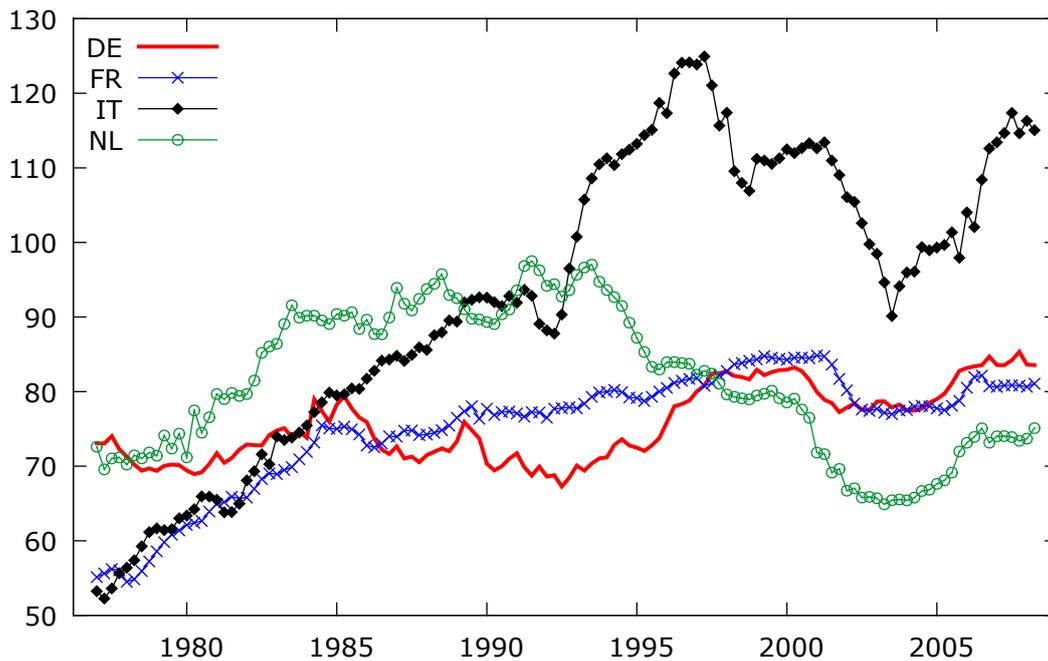


Figure 5: The tax wedge. Labor costs relative to net wages, real and in logs, times 100. This indirect measure of the total wedge includes labor taxes as well as social security and insurance contributions. (It also accounts for the price wedge between consumer and producer prices, thereby also capturing changes of indirect taxes, but the variation of the price wedge is small over the sample compared to the other components.)

the provision of public goods, a large part is used for compulsory insurance of sickness, unemployment, and retirement, all of which have intrinsic value for economic agents and would have to be privately purchased out of the gross labor income, too.⁸ Nevertheless, given the difficulty of disentangling the various degrees of inefficiencies of the tax and expenditure systems, we will use a comprehensive but widespread wedge concept, where all the aforementioned considerations are ignored.

Figure 5 shows the development of the log-ratio of total labor compensation paid by employers relative to immediately disposable real labor income. We see that over the sample this measure rose in all countries except the Netherlands, where it rebounded from a maximum in the early 1990s. (Italy's wedge also rebounded in 1998 after pension system reforms, but has not recovered to the levels of the 1980s.)

⁸This reasoning does not negate the fact that inefficiencies of the public insurance schemes represent an excess burden. Pension contributions for example surely imply a tax component since an individual could get higher (risk-adjusted) rates of return on the world capital market. However, the resulting loss in each period is an order of magnitude smaller than the total contribution value.

3.2 Determining the number of stochastic trends

Germany (DE): Germany requires special modeling because of the unification break. Busetti (2002) extends the Nyblom-Harvey test to deal with deterministic shifts at known time, which exactly corresponds to the German test problem. The corresponding co-integration test that deals with such breaks but does not require separate sets of critical values is the Saikkonen and Lütkepohl (2000) test. For estimation of the VECM, we include a level shift dummy in the model that is restricted to the co-integration space, and further (lagged) impulse dummies for the year 1991, eliminating the transition phase from the sample. In order to capture the auto-correlation features of the data we chose a lag length of $k_{DE} = 3$ which is slightly above the indication of only two lags by information criteria. There is an outlier in 1984q2 which we eliminate by a corresponding impulse dummy. The characteristic roots (see table 1) then show two very clear unit roots, and two further potential unit roots.

Indeed, the formal Busetti (2002) tests in table 2 clearly reject the existence of up to only two stochastic trends in favor of more than two unit roots. However, the significance level at which the hypothesis $6 - r_{DE} = 3$ (against $6 - r_{DE} > 4$) can be rejected already depends somewhat on the lag truncation choice.⁹ Nevertheless up to five common trends seem easily compatible with the data. Turning to the more standard co-integration tests, in table 3 we see that the nominal results indicate a co-integrating rank of $r_{DE} = 3$. The contradictory results of the two test types suggest to choose a rank 2 or 3. Stability tests and recursive estimates suggest that the lower rank $r_{DE} = 2$ is a safer choice.

France (FR): As for Germany diagnostic tests indicated a lag length of $k_{FR} = 3$ (in levels, as before). The characteristic roots of the system show three clear unit roots, and one or two further roots not too far from unity. The Nyblom-Harvey common-trends test is somewhat more significant than in the German case, in the sense that even the hypothesis $6 - r_{FR} = 4$ against $6 - r_{FR} > 4$ is rejected at the 5% level for reasonable lag truncations. The “standard” co-integration tests (again see table 3) indicate a rank of $r_{FR} = 2$. Given the choice between rank 1 or 2 for the French data we opt for $r_{FR} = 2$ because the estimates are sufficiently stable, and also because it is useful to treat the countries in a similar way, if at all possible.

Italy (IT): The lag length recommendations of information criteria range from 1 through 5, and it turns out that a choice of $k_{IT} = 4$ is needed to satisfy diagnostic tests. Apart from

⁹The lag truncation determines the lag window that is used to calculate the estimated long-run variance of the time series in the nonparametric correction of the test statistic for auto-correlated series. It is a well-known problem that the choice of that bandwidth parameter often affects the results substantially in real-world samples. Typically, too small a lag truncation value implies an oversized test, whereas a value that is too large reduces power considerably. Unfortunately, in contrast to the uni-variate case, in a multivariate setting the literature does not provide a data-based procedure to choose an optimal bandwidth.

Table 1: Largest characteristic roots of the models

Germany		France		Italy		Netherlands	
values	mod.	values	mod.	values	mod.	values	mod.
$0.976 \pm 0.010i$	0.976	0.992		$0.963 \pm 0.008i$	0.963	0.997	
$0.923 \pm 0.124i$	0.931	$0.971 \pm 0.036i$	0.972	$0.933 \pm 0.123i$	0.941	0.959	
$0.760 \pm 0.259i$	0.803	$0.881 \pm 0.139i$	0.892	$0.788 \pm 0.363i$	0.868	$0.924 \pm 0.127i$	0.933

Notes: Roots of the companion matrix of the respective systems. All countries: $N = 6$ variables in the system, no trend as regressor. For system specification details see the text.

two clear unit roots there exist at least two more potential unit roots in this model. The common-trends yield roughly the same results as for France, whereas the co-integration tests strike us as implausible: They would tend to indicate even a rank of 5, implying pairwise co-integration between such variables as productivity levels and inflation. Also, for a rank $r_{IT} = 3$ the stationarity of unemployment is only barely rejected (see table 4). Therefore we follow the French example and choose $r_{IT} = 2$ again.

Netherlands (NL): Here three lags ($k_{NL} = 3$) was sufficient to model the dynamics of the data, although this again slightly exceeds the nominal recommendation of information criteria. After removing two outliers in the system (1979q1 and 1980q2) by means of impulse dummies, in the resulting system there are two clear unit roots, whereas the status of the two remaining potential unit roots is borderline. The Nyblom-Harvey test produces somewhat less significant results than for the other countries. The Johansen co-integration test indicates a rank $r_{NL} = 3$, and given the satisfactory stability of the estimates that is the rank we choose for the Dutch data.

