CAHIER D'ETUDES
WORKING PAPER

N° 4

POTENTIAL OUTPUT AND THE OUTPUT GAP IN LUXEMBOURG:

SOME ALTERNATIVE METHODS

by Paolo GUARDA
June 2002

BANQUE CENTRALE DU LUXEMBOURG

CAHIER D'ETUDES
WORKING PAPER

N° 4

POTENTIAL OUTPUT AND THE OUTPUT GAP IN LUXEMBOURG:

SOME ALTERNATIVE METHODS

by Paolo GUARDA
June 2002

BANQUE CENTRALE DU LUXEMBOURG

Address :	2, Boulevard Royal - L-2983 Luxembourg
Telephone :	(+352) 4774 - 1
Fax :	(+352) 4774 - 4901
Internet :	http://www.bcl.lu
E-mail :	sg@bcl.lu
Télex :	2766 IML LU

Potential output and the output gap in Luxembourg : some alternative methods*.

by Paolo GUARDA**

Résumé

L'écart de production est la différence entre le niveau de production observé dans l'économie et son niveau potentiel. A court terme, la production peut dépasser son niveau potentiel (un écart de production positif) uniquement en présence de niveaux anormaux de participation sur le marché de l'emploi, d'utilisation de capacités, et/ou de progrès technologique. Cependant, un écart de production positif a tendance à générer des pressions inflationnistes sur les marchés de facteurs de production. Une fois que l'inflation s'accélère, la production doit baisser en dessous de son niveau potentiel (un écart de production négatif) afin d'augmenter les ressources disponibles et réduire la pression sur les prix. L'analyse macro-économique se sert souvent de l'écart de production pour évaluer les pressions inflationnistes actuelles et futures. Cette étude décrit plusieurs méthodes alternatives pour l'estimation de l'écart de production. Six de ces méthodes ont été appliquées à des données annuelles pour le Luxembourg. Par la suite, ces différentes mesures de l'écart de production ont été comparées et évaluées en termes de leur contribution aux prévisions d'inflation. Les méthodes basées sur les modèles à composantes inobservées sont généralement préférables à des méthodes plus simples et plus diffusées (à savoir les tendances linéaires ou le filtre de Hodrick-Prescott). Les méthodes multivariées qui tiennent compte de l'évolution simultanée de plusieurs variables économiques sont généralement privilégiées aux méthodes univariées qui se limitent à l'évolution passée de la production.

Abstract

The output gap is defined as the difference between the observed level of an economy's output and its trend or potential level. In the short term, an economy can produce above its potential level (a positive output gap) through unusually high levels of labour force participation, capacity utilisation, or technical progress. However, a positive output gap tends to generate inflationary pressures on the markets for factors of production. Once inflation accelerates, output will have to fall below its potential level (a negative output gap) to increase available resources and reduce the pressure on prices. Therefore, measures of the output gap are often used in macroeconomic analysis to assess current and future levels of inflationary pressures in the economy. This study reviews several of the many alternative methods of estimating output gaps and applies six of these to annual data for Luxembourg. These different measures of the output gap are then compared and evaluated in terms of their contribution to inflation forecasting. Methods based on unobserved components models tend to do better than simpler, better known methods (i.e. linear trends, the Hodrick-Prescott filter). Multivariate methods that consider the simultaneous evolution of several different economic variables tend to do better than univariate methods that limit themselves to the output series itself.

---

* The opinions expressed in this paper do not necessarily reflect the views of the Banque centrale du Luxembourg. The author is grateful for technical assistance provided by Luisito Bertinelli of CORE at the Université Catholique de Louvain, Simon Peters of Manchester University, and Christophe Planas of the EU Commission Joint Research Centre in Ispra. Helpful comments were provided by Francisco Nadal De Simone of the IMF and several colleagues at the BCL.

** Monetary, Economic and Statistics Department
E-mail: mes@bcl.lu

Table of contents

List of tables

List of figures

Potential output and the output gap in Luxembourg:
some alternative methods.

Introduction

The study of macroeconomic fluctuations involves distinguishing short-term cyclical fluctuations in output and underlying long-term developments. The output gap is defined as the difference between the observed level of an economy's output and its trend or potential level. In the short term, an economy can produce above its potential level (a positive output gap) through unusually high levels of labour force participation, capacity utilisation or technical progress. However, a positive output gap tends to generate inflationary pressure on the markets for factors of production. Once inflation accelerates, output will have to fall below its potential level (a negative output gap) to increase available resources and reduce the pressure on prices. Therefore, measures of the output gap are often used in macroeconomic analysis to assess the current and future levels of inflationary pressures in the economy. The output gap can also be used as a diagnostic device to evaluate whether output and inflation forecasts are consistent. IMF, OECD, or European Commission forecasts of Luxembourg's GDP and inflation all imply a given path of potential output growth. Estimates of potential output growth (past and future) can also be used to assess the realism of the growth rate assumptions required to finance the government budget or existing commitments on public pensions. Finally, output gap estimates can contribute to the formulation of monetary policy by giving an indication of where the economy is situated in the business cycle and the outlook for inflation. The following section reviews a variety of alternative methods for calculating potential output or output gaps. Section 3 applies several of these methods to annual data for Luxembourg. Section 4 provides some formal comparisons and assessments of the different output gap measures obtained. The final section raises some issues concerning the uncertainty attached to any measure of the output gap and offers some conclusions.

1. A Survey of Existing Potential Output Measures

There is no standard way to measure potential output. Different methods usually produce similar *growth rates* of potential output, but the estimated *level* of potential output is often quite different depending on the method chosen or even the specific assumptions made within a given method. This is alarming, as the policy consequences will differ according to the estimated sign and size of the output gap. Since these can differ across methods, this means that alternative methods can sometimes give contradictory policy advice.

Broadly speaking, there are two main approaches to measuring potential output[1]. The first includes *univariate* methods that identify the output gap solely from the past behaviour of the output series without referring to any other macroeconomic variables. These methods are essentially statistical as they are only based on some (explicit or implicit) assumption about the dynamics of the output series. Univariate methods really focus on isolating a *trend* measure of output that is assumed to be close to the potential level. The second approach to measuring potential output includes *multivariate* methods that also consider the evolution of other macroeconomic variables. This enables such methods to exploit relationships derived from economic theory (such as the Phillips curve) to obtain a measure of potential output that is closer to the notion of the sustainable aggregate supply capabilities of an economy.

1.1 Univariate Methods

Univariate methods decompose observed output into a trend component with a smooth evolution through time and a cyclical component whose evolution is more volatile. Changes in the trend component are generally interpreted as the result of changes that have long-term effects on the level of supply. For example, the trend component may increase because of an expanded labour supply reflecting population growth, immigration or greater participation in the workforce. Alternatively, the trend component may increase because of capital accumulation or technical progress. Changes in the cyclical component are usually interpreted as the result of changes that have only short-term effects on the level of supply. For example, the cyclical component may reflect temporary disequilibria in the markets for goods, labour, or capital due to price stickiness or costs of adjustment.

However, univariate methods are not based on any particular interpretation of the trend and cyclical components they estimate. They are based only on assumptions specifying how the different components evolve through time. Different univariate methods are distinguished by the assumptions they make regarding these dynamics. Three types of univariate methods are discussed below: linear trends, univariate filters, and univariate unobserved components models[2].

1 Since the output gap is the difference between observed output and potential output, measuring potential output and measuring the output gap are used interchangeably in the text that follows.

2 Hybrid methods implementing univariate filters or unobserved components models in a multivariate context are discussed in the next section.

1.1.1 Linear Time Trends

Probably the simplest method of isolating trend output is to assume that potential output grows at a constant exponential rate (i.e. a log-linear time trend). In this case, it is sufficient to regress the log of output on a linear function of time.

$$y_t = \mu + \beta t + \varepsilon_t$$

where y_t is the natural logarithm of the observed level of output in period *t*. The estimated coefficient β provides a measure of the hypothesised constant rate of growth of potential output. Thus regression results can be used to decompose output into a deterministic linear trend represented by $(\mu + \beta t)$ and a zero-mean cyclical component ε_t represented by the estimated residuals. Of course, for ordinary least squares to provide efficient estimates, the usual assumptions must be satisfied. That is to say that the errors must be independently and identically distributed with mean zero and constant variance.

As pioneered by Okun in the early 1960's, this method worked fairly well during the post-war period that was characterised by relatively constant growth in output[3]. In fact, linear time trends had become the standard method of estimating potential output when the 1970's brought in major supply shocks. These induced serial correlation and non-normality in the regression residuals, leading to parameter instability and unrealistic results. Initial efforts to salvage the time trend method relied on dummy variables to account for the major supply shocks. The 1970's slowdown in productivity growth was accommodated by introducing kinks in the time trend, allowing β to shift at periods of structural change. These extensions were only arbitrary "ex post" responses to the failure of the linear time trends model, but the "split time trend" or "segmented trends" method was still in use at the end of the 1980's[4]. However, it became apparent that time trend methods could systematically mislead policymakers regarding both the level of potential output and the uncertainty surrounding its forecasts (Stock and Watson 1988b).