With these choices of the co-integration rank the estimated co-integration spaces seem reasonably stable over the analyzed samples, see figure 6.

4 Estimation results

In this section we present the estimated paths of the natural unemployment rates as implied by the permanent-transitory decompositions. We also discuss the identified co-integration relations (4.3) and perform a quantitative evaluation of the estimated natural components (4.4).

Table 2: Common-trends tests

lag trunc.	Dimension of stochastic trends, $H_0 : K = K_0$ vs. $H_1 : K > K_0$						
	0	1	2	3	4	5	
DE	1	2.097***	0.901***	0.539***	0.229***	0.077**	0.019
	2	1.470***	0.638***	0.382***	0.170***	0.065*	0.022
	3	1.159***	0.510***	0.308***	0.144**	0.063*	0.0267
	4	0.969***	0.432***	0.263***	0.127**	0.059*	0.0270*
FR	1	2.506***	1.037***	0.544***	0.294***	0.136***	0.041**
	2	1.737***	0.740***	0.390***	0.213***	0.103***	0.034*
	3	1.350***	0.590***	0.313***	0.175***	0.087**	0.032*
	4	1.119***	0.501***	0.270**	0.154**	0.080**	0.031*
IT	1	2.641***	1.107***	0.539***	0.303***	0.123***	0.046**
	2	1.831***	0.790***	0.387***	0.224***	0.097**	0.035**
	3	1.426***	0.631***	0.312***	0.185***	0.085**	0.030*
	4	1.181***	0.534***	0.267**	0.161**	0.076**	0.027
NL	1	2.502***	0.985***	0.458***	0.230***	0.109***	0.022
	2	1.732***	0.702***	0.338***	0.177***	0.089**	0.024
	3	1.345***	0.559***	0.277**	0.151**	0.078**	0.024
	4	1.109***	0.472***	0.239**	0.132*	0.070*	0.023

Notes: Results of Nyblom and Harvey (2000)-type tests for given number of common trends against a higher number under the alternative, allowing for deterministic trends. Buseti (2002)-tests with level shift (at a sample fraction of 0.48) for Germany. The underlying estimation of the long-run variance uses the Bartlett weight function with a lag window that is truncated at the indicated lag. Nominal significance indicated by *** (1%), ** (5%), and * (10%); all critical values for $N = 6$ taken from Buseti (2002).

Table 3: Co-integration tests

$H_0 :$		r=0	r=1	r=2	r=3	r=4
Germany	S&L	0	.006	.004	.093	.538
France	Joh	.002	.050	.199	.181	.177
	S&L	.001	.036	.228	.010	.208
Italy	Joh	0	0	.003	.011	.005
	S&L	.006	.010	.035	.036	.060
Netherlands	Joh	0	.001	.020	.263	.675
	S&L	.003	.007	.071	.413	.184

Notes: Numbers are asymptotic p-values. “S&L” indicates the Saikkonen and Lütkepohl (2000)-type test (performed with JMulti), “Joh” stands for the standard Johansen procedure. For samples and detailed system specifications see the text. It is not reasonable to test $H_0 : r = 5$ in this setup (trending data, but no trend in the co-integration space). The Johansen test is not applicable to the German case with level shifts, see the text.

Table 4: Nominal test results against stationarity of unemployment

rank r	Germany	France	Italy	Netherlands
3	.008	.034	.041	0
2	.003	.003	.002	0

Notes: Only p-values are given, the null hypothesis is always $H_0 : u_{i,t} \sim I(0)$ in the sense that the unit vector picking unemployment lies in the co-integration space. We use these test results as additional guidance to choose ranks for which unemployment actually comes out as $I(1)$, because otherwise our exercise would be logically inconsistent. Only test results under the true co-integration rank are meaningful, therefore nominal acceptance of stationarity in the other cases should not be viewed as evidence against $I(1)$ -ness of unemployment.

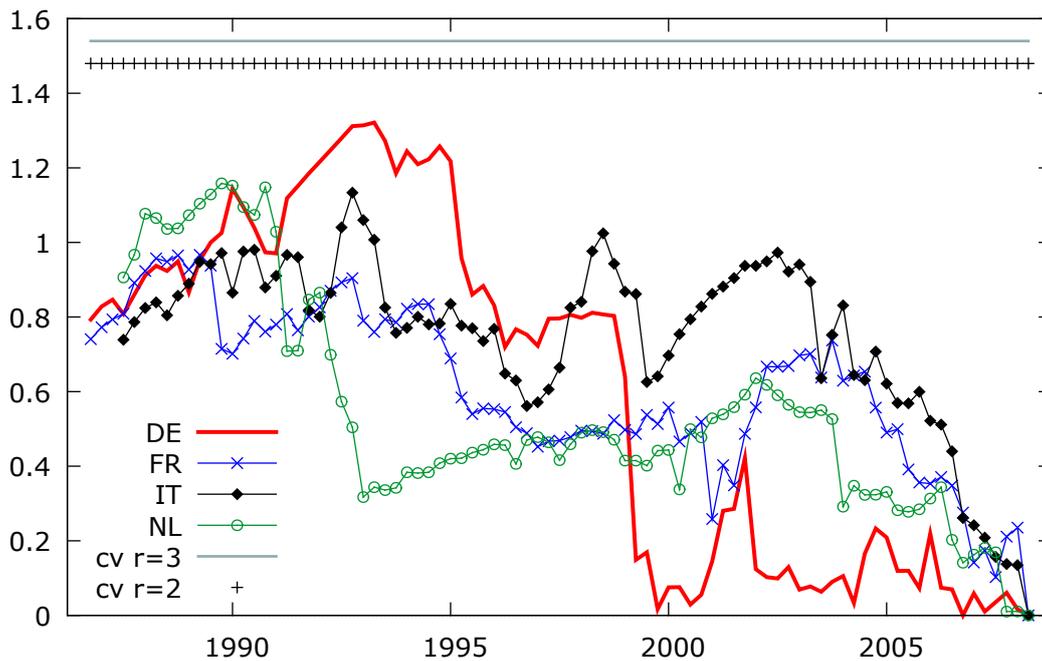


Figure 6: Stability of co-integration spaces. 5% critical values (for ranks 2 and 3) displayed for comparison; systems as described in the text.

4.1 Permanent unemployment components

Germany (figure 7): The equilibrium unemployment estimates track the long-run behavior of the unemployment rate quite well; however, the Gonzalo-Granger (GG) measure may coincide with observed unemployment rates a little too much to be useful.

In the following figure 8 we also show the OECD Nairu estimate and display a confidence band for the SWP long-run component of unemployment. We see that the observed unemployment rate falls outside the confidence band during the initial rise of German unemployment (1981/82) and especially in recent years, whereas elsewhere the estimated long-run component does not differ significantly from the observed rate. In contrast, the extremely smooth OECD Nairu is not statistically compatible with the co-integration-based estimate around 1980, 1990, and at the end of the sample. With respect to the current margin of the sample, the estimates disagree as to whether there has been excess unemployment recently in Germany before the crisis hit: the OECD measure implies that unemployment had declined too much, whereas the SWP permanent component indicates a further tendency of unemployment to fall in the absence of further shocks.