The poor performance of macro-economic models in the 1970's lead to a radical reassessment launched in part by two seminal articles. Nelson and Plosser (1982) found that growth in US output (and most other macro-economic series) was not adequately described as fluctuations around a deterministic trend. They suggested that a more appropriate model was

$$y_t = y_{t-1} + \mu + \varepsilon_t$$

where y_t is the log of output, μ is a constant term and ε_t is a stochastic disturbance. This is known as the random walk with drift, since the drift term μ represents the average growth rate of the random walk process for y_t. In such a framework the output series y_t is non-stationary because its mean and variance change over time. The shocks to y_t represented by ε_t have a permanent effect on its level (they are included in y_{t+1} the following period through

3 For a short history of this approach, see Butler (1996).

4 Giorno et al. (1995), Fisher et al. (1997), de Brouwer (1998) Botas et al. (1998) and Slevin (2001) present log linear time trend (or split time trend) estimates of potential output, but mostly for the purposes of comparison with more recent methods.

the lagged output term). Formally, y_t is said to be an *integrated* variable because its level can be obtained by summing past shocks (integrating in continuous time). In this case, y_t is integrated of order one or I(1) because first-differencing[5] is sufficient to induce stationarity (the mean and variance become constant through time). In fact, $\Delta y_t = y_t - y_{t-1} = \mu + \varepsilon_t$, which has constant mean μ and constant variance $Var(\varepsilon_t)$.

The results found by Nelson and Plosser suggested that output consisted of a non-stationary trend component plus a stationary cyclical component. They warned that in this case regressions on a linear time trend would "...confound the two sources of variation, greatly overstating the magnitude and duration of the cyclical component." A second seminal article, Nelson and Kang (1981), further developed this warning. This study found that if potential output has random walk component rather than following simple exponential growth or a linear function of time, then time trend methods will fail to deliver robust results, leading to inherent instability in output gap estimates. In particular, inappropriate detrending of time series will produce apparent evidence of cycles that are purely an artefact of the trend removal procedure. The resulting output gap measures will be entirely spurious, and the residuals of the regression would be non-stationary. This would make it difficult to interpret them as the cyclical component of output since they would not feature mean-reversion.

1.1.2 Univariate Filters

The advantage of simplicity that characterises methods based on linear time trends is shared by univariate filtering methods. However, filtering methods have the additional advantage that they allow potential output growth to change smoothly through time. This ability to react gradually to new information allows filtering methods to avoid the systematic errors to which time trend methods are prone. The simplest univariate filter is a moving average, which calculates potential output at any point in time as a weighted average of current and past values of observed output. This should help even out cyclical effects (assuming the filter is of the same length as the cycle) and allows the trend to "bend" smoothly over time rather than introducing sudden breaks as in the segmented trends model. More sophisticated filters include the band-pass filter proposed by Baxter and King (1995) and the HP filter of Hodrick and Prescott (1997). In the following we focus on the HP filter because it has proved the most popular[6] and because other univariate filters share most of its limitations.

The Hodrick-Prescott (1997) HP filter is based on the assumption that a given time series y_t is the sum of a trend or growth component g_t and a cyclical component c_t.

$$y_t = g_t + c_t \qquad \text{for } t = 1,\ldots,T$$

5 An I(2) variable requires second-differencing to induce stationarity (i.e. $x_t = x_{t-1} + y_t$), etc.

6 Although Hodrick and Prescott's article was not published until 1997, it had circulated widely in working paper format since 1981.

According to Hodrick and Prescott, "our prior knowledge is that the growth component varies 'smoothly' over time," where the measure of smoothness of the $\{g_t\}$ path is chosen to be the sum of the squares of its second difference. The cyclical component c_t represents deviations from g_t and over long time periods their average is assumed to be near zero. The growth component g_t is extracted by minimising the following loss function.

$$\underset{\{g_t\}_{t=-1}^{T}}{Min}\left\{\sum_{t=1}^{T} c_t^2 + \lambda \sum_{t=1}^{T}\left[(g_t - g_{t-1}) - (g_{t-1} - g_{t-2})\right]^2\right\}$$

The first sum represents the penalty for deviations of the observed series from the trend growth series ($c_t = y_t - g_t$), while the second sum represents the penalty for sharp changes in the trend growth component. The parameter λ is crucial, as it represents the terms on which deviations from trend are traded off against variability in the trend. The higher is λ the "stiffer" is the trend component. In fact, when $\lambda \to \infty$ the trend becomes a straight line and the HP filter gives the same result as the linear time trends method.

Unfortunately, the results can be quite sensitive to the choice of λ and there is no objective criterion by which to choose this parameter. According to Hodrick and Prescott, "if the cyclical components and the second differences of the growth components were identically and independently distributed, normal variables with means zero and variances σ_1^2 and σ_2^2 (which they are not), the conditional expectation of the g_t, given the observations, would be the solution ... when $\sqrt{\lambda} = \sigma_1 / \sigma_2$." On this basis, they recommend a value of $\lambda = 100$ for yearly data and $\lambda = 1600$ for quarterly data. These values have become "industry standards" but are actually arbitrary[7].

The HP filter proved popular because it is simple to use and can be applied mechanically to a large number of series[8]. Two arguments commonly made in its favour are that it extracts the relevant business-cycle frequencies[9] of the spectrum and that it closely approximates the cyclical component implied by reasonable time-series models of output. However, several studies raised serious doubts about these claims and about the reliability of the HP filter as a means of extracting trend components.

Harvey and Jaeger (1993) interpreted the HP filter as a set of restrictions within the framework of a structural time series or unobserved components model[10]. They found that these restrictions were coherent with US output data, for which the HP filter did seem particularly well suited. However, this was not necessarily the case for other US macroeconomic series nor was it the case for many non-US series, which are routinely detrended using the HP filter. The HP filter is optimal only for a selected class of time series processes. According to King and Rebelo (1993) this class includes cases when the series is I(2), when there are identical propagation mechanisms for innovations in the growth rate and

7 Recently, there have been some attempts to determine the value of λ endogenously. See Kaiser and Maravall (2001) and Ravn and Uhlig (2002).

8 See Ongena and Roeger (1997) on the use of the HP filter at the EU Commission.

9 As defined by NBER researchers, these are cycles lasting no less than 6 and no more than 32 quarters (see St-Amant and van Norden 1997).

10 This framework is described in more detail in the next section.

in the cycle, and when the smoothing parameter λ is known. Unfortunately, these conditions are rarely met in practice. Harvey and Jaeger showed that the HP filter can generate spurious cycles if the true data generating process does not belong to this class of time series processes. Cogley and Nason (1995) also found that the HP filter can induce spurious cycles when used with data that is integrated or nearly integrated (as most macro-economic series are believed to be).

Guay and St-Amant (1996) systematically explored the HP filter's ability to capture business cycle frequencies in the spectrum of different macro-economic time series. They found that the HP filter is inadequate when the spectrum of the series has the typical Granger shape that characterises most macro-economic variables (peak at zero frequency and bulk of the variance at low frequencies). In particular, the HP filter induces a peak inside business cycle frequencies even though it is absent in the original series. In addition, it fails to capture a significant fraction of the variance contained in business cycle frequencies but captures some variance originating outside these frequencies. Guay and St-Amant also produced Monte Carlo evidence indicating that the standard value of λ=1600 was only appropriate under implausible joint assumptions about the relative importance of demand and supply shocks and about the persistence of cycles.

The HP filter also suffers from the well-known *end-of-sample problem*. In the middle of the sample the HP is a symmetric two-sided filter as both leads and lags of output appear in the loss function. However, at the beginning or end of the sample, some leads or lags will be unavailable. This requires either gradually transforming the HP to a one-sided filter towards the edges of the sample or else generating forecasts of output outside the sample of observations. Uncertainty surrounding the estimated trend and cycle decomposition increases in both cases. St-Amant and van Norden (1997) point out that this is particularly unfortunate since the focus for policy advice is on estimating the *current* output gap at the end of the sample to help determine policy for the future. They underlined the instability of HP estimates of the output gap by showing how estimates can change dramatically when new observations are made available and incorporated into the sample. In fact, Canova (1993) underlined that if potential output includes a random walk component, rather than being a simple exponential or linear function of time, then smoothing methods will inevitably fail to deliver robust results, because they are inherently unstable as the sample size increases. The methods considered in the following section avoid this problem because they explicitly allow for a random walk component in output.

1.1.3 Univariate Unobserved Components Models

The autoregressive integrated moving average (ARIMA) models proposed by Box and Jenkins in the 1970s have been used widely to forecast economic time series. This approach refers to integrated variables with autoregressive (AR) or moving average (MA) components as ARIMA(p,d,q) where p and q respectively denote the order of the autoregressive and moving average terms and d indicates the order of integration (the number of times the series must be differenced before it is stationary). Assuming for simplicity that y_t is stationary, a pure AR(p) process can be written

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t \qquad \text{where } \varepsilon_t \text{ is white noise}^{11}$$

and a pure MA(q) process can be written

$$y_t = \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \qquad \text{where } \varepsilon_t \text{ is white noise}$$

A mixed ARMA(p,q) process can be written

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \quad \text{where } \varepsilon_t \text{ is white noise}$$

Beveridge and Nelson (1981) showed that every variable admitting an ARIMA(p,1,q) representation could be decomposed into a trend component specified as a random walk (possibly with drift) and a cyclical component that was stationary. In a sense this is obvious: if the variable under question is non-stationary before first-differencing then it must contain a random walk component. If the cyclical component is to be stationary (which seems sensible a priori) then the random walk component must be assigned to the trend. The simplest illustration of the Beveridge-Nelson result is for the case of an ARIMA(0,1,1) model. Suppose that $\Delta y_t = y_t - y_{t-1}$ follows an MA(1) process:

$$\begin{aligned} y_t &= y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1} \qquad \text{where } \varepsilon_t \text{ is white noise} \\ &= (y_{t-2} + \varepsilon_{t-1} + \theta\varepsilon_{t-2}) + (\varepsilon_t + \theta\varepsilon_{t-1}) \\ &= \sum_{r=1}^{t} \varepsilon_r + \theta \sum_{r=1}^{t-1} \varepsilon_r \\ &= (1+\theta)\left(\sum_{r=1}^{t} \varepsilon_r\right) - \theta\varepsilon_t \end{aligned}$$

11 A white noise process ε_t is such that $E[\varepsilon_t] = 0$, $Var[\varepsilon_t] = \sigma_\varepsilon^2 < \infty$, and $E[\varepsilon_t \varepsilon_s] = 0$ for all $t \neq s$.

Writing the trend component as $g_t = (1+\theta)\Sigma_r\varepsilon_r$ and the cyclical component as $c_t = -\theta\varepsilon_t$, the final expression can be rewritten

$$y_t = g_t + c_t \qquad \text{where } g_t = g_{t-1} + (1+\theta)\varepsilon_t$$

Evidently, the trend component g_t is a random walk with no drift and the cyclical component c_t is stationary.

This result is coherent with the evidence of high persistence found by Campbell and Mankiw (1987) when they fitted ARIMA models to quarterly US GDP series. They set out to assess whether output fluctuations were transitory and found a degree of persistence so high that it was incompatible with the premise that fluctuations in output are dominated by temporary deviations from a natural rate of output. It seemed more coherent with the idea that trend output followed a random walk.

The idea that macroeconomic time series were characterised by *variable trends* rather than deterministic trends fixed through time was emphasised by Stock and Watson (1988b). They suggested modelling these variable trends as random walks. In this case, what they called a *stochastic trend* would be increasing each period by some fixed amount (say 1 percent) on average. However, in any given period the change in the trend would deviate from its average by some non-forecastable amount. The stochastic trend assumption was also used in Watson (1986) to approach the problem of decomposing observed series into a trend and a cycle. However, Watson drew on the class of unobserved components (UC) models[12]. As an example, Watson defined the trend as a random walk with drift and the cycle as a stationary autoregressive process of second order. Letting y_t denote the natural log of observed output, g_t the trend (or growth) component and c_t the cyclical component:

$$y_t = g_t + c_t$$

$$g_t = g_{t-1} + \mu_g + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \upsilon_t \qquad \upsilon_t \sim iid(0, \sigma_\upsilon^2)$$

where ε_t and υ_t are uncorrelated white noise disturbances respectively representing permanent and transitory shocks to real output. For the AR(2) process driving the cyclical component to be stationary, the coefficients ϕ_1 and ϕ_2 must sum to less than unity (otherwise the υ_t shock will have permanent effects on the cyclical component). Note that the linear time trend represents the special case when the shock in the random walk has zero variance, $\sigma_\varepsilon^2 = 0$.

12 See Engle (1978), Harvey and Todd (1983) and Harvey (1985). Overviews of the UC class of models are available in Harvey (1989), Hamilton (1994) and Maravall (1995).

Harvey and Jaeger (1993) provided a more general model in the unobserved components approach. In this case, output was decomposed into three terms, the trend, the cycle and an additional irregular component ε_t. The trend component was specified as a random walk with drift, but in addition the drift term was also allowed to evolve over time according to a random walk.

$$y_t = g_t + c_t + \varepsilon_t \qquad \varepsilon_t \sim nid(0, \sigma_\varepsilon^2)$$

The trend component is a random walk with changing drift (local linear trend) defined as

$$g_t = g_{t-1} + \mu_{t-1} + \eta_t \qquad \eta_t \sim nid(0, \sigma_\eta^2)$$

$$\mu_t = \mu_{t-1} + \zeta_t \qquad \zeta_t \sim nid(0, \sigma_\zeta^2)$$

where μ_t is the slope, which itself follows a random walk, and the normal white noise disturbances, η_t and ζ_t, are independent of each other. The stochastic cycle is generated

$$c_t = \rho \cos(\lambda_c c_{t-1}) + \rho \sin(\lambda_c \tilde{c}_{t-1}) + \psi_t \qquad \psi_t \sim nid(0, \sigma_\psi^2)$$

$$\tilde{c}_t = -\rho \sin(\lambda_c c_{t-1}) + \rho \cos(\lambda_c \tilde{c}_{t-1}) + \tilde{\psi}_t \qquad \tilde{\psi}_t \sim nid(0, \sigma_\psi^2)$$

where ρ is a damping factor between zero and one, λ_c is the frequency of the cycle in radians, and the white noise disturbances ψ_t and $\tilde{\psi}_t$ are independently distributed. The cyclical component is equivalent to an ARMA(2,1) process with restrictions constraining the AR parameters to fall in the region corresponding to complex roots. This constraint is desirable a priori since the purpose is to model stochastic cycles.

Unobserved Components models such as Watson (1986) and Harvey and Jaeger (1993) have the advantage that they provided lower estimates of persistence than the ARIMA models fitted by Campbell and Mankiw (1987). Compared to filtering methods, UC models have two additional advantages. First, they can provide an estimate of the degree of uncertainty attached to the estimated level of potential output (or output gap) in the form of confidence intervals. Second, they have no end-of-sample problem and can actually provide out-of-sample forecasts on the observable variables, which can be used to assess the goodness of fit of the adopted specification, as well as ensuring policy relevance of the method.

Unfortunately, Quah (1992) showed that there is an infinite number of ways to decompose any given time series into trend and cyclical components. In effect, the amount of smoothness in the trend is arbitrary and depends on a priori assumptions about the dynamics of the components. In particular, assuming that the trend component follows a random walk (the "stochastic trend" specification chosen by most of the studies cited in this section) biases the analysis to find an important permanent component. This lead Quah to comment that

“Because studying the univariate time series characterizations of a variable leaves unidentified the sources of that variable’s fluctuations, without additional ad hoc restrictions those (univariate) characterizations are completely uninformative about the relative importance of the underlying permanent and transitory components.” The following section describes methods that adopt a multivariate framework, addressing this issue by analysing the joint dynamics of several economic variables.

1.2 Multivariate Methods

In the multivariate context too there is a variety of methods for estimating potential output and the output gap. These are generally based on a structural model of how different macro-economic variables interact. Possibly the most intuitive method is the production function approach, which postulates an aggregate production function (usually of Cobb-Douglas form) and attempts to estimate potential output as the maximum level of output that is feasible using a level of the inputs compatible with long-run equilibrium. Among the other multivariate approaches, there are several hybrid methods, combining univariate techniques (the HP filter, Beveridge-Nelson decomposition, or Unobserved Components models) with information derived from the dynamics of other macroeconomic variables. Yet another multivariate approach is based on the Vector Autoregression (VAR) and involves imposing identifying restrictions that make it possible to separate permanent from transitory shocks to output.

1.2.1 Production Function Approach

The Production Function Approach is really an extension of growth accounting. It provides an analysis of the key economic factors underlying the evolution of output in the medium term, including capital accumulation and changes in labour supply due to demographics, migration and shifts in labour participation. The key assumption is that the production process can be represented by an aggregate production function. Potential output is then calculated as the output of this function when all factors of production are at their “normal” or “natural” levels. Although the production function can take several functional forms (i.e., constant elasticity of substitution[13], translogarithmic, etc.) the Cobb-Douglas form is usually the one implemented:

$$Y_t = A_t(K_t)^{1-\alpha}(L_t)^{\alpha}$$

where Y_t is the level of output, K_t and L_t the level of factor inputs (capital and labour) and A_t is the scale factor that will change with technological progress. As written above, the Cobb-Douglas imposes constant returns to scale as the exponents on the capital and labour inputs sum to unity. This means that a proportionate increase in all factors will lead to an increase in output of the same proportion (assuming constant technology $A_t = A_{t-1}$). Increasing or decreasing returns to scale are possible in special circumstances but are generally not

13 The CES specification is implemented by Bolt and van Els (2000) and by Dimitz (2001).

considered relevant in this context. As is well known, the Cobb-Douglas becomes linear after taking logs. Letting lower case letters denote the natural logarithm of the corresponding upper case variable, taking first differences ($\Delta x_t = x_t - x_{t-1}$) and rearranging provides the "Solow residual" measure of technological progress:

$$\Delta TFP_t = \Delta a_t = \Delta y_t - (1-\alpha)\Delta k_t - \alpha \Delta l_t$$

For a given measure of α, this equation indicates the amount of the increase in output Y that cannot be accounted for by the increase in factor inputs K and L. This "extra" growth attributed to technological progress is known as the increase in Total Factor Productivity (TFP). TFP represents the productivity associated with the combination of all the factors of production, as opposed to the productivity of labour or capital in isolation. Note that TFP growth is a broader term than technological progress in that it includes not only increases in knowledge but also increases in efficiency due to the re-organisation of production processes.