France (figure 9): Here the GG and SWP estimates are quite similar. The confidence band for the SWP component in figure 10 covers the observed unemployment rate almost all of the time, whereas the OECD Nairu falls outside that band at several occasions (early and mid 1990s). In general it turns out that the co-integration-based measures appear more

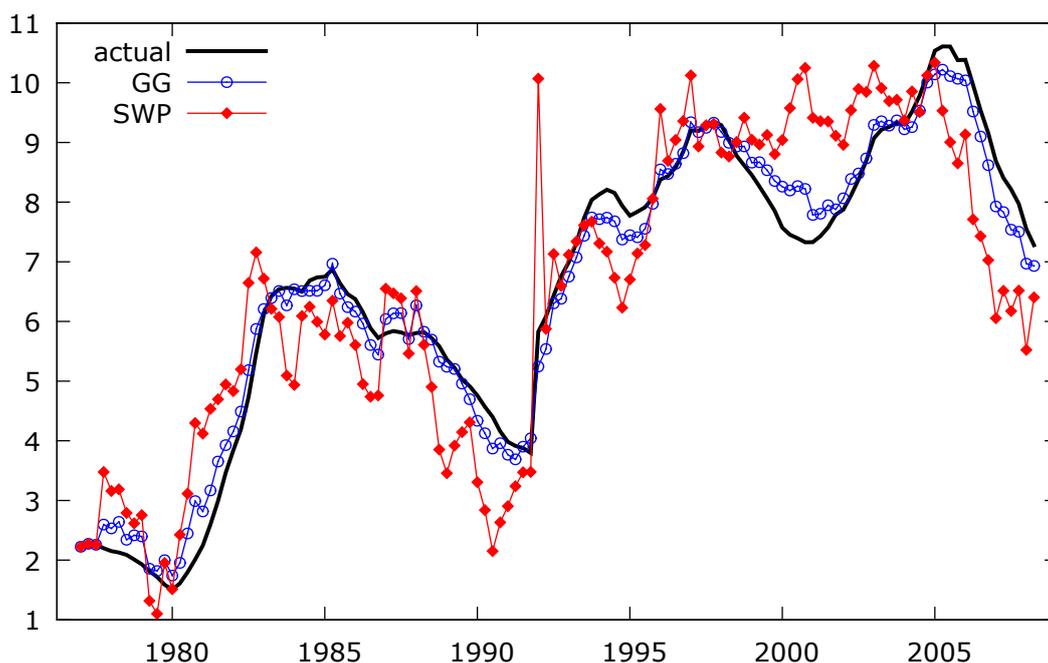


Figure 7: Germany: Gonzalo-Granger (GG) and Stock-Watson-Proietti (SWP) unemployment components compared to actual developments.

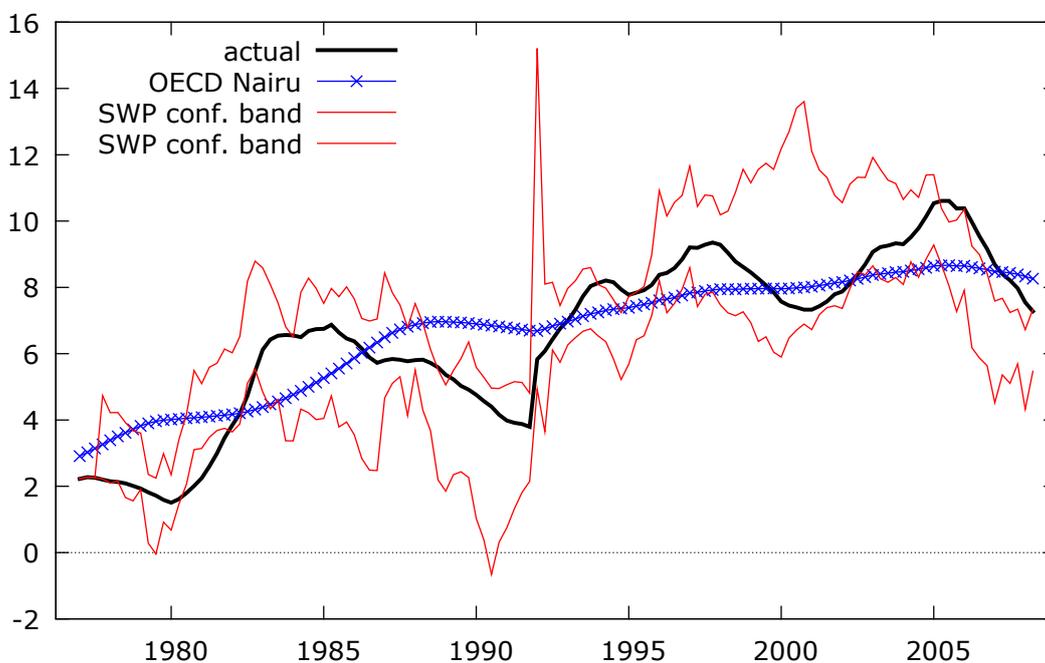


Figure 8: Germany: Stock-Watson-Proietti (SWP) 90% confidence band compared with actual unemployment and the OECD Nairu measure. The confidence band is computed with the delta method.

volatile in the medium run than the actual unemployment rate, whereas the OECD Nairu of course is again a smoothed version of the observed rate.

Italy (figure 11): As in the German case the GG measure tracks the observed unemployment rate closely, perhaps too closely to be useful in practice. However, the SWP measure is quite interesting as it lies consistently above the actual rate during the 1980s when unemployment was rising, and it was consistently below the observed rate during 1998-2005, when unemployment was falling. On the other hand figure 12 shows that these deviations are only borderline significant given the relatively high estimation uncertainty. In contrast, the OECD Nairu is often not compatible with our permanent components, and in general seems to imply excess unemployment for almost all of the Italian sample.

Netherlands (figure 13): Here the GG and SWP permanent components are quite similar from the mid 1990s on, whereas they differ in the first half of the sample. Figure 14 shows again an OECD Nairu estimate which is extremely smooth but which is nevertheless only sometimes significantly different from our estimated SWP long-run unemployment component, especially in the mid 1990s. The largest significant deviations of observed unemployment from the SWP estimate occurred in the mid 1980s, when the unemployment inertia was lagging the equilibrium movements. At the recent end of the sample actual and natural unemployment are very close to each other.

4.2 Restrictions on α and α_{\perp}

Since the common permanent factors driving the system in the Gonzalo-Granger sense are given by $\alpha'_{i\perp} y_{i,t}$, it is worthwhile to analyze the $\alpha_{i\perp}$ in more detail. We do that mainly by testing variables for weak exogeneity, meaning that the corresponding rows of α_i are zero, which in turn implies that some columns of $\alpha_{i\perp}$ are unit vectors. Thus, if a variable is weakly exogenous, it is identified as constituting one component of the $(6 - r_i)$ common permanent factors. As a complement to these tests, we also apply direct tests on $\alpha_{i\perp}$ as described by Gonzalo and Granger (1995) and implemented by the authors as gretl scripts, testing whether a variable can be excluded from the common factors, i.e. whether the corresponding row of $\alpha_{i\perp}$ is zero. The relation between the tests is as follows: If a variable is weakly exogenous, it should not be possible to exclude it from the common factors, and if it can be excluded from the common factor, it should not be weakly exogenous. Of course, statistical uncertainty can blur this logic in reality.

Not surprisingly, from the weak exogeneity tests in table 5 we first learn that inflation and labor costs are not weakly exogenous anywhere. In contrast, productivity seems to be one of the GG factors driving the systems of Germany and France. The external price wedge is also weakly exogenous in those countries, as well as in the Netherlands. Interestingly, the tax wedge is only weakly exogenous in Italy, which would support the