The crucial parameter α that needs to be estimated is the elasticity of output with respect to labour. If capital and technological progress are constant, a 1% increase in labour will lead to an α% increase in output. Under the conditions of perfect competition in markets for goods and labour, α should coincide with the labour share (that is to say the wage bill divided by gross value added). However, the labour share as calculated from national accounts data varies over time and values of α obtained by direct estimation of the production function can differ markedly. This has lead several authors to attempt to estimate the production function parameters more accurately by imposing cross-equation restrictions within a system of simultaneous equations (Adams and Coe 1990, Layard et al. 1991, Fisher et al. 1997, Slevin 2001). In practice, however, the historical average of the labour share is often used or the labour share is smoothed using the HP filter.

The production function approach has been used widely by international organisations such as the IMF (De Masi 1997) and the OECD (Giorno et al. 1995). It can be extended to include other inputs (such as energy or imported materials) or to decompose the labour input to take account of long-term trends. For example, labour input, as measured by the number of hours worked, will be affected by changes in the patterns of part-time work, in the rate of participation in the labour force, in the age structure of the population and in the rates of population growth and migration. Using the production function approach, separate assumptions about these different factors can be combined to produce forecasts of potential output.

However, the production function approach also suffers from certain disadvantages. First, the Cobb-Douglas specification is necessarily a simplistic representation of the production technology. In addition, the approach may be subject to omitted variable bias due to improper use of value-added data or the assumption of perfectly competitive markets for inputs and output. Second, there are important measurement problems with the data, not just with respect to the number of hours worked (where data is often partial or unavailable) but especially with respect to capital inputs. Apart from the need to correct for variable capacity utilisation, measures of the capital stock are notoriously unreliable and subject to many methodological difficulties. Third, it is hard to make sense of the notion of “normal” or “natural” level of inputs. Potential output cannot be defined as the maximum level of output feasible in the engineering sense because full capacity use of the capital stock would be unsustainable in the long run. This is not only because of routine breakdowns and servicing requirements for regular maintenance but also because technical progress requires time to replace obsolete capital stock and to learn to use new technology. Operating at the maximum level of output would also imply an extraordinarily high employment rate that could be socially costly, as it would involve sustained overtime and high participation among all social groups including the young, the elderly, and the handicapped. Some assumption is needed as to the maximum “sustainable” level of the inputs corresponding to a “normal” level of capacity usage, workforce participation, or unemployment. Unfortunately, the HP filter has often been implemented here to extract a trend level of capacity usage or unemployment. This revives the problems discussed above, i.e. the end-of-sample bias. The problem of smoothing output is simply shifted to the problem of smoothing inputs. Finally, Solow found that TFP growth accounted for a surprisingly important part of overall growth. However, since the production function approach derives TFP as a residual, it is treated as exogenous, leaving it unexplained by economic theory. This gives no guidance as to the appropriate assumptions concerning TFP growth when forecasting into the future. Practitioners of the production function approach have often resorted to linear time trends or the HP filter to extract a trend TFP to be projected into the future. But in applying these techniques they run into the same problems as the simple statistical methods that they often dismiss.

1.2.2 Hybrid Methods

There have been several attempts to improve upon the univariate methods described previously by combining them with economic theory in a more coherent manner. The HP filter, the Beveridge-Nelson decomposition, and Unobserved Components models have all been implemented in a multivariate context. This allows them to link output growth to the dynamics of other macroeconomic variables in an attempt to include structural information derived from economic theory.

1.2.2.1 The Multivariate HP Filter

As noted previously, the univariate HP filter chooses the trend as the solution to

$$\underset{\{g_t\}_{t=-1}^{T}}{\text{Min}}\left\{\sum_{t=1}^{T} c_t^2 + \lambda \sum [(g_t - g_{t-1}) - (g_{t-1} - g_{t-2})]^2\right\}$$

where c_t is the cyclical term (the deviation of observed output from trend) and g_t is the trend. The multivariate implementation of the HP filter includes an additional term

$$\underset{\{g_t\}_{t=-1}^{T}}{\text{Min}}\left\{\sum_{t=1}^{T} c_t^2 + \lambda_g \sum_{t=1}^{T} [(g_t - g_{t-1}) - (g_{t-1} - g_{t-2})]^2 + \lambda_\varepsilon \sum_{t=1}^{T} \varepsilon_t^2\right\}$$

where $\varepsilon_t = z_t - f(g_t, x_t)$, z_t is some other economic variable of interest, and f(.) models z_t as a function of some explanatory variables x_t and the estimated trend g_t. The new term in ε_t is the residual from this structural relationship as $z_t = f(g_t, x_t) + \varepsilon_t$. As a result of the additional term, the HP filter chooses the trend to simultaneously minimise deviations of output from trend, minimise changes in the trend's growth rate, and maximise the ability of the trend to fit the additional structural relationship. The two parameters λ_g and λ_ε now represent the relative weights of these different objectives.

Of course, the reasoning can be extended to include the residuals from more than one structural relationship. The only requirement is that the trend term appear as an explanatory variable. The above equation can be generalised to include an arbitrary number *n* of structural relationships, each including the common trend g_t:

$$\underset{\{g_t\}_{t=-1}^{T}}{\text{Min}}\left\{\sum_{t=1}^{T} c_t^2 + \lambda_g \sum_{t=1}^{T} [(g_t - g_{t-1}) - (g_{t-1} - g_{t-2})]^2 + \sum_{i=1}^{n} \lambda_i \sum_{t=1}^{T} \varepsilon_{it}^2\right\}$$

Laxton and Tetlow (1992) implemented the multivariate HP filter using two different structural relationships based on the difference between potential output and its actual level (the output gap). The first was the (expectations augmented) Phillips curve, which states that inflation will be above expectations when output is above the (non-accelerating inflation) level of potential output:

$$\pi_t = \pi_t^e + A(L)(y_t - g_t) + \varepsilon_{\pi,t}$$

where π_t is the level of price inflation, π_t^e is the expected level of inflation, $(y_t - g_t)$ is the output gap (the deviation of observed output from trend or potential) and $\varepsilon_{\pi,t}$ is a stochastic error term associated with this relationship. The notation A(L) denotes a polynomial in the lag operator[14] meaning that past values of the output gap can have an effect on the deviation of inflation from its expected values.

14 The lag operator L is defined $Lx_t = x_{t-1}$, and $L^2x_t = L(Lx_t) = x_{t-2}$ so that in general $L^m x_t = x_{t-m}$. A polynomial in L of order *n* has the following form: $A_0 + A_1L + A_2L^2 + \ldots + A_nL^n$.

The second relationship used by Laxton and Tetlow is Okun's law. This relationship is really an empirical regularity rather than a law. As observed in US data, it holds that unemployment declines 1% for every 2.2 percentage points of output above the trend rate. As implemented by Laxton and Tetlow, Okun's law explains the level of the unemployment rate by reference to the underlying "structural" or "natural" rate of unemployment and past deviations from the level of potential output.

$$u_t = u_t^* + B(L)(y_t - g_t) + \varepsilon_{u,t}$$

where u_t is the unemployment rate, u_t^* is the structural unemployment rate, $(y_t - g_t)$ is the output gap again and $\varepsilon_{u,t}$ is another stochastic error term. The polynomial in the lag operator B(L) indicates that past values of the output gap can have an effect on the current unemployment rate. Of course, implementing this equation requires some measure of the structural or natural rate of unemployment. This is usually provided by the non-accelerating inflation rate of unemployment (NAIRU) which can be estimated by a variety of different methods. An added complication is that u_t^* can change over time. Butler (1996) constructed a measure of u_t^* that took account of the aggregate participation rate in short-run equilibrium. He called the resulting measure the equilibrium labour input. Conway and Hunt (1997) extended the multivariate HP filter to include a third structural relationship based on capacity utilisation

$$cu_t = cu_t^* + C(L)(y_t - g_t) + \varepsilon_{c,t}$$

where cu_t is the an indicator of capacity utilisation derived from survey data and the "normal" level of capacity utilisation cu_t^* was estimated as a constant.