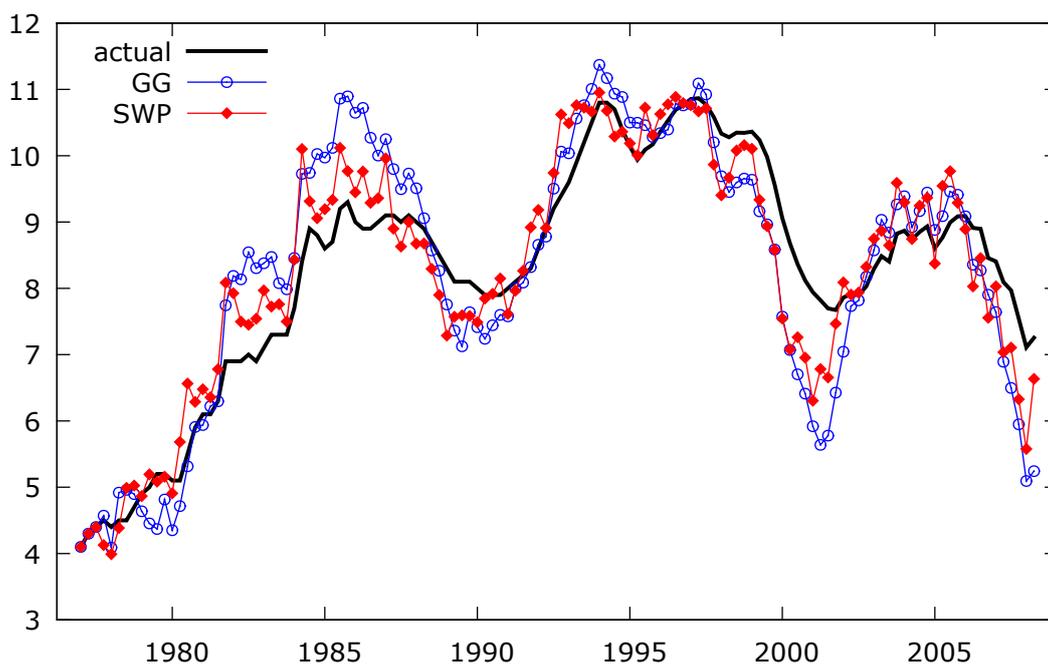


Figure 9: France: Gonzalo-Granger (GG) and Stock-Watson-Proietti (SWP) unemployment components.

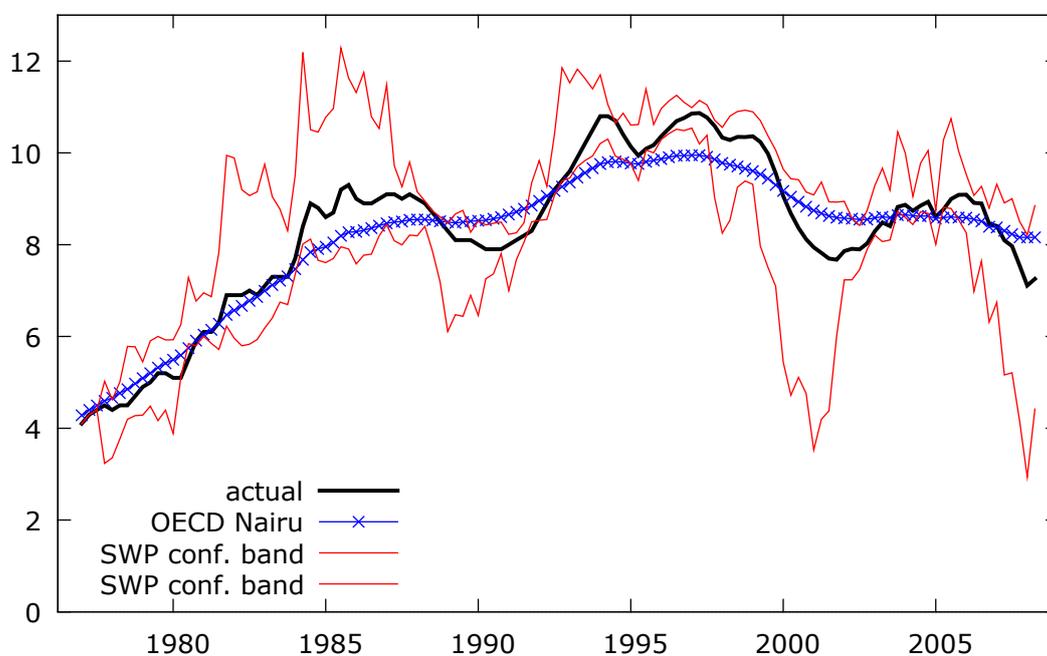


Figure 10: France: Stock-Watson-Proietti (SWP) 90% confidence band compared with actual unemployment and the OECD Nairu measure. The confidence band is computed with the delta method.

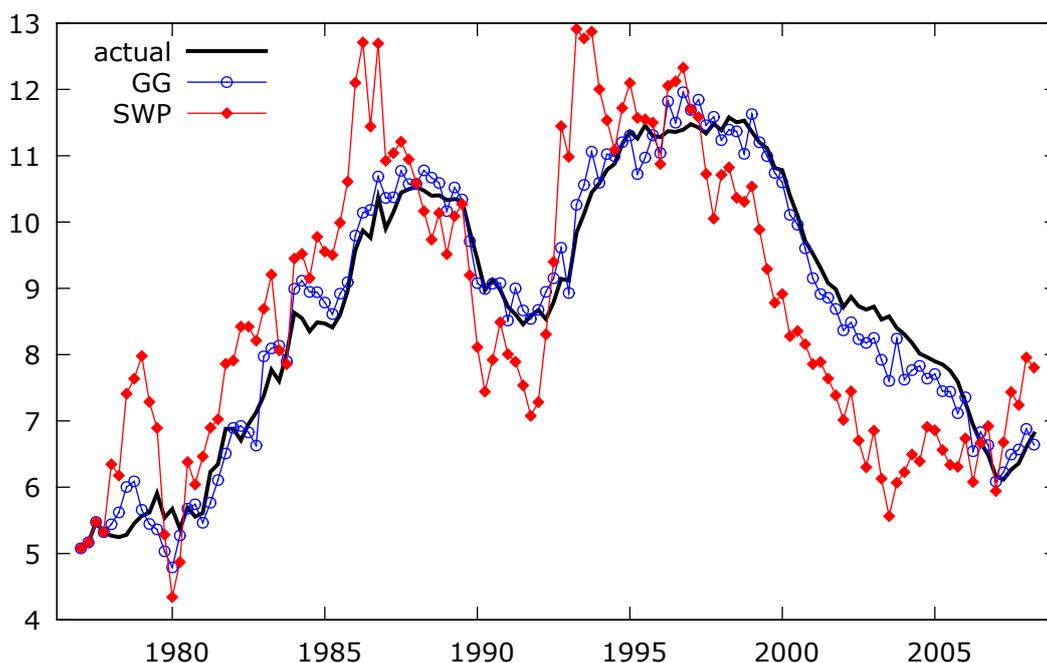


Figure 11: Italy: Gonzalo-Granger (GG) and Stock-Watson-Proietti (SWP) unemployment components compared to actual developments.

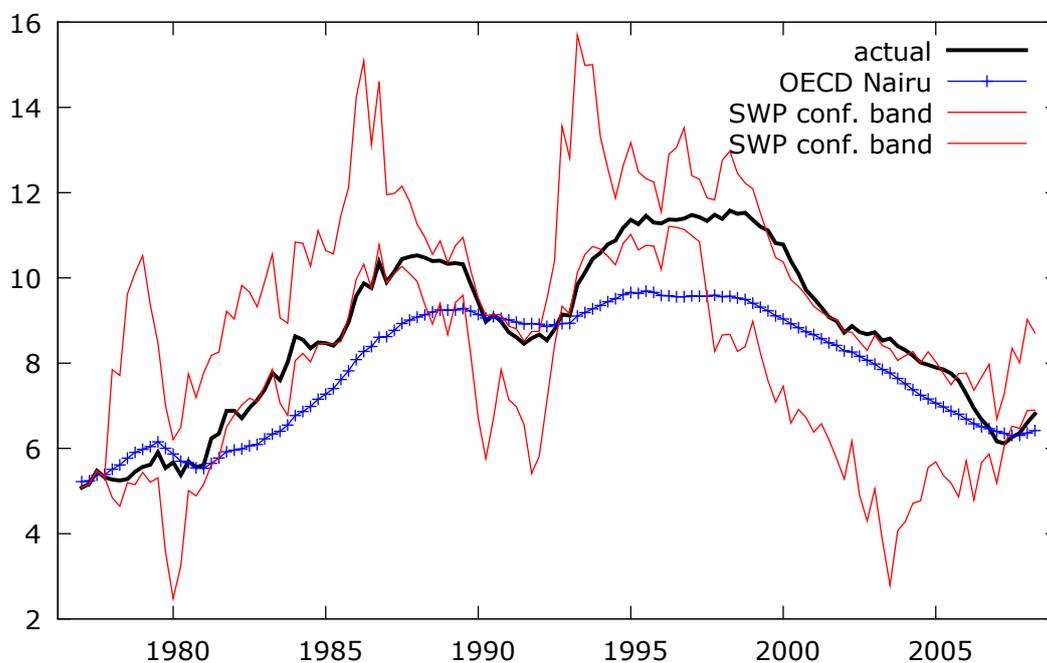


Figure 12: Italy: Stock-Watson-Proietti (SWP) 90% confidence band compared with actual unemployment and the OECD Nairu measure. The confidence band is computed with the delta method.

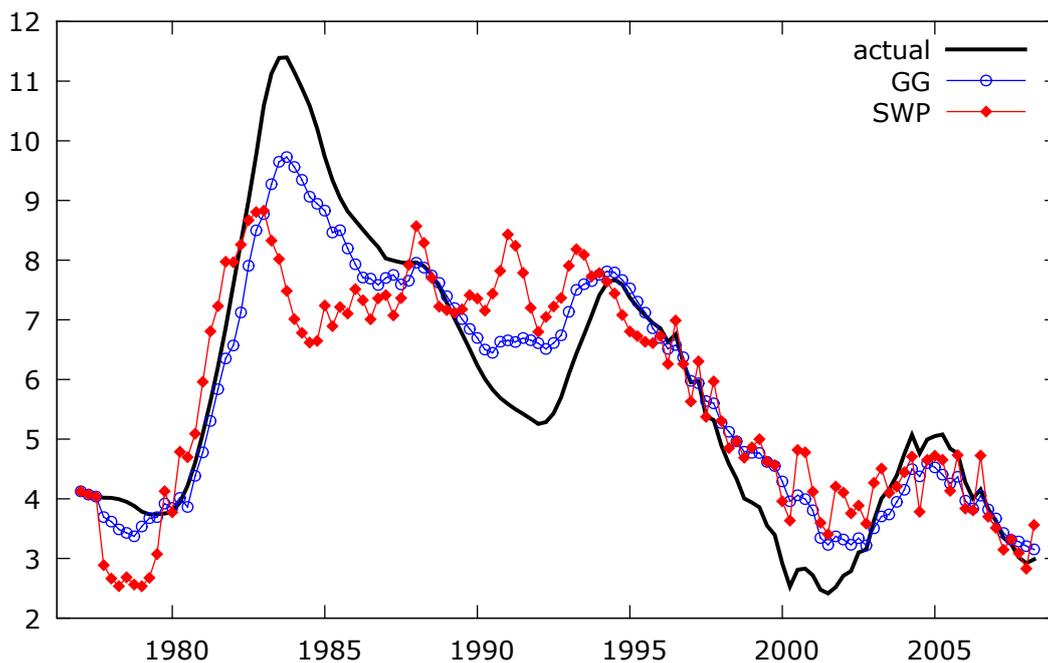


Figure 13: Netherlands: Gonzalo-Granger (GG) and Stock-Watson-Proietti (SWP) unemployment components compared to actual developments.

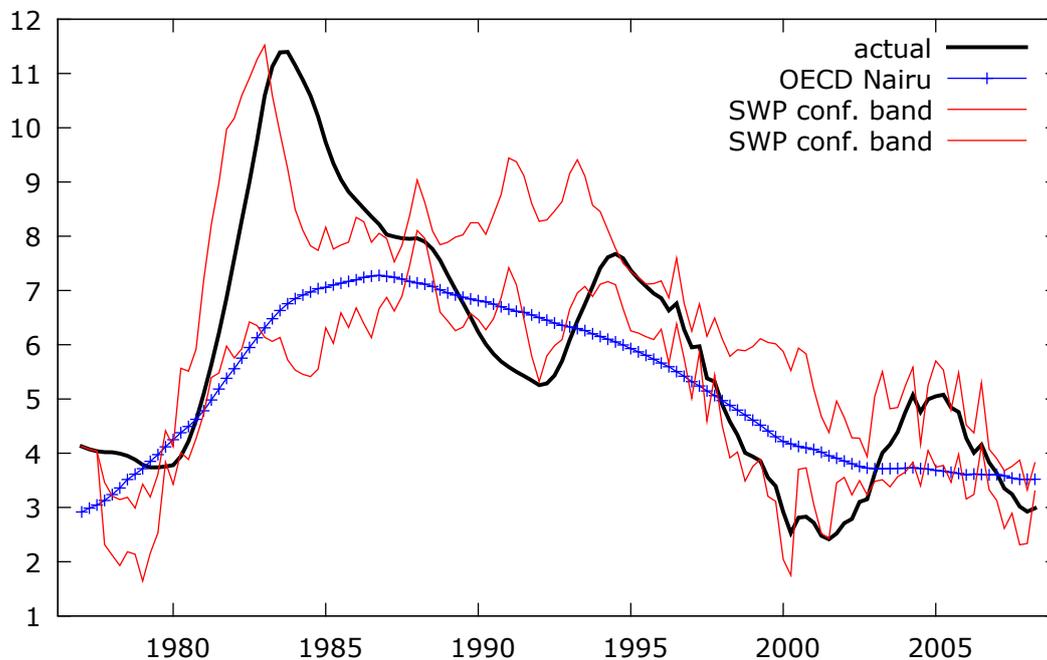


Figure 14: Netherlands: Stock-Watson-Proietti (SWP) 90% confidence band compared with actual unemployment and the OECD Nairu measure. The confidence band is computed with the delta method.

argument that its rise in the other countries was at least partly an endogenous reaction to increased public financing requirements. Nevertheless that rise could of course still have perpetuated the initially rising unemployment rates. In Italy the weak exogeneity test result for the tax wedge could stem from the reform of social security contributions in the late 1990s.

Finally, in Italy the unemployment rate is also weakly exogenous and thus forms one of the GG factors. This feature explains why the GG permanent component for Italy was so close to the actual unemployment rate. Since weak exogeneity means that a variable is not (Granger-) caused by others in the long run, we must acknowledge that it is therefore difficult to relate long-run Italian unemployment developments to the other variables in the model. While the nominal p-values of unemployment in France and the Netherlands do not clearly suggest to reject weak exogeneity, it turns out that after imposing the other clear-cut restrictions on α_{FR} and α_{NL} the unemployment equations have at least one significant loading coefficient and thus unemployment is not weakly exogenous in those countries. For the Netherlands this can also be seen in the lower p-value of the joint test of weak exogeneity of the external price wedge and unemployment.

The results of the inverse testing approach reported in table 6 are compatible with the exogeneity findings.¹⁰ In general, not many variables might be excluded from the GG common factors; those would therefore not contribute to driving the system in the long run: inflation in Germany and the Netherlands and the tax wedge in Germany. The latter result seems to confirm the findings of Bauer and Riphahn (2002) who use different models and methods to analyze the effects of payroll taxes in Germany.

Note that the weak exogeneity restrictions together with the factor exclusion tests mean a partial identification of the four-dimensional GG factors for Germany: Productivity and the external price wedge each form a separate factor, and the remaining two factors must be given by different combinations of unemployment and labor costs. For the other countries the test results yield more complex pictures. But bear in mind that this interpretation relates to the GG factors, not to the permanent components from the common-trends approach. For a direct interpretation of the common-trends based measures (SWP) the permanent shocks would have to be identified along the lines of the structural VAR literature. However, it is well known that especially for systems with more than two or three variables the number of restrictions needed to identify the shocks becomes uncomfortably large, and many restrictions then become incredible. Therefore shock identification is beyond the scope of this paper.

¹⁰Only the p-value of the exclusion test of the German external price wedge $f_{x_{DE,t}}$ is too high compared to conventional significance levels.

Table 5: Weak exogeneity tests

	DE (r=2)	FR (r=2)	IT (r=2)	NL (r=3)
unemployment	.007	.090	.817	.135
inflation	0	0	.003	.001
external price wedge	.584	.817	0	.275
tax wedge	0	.028	.775	.004
labor costs	0	.004	.001	.003
productivity	.438	.288	.033	.002
<i>selected combinations</i>				
prod. & ext. pr. wedge	.477	.385	–	–
unempl. & ext. pr. wedge	–	–	–	.098
unempl. & tax wedge	–	–	.920	–

Notes: Tests whether variables do not adjust to equilibrium deviations, i.e. tests of rows of α filled with zeroes; numbers are p-values. Relevant distribution is $\chi^2(xr)$, where x is the number of tested variables; given numbers are p-values.