Before the filtering problem can be solved, the parameters attached to the lagged output gap terms (implicit in the polynomials A(L), B(L) and C(L)) need to be estimated. This leads to an iterative procedure whereby an initial estimate of the trend series is used to estimate these parameters, the loss function is minimised and the resulting output gap is then plugged into the regressions for a new estimate of these parameters. A separate problem arises in choosing the λ_i weights attached to the different parts of the loss function. As the fit of the individual equations usually varies dramatically, their squared residuals need to be weighted differently in the loss function.

The multivariate HP filter has been criticised on several grounds (see St-Amant and van Norden 1997). First, the estimated output gap can be sensitive to the specification of the structural equations. Second, in practice structural information has not much improved performance at the end of the sample. Third, the structural parameters cannot be estimated at an acceptable level of precision.

1.2.2.2 Multivariate Unobserved Components

Compared to smoothing methods such as the HP filter, the unobserved components models already had several advantages in the univariate context. First, unlike the HP filter, unobserved components models can provide a measure of the uncertainty with which potential output is measured in the form of confidence intervals. Second, unobserved components models can easily generate forecasts that can be used in-sample to check goodness of fit or out-of-sample to produce policy advice. Third, unobserved components models provide a flexible balance between structure and parsimony. This makes them easy to extend to a multivariate context.

Kuttner (1994) modified the univariate stochastic trend specification in Watson (1986) by simply adding a fourth equation representing the Phillips curve.

$$y_t = g_t + c_t$$

$$g_t = g_{t-1} + \mu_g + \varepsilon_t \qquad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \upsilon_t \qquad \upsilon_t \sim iid(0, \sigma_\upsilon^2)$$

$$\Delta\pi_t = \mu_\pi + \gamma\Delta y_{t-1} + \beta c_{t-1} + \nu_t + \delta_1 \nu_{t-1} + \delta_2 \nu_{t-2} + \delta_3 \nu_{t-3} \qquad \nu_t \sim iid(0, \sigma_\nu^2)$$

where the first equation decomposes output into trend and cycle, the second equation specifies that the trend component of output follows a random walk with drift μ_g, the third equation indicates that the cycle in output follows an MA(2) process, and the fourth equation is the modified Phillips curve.

It may be surprising in this last equation that changes in inflation, $\Delta\pi_t$, appear on the left hand side instead of the level of inflation, π_t, but this is a direct simplification of the expectations-augmented Phillips curve under the assumptions that expected inflation is totally backward looking. In addition, this specification is consistent with evidence that the unit root hypothesis cannot be rejected in the inflation process for many countries. Past changes in output (Δy_{t-1}) appear among the explanatory variables to capture the positive correlation between inflation and lagged real output growth. The lagged cyclical component (c_{t-1}) captures the impact of the output gap in the previous period, so the β coefficient can be interpreted as the slope of the Phillips curve. Note that the presence of the drift term μ_π and the MA(3) error process for ν_t in the fourth equation were justified on empirical grounds[15].

15 Kuttner found significant fourth order autocorrelation in quarterly US inflation series.

Gerlach and Smets (1997) extended Kuttner's analysis by generalising the Phillips curve to include log changes in the nominal trade-weighted exchange rate and in nominal oil prices. These variables were included to capture the effects of temporary relative price shocks to inflation. Kichian (1999) took the Gerlach-Smets model one step further, focussing on alternative specifications for the trend component of output and the treatment of inflation expectations, as well as considering core inflation measures.

Apel and Jansson (1999a,b) took a different approach, arguing that potential output and the NAIRU should be estimated simultaneously as they were related concepts which both affected inflation developments. Their multivariate unobserved components model included not only a Phillips curve equation but also an Okun's law relationship.

$$\pi_t = (1-\rho(1))\pi^* + \rho(L)\pi_{t-1} + \eta(L)(u_t - u_t^n) + \omega(L)z_t + \varepsilon_t^{pc} \qquad \varepsilon_t^{pc} \sim iid(0, \sigma_{pc}^2)$$

$$y_t - y_t^p = \phi(L)(u_t - u_t^n) + \varepsilon_t^{ol} \qquad \varepsilon_t^{ol} \sim iid(0, \sigma_{ol}^2)$$

$$u_t^n = u_{t-1}^n + \varepsilon_t^n \qquad \varepsilon_t^n \sim iid(0, \sigma_n^2)$$

$$y_t^p = \alpha + y_{t-1}^p + \varepsilon_t^p \qquad \varepsilon_t^p \sim iid(0, \sigma_p^2)$$

$$u_t - u_t^n = \delta(L)(u_{t-1} - u_{t-1}^n) + \varepsilon_t^c \qquad \varepsilon_t^c \sim iid(0, \sigma_c^2)$$

The first two equations represent a Phillips curve and an Okun's law relationship and provide the identifying equations of the system. The Phillips curve specification follows the triangle model of inflation in Gordon (1997). The initial term allows for an equilibrium level of inflation in the long term (π^*), which will drop out if inflation follows a random walk. The other terms represent the three components underlying the triangle model of inflation. First, inflation inertia (π_{t-1}) reflecting backward looking expectations or nominal price rigidities. Second, excess demand ($u_t - u_t^n$) measured by the gap between actual unemployment and the NAIRU. Third, a vector of supply shocks (z_t) including shocks to oil, import prices, productivity, etc. The second equation of the system represents Okun's law, translating fluctuations in cyclical output into fluctuations in cyclical unemployment (i.e. the unemployment gap).

The other three equations in the system specify the dynamics of the unobserved components (NAIRU, potential output and cyclical unemployment). In line with usual practice in the unobserved components literature, the NAIRU and potential output are assumed to follow stochastic trends. For the NAIRU this is specified as a pure random walk and for potential output as a random walk with drift.

Rasi and Viikari (1998) applied several variants of the Apel-Jansson multivariate unobserved components model to Finnish data, including both the level and the changes of the unemployment gap. Since the changes in the unemployment gap seemed more significant in explaining inflation developments, Rasi and Viikari concluded that there was evidence suggesting hysteresis effects. Fabiani and Mestre (2001) explored alternative specifications within the Apel-Jansson approach using euro area aggregate data. They found better performance incorporating a stochastic drift or local linear trend in the NAIRU dynamics as well as those of potential output. Fabiani and Mestre used bootstrap techniques to evaluate the uncertainty surrounding the estimates. Ross and Ubide (2001) evaluated several different measures of the euro area output gap, including variants of the Apel-Jansson approach. These included a NAIRU following a pure random walk, one with a stochastic trend, and one allowing for hysteresis effects.

Cerra and Saxena (2000) not only updated the Apel-Jansson results but also considered an alternative multivariate unobserved components model that included common permanent and cyclical components. In this framework, the growth rate of output may switch according to an unobserved state variable. This extension made it possible to allow for asymmetries in the business cycle, expansions being more persistent than contractions.

1.2.2.3 The Multivariate Beveridge-Nelson Decomposition

As mentioned in the univariate context, Beveridge and Nelson (1981) showed that any integrated variable admitting an ARIMA(p,1,q) representation could be decomposed into a stochastic trend and a stationary cyclical component. Stock and Watson (1988a) and King, Plosser, Stock and Watson (1991) applied multivariate versions of the Beveridge-Nelson decomposition to extract common stochastic trends from a set of variables. Evans and Reichlin (1994) considered the multivariate generalisation of the Beveridge-Nelson decomposition when the information set includes other I(1) and/or stationary variables.

Consider a $n \times 1$ stationary vector $W_t' = [\Delta X1_t \; X2_t]$, whose first n_1 elements, $\Delta X1_t$, are first differences of I(1) variables, and remaining n-n_1 elements, $X2_t$, are the levels of stationary variables. The vector W_t admits an MA representation $W_t = D + A(L)\upsilon_t$ where D is a deterministic n-dimension vector representing a constant term, $A(L) = \Sigma_j A_j L^j$ is a rational function of the lag operator, A_j is an $n \times n$ matrix, and υ_t is an n-dimension white noise innovation such that $E[\upsilon_t]=0$ and $E[\upsilon_t \upsilon_t'] = \Omega$ with Ω a positive definite matrix. The MA representation can be partitioned as follows:

$$W_t = \begin{pmatrix} \Delta X1_t \\ X2_t \end{pmatrix} = D + \begin{pmatrix} A1(L) \\ A2(L) \end{pmatrix} \upsilon_t$$

From which the multivariate generalisation of the Beveridge-Nelson decomposition can be written:

$$W_t = \begin{pmatrix} \Delta X1_t \\ X2_t \end{pmatrix} = \begin{pmatrix} D1 \\ 0 \end{pmatrix} + \begin{pmatrix} A1(1) \\ 0 \end{pmatrix} \upsilon_t + \begin{pmatrix} (1-L)\tilde{A}1(L) \\ (1-L)A2(L) \end{pmatrix} \upsilon_t$$

Where the sum of the first two terms on the right-hand side is interpreted as the change in the non-stationary trend component, while the third term is interpreted as the change in the stationary cyclical component. If the variables in X1 are not cointegrated[16], then A(L) can be inverted and the parameters can be estimated from a standard vector autoregression[17]. Otherwise, if the variables in X1 are cointegrated the standard error correction framework can be adapted to include the stationary variables in X2.