Table 6: Exclusion tests from the common factors

	DE (r=2)	FR (r=2)	IT (r=2)	NL (r=3)
unemployment	.025	.002	0	.002
inflation	.320	.005	0	.162
external price wedge	.107	0	.023	0
tax wedge	.366	.003	0	.023
labor costs	.048	.025	.001	.004
productivity	0	.003	0	.031

Notes: Tests whether an individual variable can be excluded from the common permanent components in the Gonzalo-Granger sense, i.e. restriction tests on α_{\perp} . Null hypothesis: The respective variable is *not* part of the permanent factors driving the system. Relevant distribution is $\chi^2(K)$, numbers are p-values.

Table 7: Test for long-run wage-productivity homogeneity

Germany	France	Italy	Netherlands
0.051	0.016	0	0.086

Notes: Likelihood ratio test results of the restriction of a 1:1 relation between real hourly labor costs and real hourly labor productivity in all equilibrium relations. I.e., under the null hypothesis each coefficient of labor costs would have to be equal to the negative coefficient of labor productivity in each co-integration vector. If the null hypothesis were true, the long-run relationships could be expressed by substituting the labor share for labor costs and productivity. The chosen co-integration ranks are the same as in the rest of the analysis.

4.3 Identification of the co-integration relations

Since the permanent-transitory decompositions based on co-integrated systems always take into account the full co-integration space, it is not necessary to identify the individual co-integration relations. This is why we could present the permanent unemployment components without having discussed identification. However, for a fuller understanding of what connects the variables within the four countries it is useful to analyze the estimated co-integration spaces in more detail. To start with the interpretation of the co-integration space we first test the restriction of unit proportionality between productivity and labor costs as mentioned earlier. Table 7 shows that this is not a safe restriction to impose, and using labor shares in all co-integration relations is therefore not possible.

In table 8 we therefore report statistically acceptable and economically plausible estimates of the co-integrating relations. We used over-identifying restrictions and the weak exogeneity restrictions from the previous subsection were also imposed. It turns out that due to apparent cross-country differences we cannot apply a single identification scheme for all systems, but variants of Phillips curves apply to all countries.

In Germany, we find a Phillips curve which primarily links inflation and unemployment, but where labor costs relative to the external price wedge would also affect inflation. The second relation can be interpreted to say that a higher tax wedge will go hand in hand with unemployment, and that higher labor costs relative to (marked-down) productivity would tend to raise unemployment.

For France, the first relation again is a Phillips curve variant, albeit one where both the unemployment rate and a quasi-labor share appear (with plausible signs). In the second relation unemployment is again linked to the tax wedge, but in contrast to the German case productivity is now compared to the external price wedge instead of labor costs.

The Italian model is special in so far as the tax wedge does not enter any cointegration relationship. Combined with the finding that it is also weakly exogenous this means that

Table 8: Identified co-integration relations

	DE		FR		
unemployment	0.96 (0.062)	1	0.54 (0.129)	1	
inflation	1	0	1	0	
ext. pr. wedge	0.125 (0.018)	0	0	-0.15 (0.037)	
tax wedge	0	-0.68 (0.072)	0	-0.51 (0.073)	
labor costs	-0.131 (0.030)	-1.79 (0.237)	-0.73 (0.078)	0	
productivity	0	1.36 (0.176)	0.60 (0.057)	0.23 (0.054)	
unification shift	-2.44 (0.859)	7.98 (2.485)	-	-	
weakly exog. over-id. test	ext. pr. wedge, prod. $\chi^2_6 = 6.92, p = 0.328$		ext. pr. wedge, prod. $\chi^2_6 = 6.86, p = 0.334$		
	IT		NL		
unemployment	0	1	0.78 (0.132)	1.59 (0.390)	-4.98 (0.708)
inflation	1	0	1	0	0
ext. pr. wedge	0	0.71 (0.125)	0.05 (0.015)	0.35 (0.045)	0
tax wedge	0	0	0	0	1
labor costs	-0.36 (0.046)	-1.68 (0.266)	0	1	3.90 (0.816)
productivity	0.36	0	0	-1	-2.55 (0.523)
weakly exog. over-id. test	unempl., tax wedge $\chi^2_9 = 6.83, p = 0.655$		ext. pr. wedge $\chi^2_5 = 4.45, p = 0.487$		

Notes: Standard errors below point estimates. Coefficients without standard errors are restricted. The over-identification tests refer to all restrictions on α and β .

the tax wedge is quite detached from the other variables in the long run. Also somewhat surprisingly, the Italian Phillips curve variant does not contain the unemployment rate; instead it is only the labor share which affects inflation. As a mirror image of the French second co-integration relation we now have labor costs relative to the external price wedge which are related to unemployment (of course with the opposite sign).

Finally turning to the Netherlands, the Phillips curve with unemployment again surfaces, as in Germany and France. At least statistically the external price wedge seems to be needed there, too, but with a very small coefficient. Next, in the second relation we find a quite clear wage curve: labor costs are homogeneous with respect to productivity, and are negatively affected by high unemployment as well as low competitiveness (low relative import prices). The last relation is denominated in terms of the tax wedge and explains it with unemployment and labor costs corrected for (marked-down) productivity.

4.4 Evaluation

Now we turn to a quantitative evaluation of the various measures of natural unemployment. Since our approach is based on reduced-form systems, we choose three types of predictive power as the metric with which to assess the quality of the estimates. All approaches share the fact that we use the unemployment gaps $g_{i,t}^j = u_{i,t} - u_{i,t}^j$ as the relevant predictors, where $j \in \{GG, SWP\}$ indicates which estimate is used. For comparison we also use the unemployment gap from the OECD Nairu estimates. In principle it would be possible to use the predictive power of the gaps as an objective function to be maximized. However, that would lead us directly to a state-space model with unobserved components, to be estimated by the Kalman filter. While such an approach also has its merits, we view our co-integration-based approach as a complement. The advantages of our approach are that it is transparent and that it avoids typical difficulties of complex multivariate state-space models like convergence issues of the numerical algorithms. Therefore we use the predictive power of the estimated gaps as an ex-post criterion, not as an ex-ante objective function.

The first variant starts from the premise that natural unemployment should be an attractor for observed unemployment. This means that unemployment changes should react to realizations of the unemployment gap in the following manner:

$$\Delta u_{i,t} = c_{i,1} - b_{i,1} g_{i,t-1}^j, \quad b_{i,1} > 0, \quad (9)$$

where the intercept may balance a non-zero mean of the gap over the current sample, for example. The strength of this bi-variate relation can be assessed by simply calculating the correlation coefficient, which should be negative.

The second and third variants rely on a simple Phillips-curve concept, which has always posited that a higher tightness of the labor market should tend to limit price increases.¹¹ Here we differentiate between two measures of cyclical inflation, namely using a simple first-difference filter, and the implicitly estimated inflation gap from our models. Formally this is described in the two following equations:

$$\Delta\pi_{i,t} = c_{i,2} - b_{i,2}g_{i,t-1}^j, \quad b_{i,2} > 0 \quad (10)$$

$$\pi_{i,t} - \pi_{i,t}^j = c_{i,3} - b_{i,3}g_{i,t-1}^j, \quad (11)$$

In variant (10) the lagged inflation rate included in the difference of inflation serves to balance the non-stationarity of inflation, but it is of course a predetermined variable at time t . Hence we expect a negative sign for $b_{i,2}$, reflecting the dampening effect of excess unemployment on inflation. With respect to the variant (11), however, where the contemporaneous permanent component appears along with actual inflation, things are more complicated. In a world of non-stationary inflation and unemployment it could be that past excess unemployment also affects the permanent component of inflation, and thus it is not clear a priori whether $b_{i,3}$ should be positive or negative.

The results of the three exercises are displayed in table 9. The most important insights are the following: First, in terms of predicting changes of unemployment and the level of the inflation gap (first and third panel) our measures are quite successful. Only for the Netherlands the relevant correlations are quite low (in absolute value). It is interesting that for the prediction of unemployment changes the SWP measures are most useful, whereas the inflation gap forecasts seem to work best with the GG measures.

In contrast, the predictive power of our implicit unemployment gap measures for inflation changes (middle panel) is negligible in all countries except for Italy. Finally, the OECD unemployment gaps compare surprisingly badly, showing mostly very little predictive power; the only cases where the correlations exceed 0.2 (in absolute value) are the forecasts of the GG and SWP inflation gaps in Germany and France. Even there, however, they perform much worse than our GG unemployment gaps.

5 Conclusions

In this paper we have presented an operational and directly reproducible approach to tackle the old problem of estimating the natural unemployment rate in the biggest euro-area countries, namely Germany, France, and Italy. The Netherlands were also analyzed

¹¹See for example Claar (2006) for an assessment of equilibrium unemployment measures based on inflation forecasts.

Table 9: Evaluation: predictive power of unemployment gaps

correlation between...			DE	FR	IT	NL
Δu_t	and GG gap	$(u_{t-1} - u_{t-1}^{GG})$	-0.277	-0.341	-0.291	0.103
Δu_t	SWP gap	$(u_{t-1} - u_{t-1}^{SWP})$	-0.434	-0.599	-0.463	-0.216
Δu_t	OECD Nairu gap	$(u_{t-1} - u_{t-1}^{OECD})$	-0.056	-0.107	-0.077	0.052
$\Delta \pi_t$	GG gap		-0.125	0.044	-0.425	-0.028
$\Delta \pi_t$	SWP gap		-0.151	0.062	-0.124	-0.002
$\Delta \pi_t$	OECD Nairu gap		-0.005	-0.042	-0.032	-0.025
GG inflation gap	GG gap	$(\pi_t - \pi_t^{GG})$	0.584	-0.913	0.508	-0.095
SWP inflation gap	SWP gap	$(\pi_t - \pi_t^{SWP})$	-0.178	-0.683	0.167	-0.292
GG inflation gap	OECD Nairu gap		0.323	0.474	0.125	0.006
SWP inflation gap	OECD Nairu gap		0.261	0.257	-0.016	-0.163

Notes: All panels show correlations with respect to (the lag of) the implicit unemployment gap, see equations (9), (10), and (11). **Boldface** indicates the strongest correlation per country and panel (in the upper two panels only negative correlations are considered).

because they share the feature that unemployment rates appear non-stationary over the last three decades. Our approach relies on the proper specification of co-integrated VARs and then extracts the corresponding long-run component of unemployment by means of permanent-transitory decompositions. While this method is mostly non-structural, it is nevertheless influenced by theory. The established literature on the wage-price mechanism, wage-bargaining theory, and unemployment determination dictated the choice of the relevant variables in the system. Also, although not crucial for the estimates of the permanent components, the estimated co-integration relations could be identified and structurally interpreted, for example as variants of Phillips curves.

In a sample spanning the period between the oil price shocks of the 1970s and the recent financial crisis we found the same number of co-integration relationships –namely two– for the big three euro area members Germany, France, and Italy. In the Netherlands even a third equilibrium relationship could be established.

Among the many possible permanent-transitory decompositions we picked the two most natural candidates: namely the Gonzalo-Granger decomposition, where the permanent components are given as linear combinations of the contemporaneous observables, and the Stock-Watson decomposition, where the long-run components describe the equilibrium location of the system which would be attained if no further shocks occurred. For the latter, we used the formulation by Proietti (1997) which expresses the components as distributed lags of the variables, instead of directly using the estimated residuals.

In terms of the permanent unemployment components, an expected but nonetheless important finding is that they are very different from the OECD Nairus, because the latter are extremely smooth by construction. However, given the standard assumption of constantly arriving permanent shocks we believe that it is quite sensible to model a relatively volatile equilibrium path instead, together with sluggish *adjustment towards* equilibrium. Especially the Stock-Watson-Proietti components track the long-run behavior of the unemployment rates well without being virtually identical to observed actual rates. In that sense they appear to be plausible measures of structural unemployment levels.

In general all included variables seemed to play a role for the long-run development of the systems. Non-causality (weak exogeneity) tests indicate that the external price wedge is a separate long-run driving force in all countries except Italy (where it nonetheless still belongs to the long-run factors in the Gonzalo-Granger sense). In Germany and France a second long-run driving factor is given by labor productivity. It is perhaps noteworthy that the tax wedge does not enter the Gonzalo-Granger long-run factors in Germany.

We evaluated all measures based on the predictive power of the implied unemployment gaps. They are generally useful predictors for future unemployment developments and for future inflation gaps estimated within the system, with the exception of the Nether-

lands. However, the unemployment gaps fail as forecasters of first differences of inflation. For comparison, the OECD Nairus do not perform well in any of the evaluations, although these assessments arguably may not be well suited for such extremely smooth series.

All in all, we do not claim that our approach solves all inherent problems of the estimation of non-stationary natural unemployment rates; for example, in the Italian model unemployment turned out as not being significantly (Granger-) caused by other variables in the long-run. But we conclude that it is a useful tool, and it is ready to be routinely applied as a complement to the other currently widespread methods.

References

- Apel, M. and Jansson, P. (1999). System estimates of potential output and the NAIRU. *Empirical Economics*, 24:373–388.
- Arestis, P. and Mariscal, I.-F. (2000). OECD unemployment: structural breaks and stationarity. *Applied Economics*, 32(4):399–403.
- Bårdsen, G., Eitrheim, Ø., Jansen, E., and Nymoen, R. (2005). *The Econometrics of Macroeconomic Modelling*. Oxford University Press.
- Bårdsen, G. and Nymoen, R. (2003). Testing steady-state implications for the NAIRU. *Review of Economics and Statistics*, 85(4):1070–1075.
- Basistha, A. and Startz, R. (2008). Measuring the NAIRU with reduced uncertainty: A multiple-indicator common-cycle approach. *The Review of Economics and Statistics*, 90(4):805–811.
- Batini, N., Jackson, B., and Nickell, S. (2005). An open-economy new Keynesian Phillips curve for the U.K. *Journal of Monetary Economics*, 52(6):1061–1071.
- Bauer, T. and Riphahn, R. T. (2002). Employment effects of payroll taxes – an empirical test for Germany. *Applied Economics*, 34(7):865–876.
- Blanchard, O. and Katz, L. F. (1999). Wage dynamics: Reconciling theory and evidence. *American Economic Review Papers&Proceedings*, 89(2):69–74.
- Blanchflower, D. G. and Oswald, A. J. (1995). An introduction to the wage curve. *Journal of Economic Perspectives*, 9(3):153–167.
- Buseti, F. (2002). Testing for (common) stochastic trends in the presence of structural breaks. *Journal of Forecasting*, 21(2):81–105.
- Claar, V. (2006). Is the NAIRU more useful in forecasting inflation than the natural rate of unemployment? *Applied Economics*, 38(18):2179–2189.
- Cottrell, A. and Lucchetti, R. (2009). *Gretl User's Guide*. Version 1.8.1; available at <http://gretl.sourceforge.net/#man>.