Evans and Reichlin showed that the relative importance of the cyclical component depends on the size of the information set and is necessarily higher in the multivariate case. This result is intuitive: since output growth can be forecasted better with multivariate models, they will ascribe more of output fluctuations to the cyclical component.

Lippi and Reichlin (1994) argued that the assumption that potential output follows a random walk is inconsistent with most economists' interpretation of productivity growth. It is generally believed that technology shocks affecting the supply side are only gradually absorbed by the economy. Adjustment costs for capital and labour, learning and diffusion processes, habit formation, and "time to build" all suggest that shocks due to technological progress will induce richer dynamics than a random walk. However, the random walk assumption is common to the Beveridge-Nelson decomposition and the unobserved components approach, whether implemented in the multivariate or the univariate framework. The structural vector autoregressions described next have the advantage that they let the data determine the shape of the diffusion process associated with permanent shocks to output.

1.2.3 Structural Vector Autoregressions

Sims (1980) proposed the Vector Autoregression (VAR) as an alternative to simultaneous equations models (see Canova 1995a,b for surveys of VAR methods). In a VAR, all variables are treated as endogenous and each is written as a linear combination of lagged values of itself and of the other variables in the system. Consider the $n \times 1$ vector W_t from the previous section. A VAR(p) model for this vector would be

$$W_t = D + A_1 W_{t-1} + A_2 W_{t-2} + \ldots + A_p W_{t-p} + \varepsilon_t$$

16 That is to say they do not share a common stochastic trend. Although cointegrated variables are non-stationary individually, there exists a linear combination of such variables that is stationary (because it eliminates the common stochastic trend).

17 See next section.

Where the A_j matrices are $n \times n$, the $n \times 1$ vector D may contain an intercept and/or deterministic trend and the vector of disturbances is such that $E[\varepsilon_t]=0$ and $E[\varepsilon_t \varepsilon_t'] = \Omega$ with Ω a positive definite matrix. The parameters in the matrices can be estimated simply by fitting separate regressions for each variable using ordinary least squares, which in this context is equivalent to the seemingly unrelated regressions estimator. In itself, this provides a useful tool for capturing the simultaneous dynamics of several time series and for producing forecasts without having to specify the future path of any exogenous variables. However, in this form the VAR has little economic content apart from the initial selection of the variables to include[18]. Sims suggested "inverting" the estimated VAR to obtain the Wold moving average representation of W_t in terms of lagged values of the disturbances ε_t. Ignoring the constant term D for simplicity, this can be obtained by solving out the lagged variables recursively.

$$\begin{aligned} W_t &= A_t W_{t-1} + \dots + A_p W_{t-p} + \varepsilon_t \\ &= (I - A_1 L - \dots . A_p L^p)^{-1} \varepsilon_t \\ &= (I + \Psi_1 L + \Psi_2 L^2 + \dots)\varepsilon_t \end{aligned}$$

The coefficients Ψ of the MA representation can be recovered as follows. Note that the VAR coefficients A_j and the MA coefficients Ψ_j must satisfy the following relationship:

$$I = (I - A_1 L - \dots - A_p L^p)(I + \Psi_1 L + \Psi_2 L^2 + \dots) = I + C_1 L + C_2 L^2 + \dots$$

Where $C_1 = C_2 = \dots = 0$. These conditions on C recursively define the MA coefficients

$$\begin{aligned} \Psi_1 &= A_1 \\ \Psi_2 &= A_1 \Psi_1 + A_2 \\ \Psi_s &= A_1 \Psi_{s-1} + A_2 \Psi_{s-2} + \dots + A_p \Psi_{s-p} \end{aligned}$$

From the MA representation, it is possible to calculate the response of individual variables to the different shocks that make up the vector ε_t. However, as estimated from the AR specification, the ε_t residuals are correlated with each other across the different equations, so they do not have a clear interpretation as shocks originating from separate sources.

Under the assumption that the elements of ε_t are linear combinations of uncorrelated structural disturbances u_t, Sims proposed recovering "orthogonalised" innovations uncorrelated with each other by using the Choleski decomposition of the estimated covariance matrix Ω. Since the covariance matrix is positive definite, it admits the factorisation $\Omega = PP'$ where P is a lower triangular matrix. The orthogonal innovations u_t can then be derived from the estimated disturbances using the formula $u_t = P^{-1}\varepsilon_t$.

18 Another a priori choice is involved in the selection of the lag length *p*, which can be based on one of several criteria (Akaike Information Criterion, Schwarz Bayesian Criterion, sequential LR test, etc.)

Consider a shock in one of the orthogonalised innovations. The impulse response of the dependent variables at lag *s* (holding the other shocks constant at zero) is then given by $\Psi_s P$. However, choosing the Choleski decomposition to derive P implies that the system has a recursive structure, with the innovation in the first equation contemporaneously affecting all the others but not being affected by them in the current period. The innovation in the second equation will be contemporaneously affected only by the first innovation but will affect all others in the current period, etc. This means that the order in which the variables are arranged in the vector W_t will affect the resulting impulse response.

The class of "structural" VARs are based on different restrictions that one can impose to derive P. Consider the matrix C representing the long-run response of the variables in the system to shocks in the "structural" innovations:

$$C = \Psi_\infty P^{-1} = (I - A_1 - A_2 - \ldots - A_p)^{-1} P^{-1}$$

Since the A_j matrices are estimated in the VAR, the P matrix could be recovered by imposing restrictions on the C matrix. Typically these take the form of zero restrictions. For example, $C_{i,j} = 0$ means that the (accumulated) long-run response of the *i*-th variable to the *j*-th structural innovation shock is zero in the long-run. Thus the estimated VAR is transformed by post-multiplying by a matrix that imposes the necessary conditions on the long-run multipliers and the residual covariance matrix. There is a unique matrix that can simultaneously diagonalise the VAR innovation covariance matrix and triangularise the matrix of long run multipliers.

Blanchard and Quah (1988) and Shapiro and Watson (1988) used this framework to separate output into its permanent and transitory components. They made the assumption that shocks that had no long-run effect on output were demand side disturbances, whereas shocks that had a long-run effect on output were supply side productivity shocks.

Blanchard and Quah specified a bivariate VAR including output and unemployment. Shapiro and Watson included output, hours worked, inflation, the nominal interest rate and real oil prices. King, Plosser, Stock and Watson (1991) extended the analysis to allow for cointegrated variables, using a structural VAR that included six variables to derive potential output. Cochrane (1994) used a similar framework to estimate potential output by exploiting the permanent income hypothesis. One implication of this theory is that consumption follows a random walk (for a constant real rate of interest). Therefore, assuming that output and consumption are cointegrated, any fluctuations in output that leave consumption unchanged must be transitory. This makes it possible to extract a measure of potential output that is constrained to be a random walk only to the extent that consumption has random walk characteristics. Claus (2000) used a structural VAR to estimate potential output using observed output, employment and a survey measure of capacity utilisation.

Compared to other methods of estimating potential output and the output gap, structural VARs have several advantages. First, they have no end-of-sample problem and can easily be used to forecast into the future. Second, they can provide confidence intervals for their estimates of potential output and the output gap. While the previous two advantages are shared with unobserved components methods, structural VARs have the added advantage that they are based on limited a priori information. This takes the form of identifying assumptions with a clear theoretical interpretation. Structural VARs also allow for richer dynamics of potential output than the strict random walk assumed in unobserved components models (Dupasquier et al. 1997).

On the other hand, Cooley and Dwyer (1998) warned that structural VAR results are likely to be quite sensitive to mis-specification in the identifying assumptions. However, since these assumptions are necessary for identification, by their very nature they cannot be tested. The identification strategy will be inappropriate if the included variables do not provide a good indication of cyclical developments in output. Finally, most estimates of output gap produced using structural VARs have found wider confidence intervals, suggesting greater uncertainty.

2. Estimation Results

The level of potential output is usually estimated using quarterly series to account for the short-term dynamics of the business cycle. Unfortunately, quarterly national accounts for Luxembourg have yet to be published, so this study relies on annual series for output and prices. Output, Y_t, is measured by Gross Domestic Product at 1995 market prices (in national currency). Statec, the Luxembourg national statistical institute, has only published ESA95 national accounts for the years 1995 to 2001 so this output series was spliced with ESA79 data drawn from the AMECO[19] database of the European Commission. Prices, P_t, were measured by the Harmonised Index of Consumer Prices (1995 = 100), available since 1995, and spliced with the National Index of Consumer Prices (NICP) for previous years.

Table 1: Autocorrelation functions

		Lag							
		1	2	3	4	5	6	7	8
y_t	**acf**	0.920	0.838	0.758	0.680	0.606	0.535	0.463	0.389
	pacf	0.920	-0.051	-0.034	-0.033	-0.019	-0.028	-0.052	-0.058
P_t	**acf**	0.948	0.891	0.830	0.767	0.702	0.634	0.564	0.491
	pacf	0.948	-0.075	-0.060	-0.059	-0.056	-0.056	-0.065	-0.064

The first step is to assess the stationarity characteristics of the data. Table 1 presents the sample autocorrelation and partial autocorrelation functions of the series for the period 1960-2001. Lower case letters indicate natural logarithms of the upper case variable.

Both series appear to be non-stationary in levels, as the autocorrelation function at lag one is near unity and it declines only very gradually, indicating substantial persistence. The partial autocorrelation function cuts off after lag one, suggesting a pure AR process is a more likely candidate than a pure MA process.

In table 2 the augmented Dickey-Fuller (ADF) test and the Phillips-Perron test confirm the presence of non-stationarity. The null hypothesis of a unit root cannot be rejected at conventional significance levels for any of the series and under all the specifications of the tests considered. For both tests, regressions were run with a constant only and with a constant and a deterministic trend. Up to eight lags of the dependent variable were included in the regression to account for any residual autocorrelation.

19 The AMECO codes for the series Y_t and P_t are LUX.1.1.0.0.OVGD and LUX.3.0.0.0.ZCPIN.

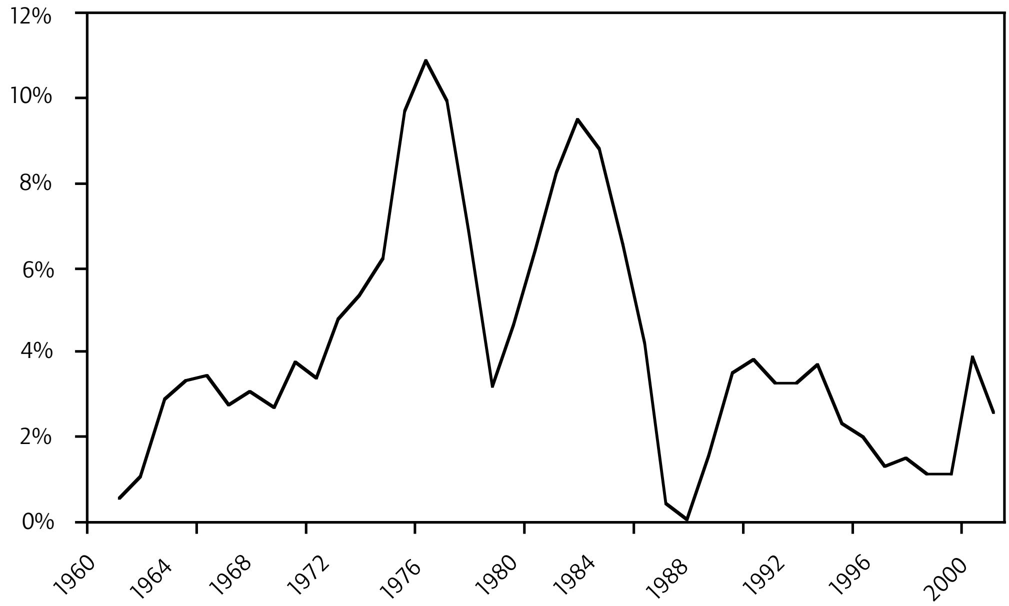
Table 2: Unit Root Tests

	log output (y_t)				log prices (p_t)			
	ADF		Phillips-Perron		ADF		Phillips-Perron	
lags	constant	trend	constant	trend	constant	trend	constant	trend
1	1.51	-0.77	1.72	-0.62	-1.67	-1.95	-0.96	-0.36
2	1.49	-0.52	1.78	-0.60	-1.52	-0.51	-0.89	-0.63
3	1.56	-0.42	1.87	-0.56	-1.38	-0.65	-0.86	-0.78
4	1.51	-0.97	1.86	-0.59	-1.75	-0.98	-0.85	-0.87
5	0.99	-1.53	1.74	-0.69	-2.09	-1.47	-0.84	-0.93
6	1.07	-0.95	1.73	-0.72	-2.20	-1.31	-0.83	-0.98
7	0.98	-0.77	1.76	-0.71	-2.41	-1.99	-0.82	-1.03
8	0.95	-0.78	1.81	-0.69	-2.76	-2.09	-0.82	-1.09

None of these statistics is significant at the 1% or 5% level. However, at the 10% level of significance, the ADF test for log prices with eight lags and a constant but no trend can reject the unit root hypothesis. However, in view of the discussion of deterministic trends cited above, it is preferred to maintain the hypothesis of a unit root in the price level. In general, the effects on inference of erroneously assuming that a variable is stationary around a deterministic trend when it actually includes a random walk are usually far worse than the effects of erroneously assuming a variable follows a random walk when it is actually stationary around a deterministic trend. Both the estimated autocorrelation functions and the unit root tests are qualitatively unchanged when observations before 1970 or before 1980 are dropped from the sample.

Figure 1 illustrates the evolution of consumer price inflation in Luxembourg. The two major episodes associated with the oil price shocks of the 1970s are clearly visible. While the effects of the first oil price shock peaked rapidly in 1974-1975, inflation following the second oil price shock only peaked in 1982-1983, when the Luxembourg Franc was devalued along with the Belgian Franc. A period of sharp disinflation followed, aided in part by the bursting of the bubble in the US dollar in 1985. The Luxembourg economy grew very rapidly in 1988 and 1999, when inflation began to rise again. There followed a period of steady disinflation over the years 1993-1999 and the end of the decade was marked by another spike in prices.

Figure 1: Consumer Price Inflation in Luxembourg



2.1 Linear trends and the HP filter

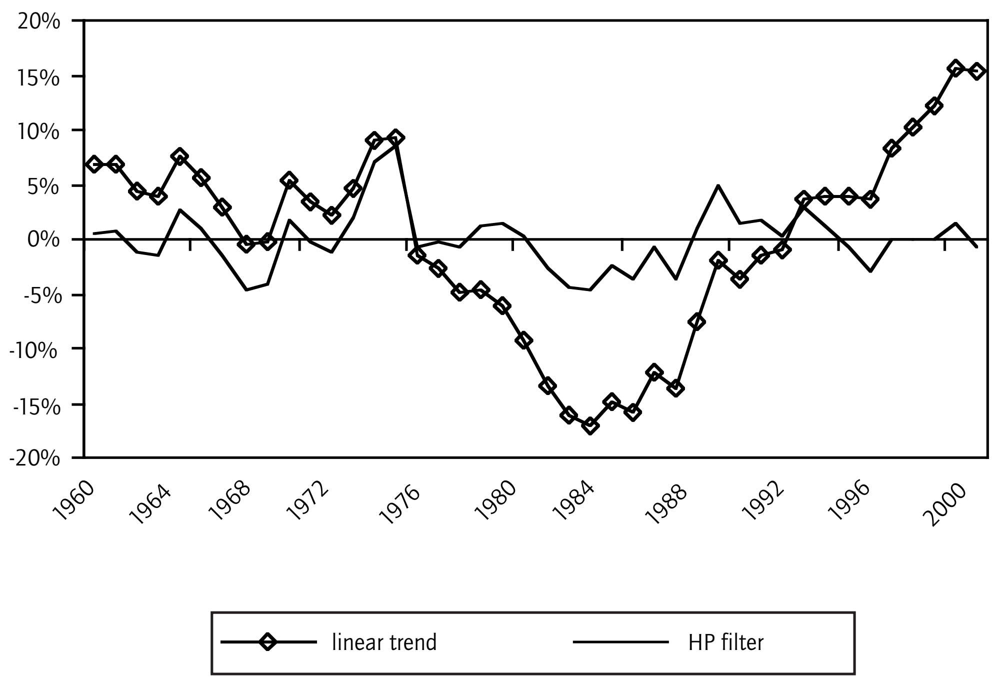
The simplest measures of potential output reviewed above were based on the linear time trend and on the Hodrick-Prescott filter. These are compared below. With a sample covering 1960-2001, regressing log output on a constant and a deterministic time trend yields the following results:

$$y_t = \underset{(0.027)}{8.12} + \underset{(0.001)}{0.038t} + \varepsilon_t \qquad R^2 = 0.966 \quad DW = 0.14$$

where asymptotic standard errors of the coefficient estimates appear in parentheses. Superficially, this may seem a good regression, with a high degree of variation explained (R^2 near unity) and highly significant *t*-statistics on the estimated coefficients. The coefficient on time suggests that trend output grows by 3.8% each year. However, the residuals are autocorrelated, as will appear below and as signalled by the low Durbin-Watson statistic (DW). In the presence of residual autocorrelation, ordinary least squares still provides a consistent estimator but it is inefficient and the associated covariance matrix estimator is biased downwards. This means that the *t*-statistics do not follow the standard Student distribution and cannot be used for valid inference. In fact, an R^2 in excess of the DW statistic is the trademark of a spurious regression due to the presence of a random walk component in one of the variables.