- Fabiani, S. and Mestre, R. (2004). A system approach for measuring the euro area NAIRU. *Empirical Economics*, 29:311–341.
- Garratt, A., Robertson, D., and Wright, S. (2006). Permanent vs transitory components and economic fundamentals. *Journal of Applied Econometrics*, 21(4):521–542.
- Gonzalo, J. and Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business and Economics Statistics*, 13(1):27–35.
- Gordon, R. J. (1997). The time-varying NAIRU and its implications for economic policy. *Journal of Economic Perspectives*, 11:11–32.
- Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.
- King, R. G., Plosser, C. I., Stock, J. H., and Watson, M. W. (1991). Stochastic trends and economic fluctuations. *American Economic Review*, 81(4):819–840.
- Krätzig, M. (2004). The software JMulTi. In Lütkepohl, H. and Krätzig, M., editors, *Applied Time Series Econometrics*, chapter 8. Cambridge University Press.
- Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root, how sure are we that economic time series have a unit root? *Journal of Econometrics*, 54:159–178.
- Lanne, M., Lütkepohl, H., and Saikkonen, P. (2003). Test procedures for unit roots in time series with level shifts at unknown time. *Oxford Bulletin of Economics and Statistics*, 65(1):91–115.
- Laubach, T. (2001). Measuring the NAIRU: Evidence from seven economies. *The Review of Economics and Statistics*, 83(2):218–231.
- Layard, R., Nickell, S., and Jackman, R. (1991). *Unemployment – Macroeconomic Performance and the Labour Market*. Oxford University Press.
- Nickell, S. J. (1998). Unemployment: Questions and some answers. *Economic Journal*, 108(448):802–816.
- Nyblom, J. and Harvey, A. (2000). Tests of common stochastic trends. *Econometric Theory*, 16(2):176–99.
- Papell, D. H., Murray, C. J., and Ghiblawi, H. (2000). The structure of unemployment. *Review of Economics and Statistics*, 82(2):309–315.
- Pissarides, C. A. (1998). The impact of employment tax cuts on unemployment and wages; the role of unemployment benefits and tax structure. *European Economic Review*, 42:155–183.
- Proietti, T. (1997). Short-run dynamics in cointegrated systems. *Oxford Bulletin of Economics and Statistics*, 59(3):405–422.

- Proietti, T., Musso, A., and Westermann, T. (2007). Estimating potential output and the output gap for the euro area: a model-based production function approach. *Empirical Economics*, 33(1):85–113.
- Richardson, P., Boone, L., Giorno, C., Meacci, M., Rae, D., and Turner, D. (2000). The concept, policy use and measurement of structural unemployment: Estimating a time varying NAIRU across 21 OECD countries. Technical Report 250, OECD, Paris.
- Saikkonen, P. and Lütkepohl, H. (2000). Testing for the cointegrating rank of a VAR process with structural shifts. *Journal of Business and Economic Statistics*, 18(4):451–464.
- Stock, J. H. and Watson, M. W. (1988). Testing for common trends. *Journal of the American Statistical Association*, 83:1097–1107.

A Unit root tests

Table 10 reports the results of augmented Dickey-Fuller (ADF) tests on all included variables. At a 5% level of significance, if all variables are truly $I(1)$ we would expect about one test rejection out of 20 cases. Here we test 24 variables and the unit root is nominally rejected for Dutch labor costs and German unemployment. But note that in the case of Dutch wages the alternative hypothesis of the unit root test is given by trend stationarity. However, a trend term is not included in the co-integration space of our statistical models in this paper, and thus it is impossible that Dutch wages could form an additional co-integration relationship trivially by itself.

For German unemployment the situation is of course more complicated, since no trend term is allowed there and because the existence of an (approximate) unit root in the unemployment series is a necessary condition for the permanent-transitory decompositions of the present study. Therefore we are obliged to investigate this issue further. We do that by applying a stationarity test where the *alternative* instead of the null hypothesis is given by the unit root; the multivariate generalization of the same test was used to assess the number of common trends in section 3.2. The results in table 11 show that the null of stationarity is now clearly rejected for German unemployment for all values of the lag truncation parameter. In contrast, for Dutch labor costs the results are ambiguous precisely because they depend on the choice of the lag truncation.

Our interpretation of the conflicting test results for German unemployment tends to view the nominal rejection of the ADF-type test as possibly spurious, because it is known that some initial values of a sample can bias the test towards rejection. Also the results of the tests in favor of $I(1)$ are even more significant than the tests in favor of $I(0)$. Therefore we continue to work under the assumption that German unemployment has a unit root. As mentioned in the introduction, this corresponds to the results that were found by Papell et al. (2000) who even allowed for more shifts under the alternative hypothesis. Arestis and Mariscal (2000) also found unit roots in unemployment for the countries of the present paper, given that we disregard their trend stationary result for Italy, because we rule out linear trends in unemployment rates, and that we require a unification shift in specifications for unified German data.

With respect to Dutch labor costs we choose to effectively treat the unit root test result as a type I error and continue to assume that Dutch labor costs are $I(1)$. Of course, the corresponding Dutch results should be interpreted with the necessary caution.

An additional concern is that inflation could be neither $I(0)$ or $I(1)$, but could be fractionally integrated with an order of integration between zero and unity. We check whether there is significant evidence of this possibility by applying tests with the alternative hypothesis of fractional integration to the inflation series of all countries, see table 12. These

Table 10: Unit root tests

	Germany	France	Italy	Netherlands
unemployment	-3.102**	-0.130	-0.633	-1.288
inflation	-2.357	-1.464	-0.830	-0.632
external price wedge	-2.741	-0.783	-2.166	-0.197
tax wedge	-0.742	0.714	0.684	-1.051
labor costs	-1.251	-0.417	-0.373	-3.032**
productivity	-1.771	-0.053	-0.134	-1.018

Notes: Augmented Dickey-Fuller-style unit root tests. Significance denoted by * (10%), ** (5%), *** (1%). For the German series the tests include a unification shift, see Lanne et al. (2003), as implemented in JMulTi; critical values for this specific test are -2.58/-2.88/-3.48 (without trend) and -2.76/-3.03/-3.55 (with trend). The tests for external relative prices, labor costs, and productivity include a trend term. All tests use a GLS de-trending/de-meaning procedure for the deterministic terms under the null. For the tests without break insignificant lagged differences were automatically removed; maximum number of lagged differences is 5 (or 8 where necessary).

Table 11: Additional stationarity tests

lag truncation	German unemployment	Dutch labor costs
1	0.734***	0.222***
2	0.499***	0.155**
3	0.382***	0.121*
4	0.312***	0.101
5	0.266***	0.088

Notes: Test of stationarity of the series against the alternative hypothesis of a unit root. The test for German unemployment is a Busetti (2002) test with a level shift for unification. The test for Dutch labor costs is a standard KPSS test (Kwiatkowski et al., 1992) with trend. Significance denoted by * (10%), ** (5%), *** (1%).

Table 12: Tests of fractional integration in inflation

Germany	France	Italy	Netherlands
-0.618 (p=0.546)	-1.225 (p=0.240)	-1.608 (p=0.129)	-1.041 (p=0.315)

Notes: Geweke and Porter-Hudak (GPH) test of $H_0 : I(0)$ versus $H_1 : I(d)$, $d \neq 0$ applied to the difference of the inflation series, such that effectively the hypothesis pair with respect to inflation is $I(1)$ vs. $I(d)$, $d \neq 1$. The German differenced inflation series were regressed on impulse dummies first to remove the potential unification break.

tests do not reject the null hypothesis that inflation is $I(1)$, either.

B Data details

All data are taken from the OECD Economic Outlook database, except where noted otherwise.

The quarterly WSSS (gross wages) series is missing for the Netherlands, such that we use an equivalent series (coe) from the Quarterly National Accounts database. Because of the availability of the unemployment series for unified Germany all data for Germany are combined from Western Germany until 1991q4, and unified Germany starting 1992q1, although political unification already took place in 1990q4.

Only total direct taxes on all types of household income are available in ECO. Therefore we use TKTRG (=capital tax and transfers received by government) and subtract that from TY (=total direct taxes), as a first approximation. The tax and social security data (TY, SSRG, TKRTKGSFRDS) are only annually available such that we use a centered moving average for interpolation to quarterly data.

Average hours per employee in OECD ECO refer to the business (private) sector. But we still take a total-economy perspective and use data for the entire economy, e.g., because unemployment cannot be distinguished between private and public sectors. The implicit assumption here is that working hours in private and public sectors are log-equal, i.e., the ratio is constant.

For non-commodity import prices (PMGSX, and PMNW) no (West) German data before 1991 is available in ECO. Instead we use (historical, and re-based) import prices excluding oil and oil products until 1990 from the German federal statistical office.