The residuals of the linear time trend regression are the simplest measure of the output gap. This is expressed in percentage of potential output in the following figure. A positive output gap signals that observed output is above potential output, suggesting that current growth relies on unsustainable levels of labour and capital inputs. This raises the possibility of the economy overheating, with increasing costs in the factor markets. A negative output gap indicates that observed output is below potential so that growth can accelerate without causing additional inflationary pressure. In the figure, the output gap measure based on the linear time trend regression is compared with the output gap measure calculated from the HP filter with smoothing parameter λ=100, the standard value for data at annual frequency.

Figure 2: Linear time trend residual and HP filter measures of output gap



The serial correlation of the linear trend regression residuals is obvious in the figure, as the gap measure is negative from 1975 to 1992 and positive thereafter. One could attempt to justify the appearance and then the disappearance of the gap over the 1980's as an indication of capacity initially growing faster than output following the introduction of the common market and free movement of goods and labour in the EU. Alternatively, one could refer to the appearance and initial development of the financial sector in Luxembourg. This period also saw the first large cross-border labour inflows. However, the size of the gap on this measure is simply unrealistic. It drops below -17% and then rises to more than +15% at the end of the sample. The HP filter, instead, can track actual output more closely, so that the implied output gap remains closer to the zero line because the trend can change smoothly over time. In this case, the largest positive gap is in 1974 (+8.5%) and the largest negative gap is in 1967 (-4.7%).

Although the rate of change of the output gap is sometimes similar on the two measures, the level is quite different and they often indicate gaps of opposite signs. It is important to remember that the HP method is based on a filter and therefore cannot produce a measure of uncertainty associated with the estimated gap. This means that it is impossible to test formally whether the two measures are statistically different. It is actually impossible to test whether the output gap as measured by the HP filter is statistically positive or negative. However, a more formal comparison of the various estimated output gap measures will be provided in Section 3.

2.2 Harvey-Jaeger univariate UC model

As discussed above, univariate unobserved components models have several limitations but they do provide an indication of the uncertainty attached to their estimate of trend. In addition, unobserved components models can easily produce forecasts out of sample. The output gap as measured by the HP filter is compared below to that produced by the Harvey and Jaeger (1993) unobserved components model. As discussed previously, this is based on the decomposition of output into trend, cycle, and irregular component. Estimation from the output series described above yields the following results:

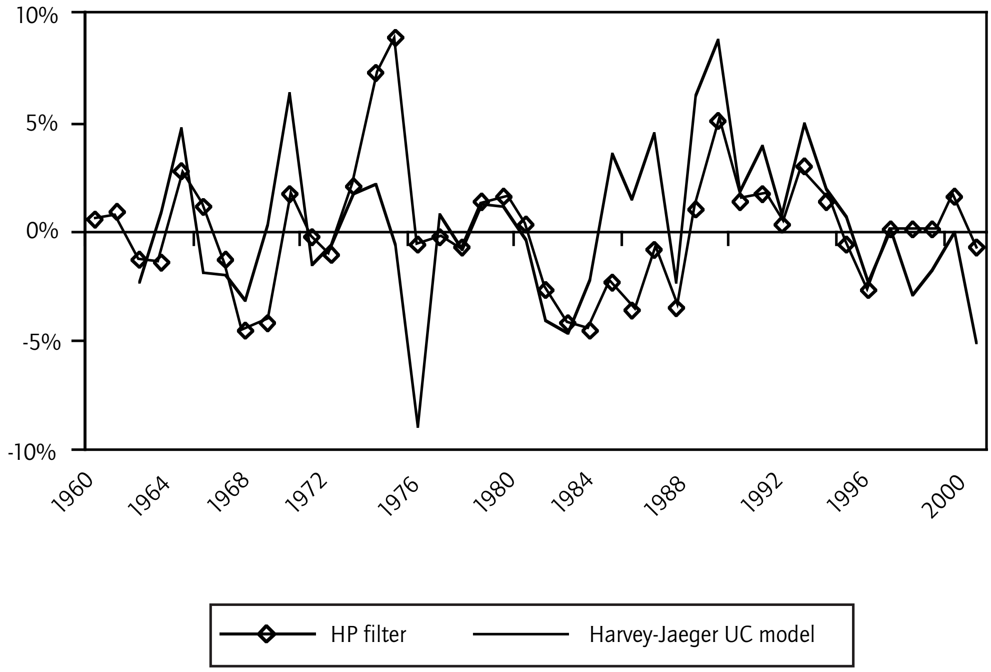
σ_ε	σ_η	σ_ζ	ρ	$2\pi/\lambda_c$	σ_ψ	SE	R_D^2	Q(14)
0.012 (0.728)	0.017 (1.000)	0.006 (0.351)	1.00	4.894	0.012	0.027	0.285	13.20

The first three parameters reported are the standard errors of the disturbances associated respectively with the irregular component (ε_t), with the level of the trend component (η_t) and with the slope of the trend component (ζ_t). The last of these is very low, indicating only gradual change in the drift of the trend component. The q-ratios that appear in parentheses suggest that none of these standard deviations is statistically significant. The parameter ρ is estimated at unity, indicating no damping (a possible sign of misspecification[20]). In the next column, $2\pi/\lambda_c$ indicates the period of the cyclical component. In this case, the period is 4.894 years, which is close to the value found by Harvey and Jaeger for the US and falls in the standard business cycle frequencies. The next parameter is the standard deviation of the disturbance associated with the cyclical component (ψ_t). The columns labelled SE and R_D^2 indicate two measures of fit, respectively the standard error of the predictions and the coefficient of determination with respect to the first differences. The last column is a diagnostic test, the Ljung-Box Q(*p*) statistic based on the first *p* residual autocorrelations. Since it follows a $\chi^2(p)$ distribution, the observed value of this test statistic cannot reject the null hypothesis of white noise residuals at conventional levels.

Apart from the period 1973-76, the output gap as measured by the univariate unobserved components model is fairly close to that produced using the HP filter. This is consistent with the results found by Harvey and Jaeger using US data. However, there are some differences at the beginning and end of the sample. On the Harvey-Jaeger UC measure, the gap drops from near zero to −5% at the end of the sample, where the HP filter moves from a small positive gap to a small negative gap. This could be due to the limitations of the HP filter near the edge of the sample, where it has to change from a two-sided filter to a one-sided filter. This means that in extracting the trend, the HP filter gives proportionately more weight to the observations near the edge than it does when it is in the middle of the sample. The difference in size of the gap in the end-of-sample period according to the two measures is critical for the formulation of policy advice. The HP filter suggests that output is near potential, while the Harvey-Jaeger UC model indicates a substantial negative output gap, suggesting that trend output is growing faster than observed output, leading to spare capacity.

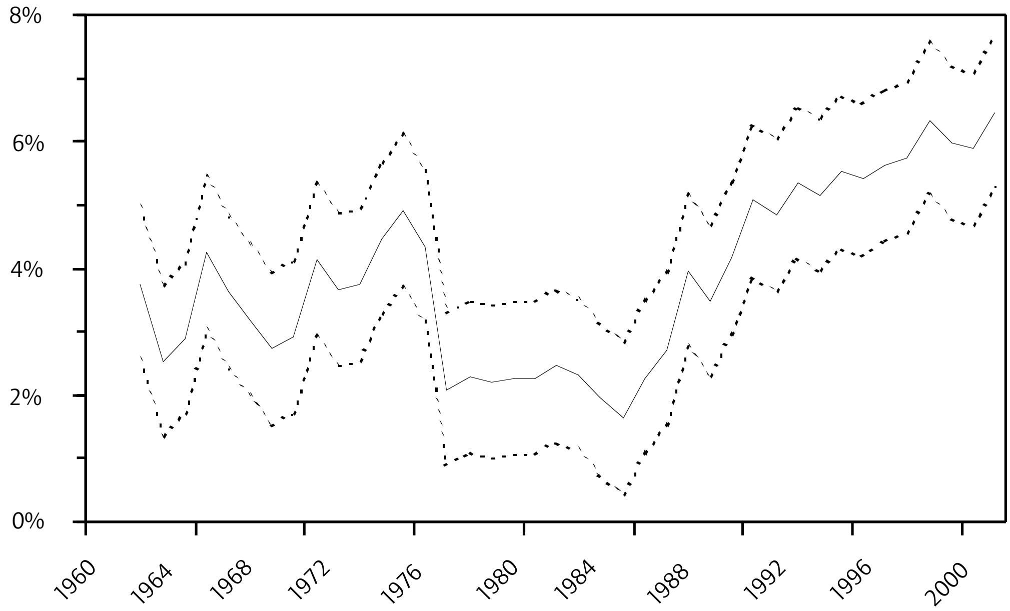
20 Other specifications of univariate UC models appear in Harvey (1985), Watson (1986) and Clark (1987).

Figure 3: HP filter and Harvey-Jaeger UC measures of output gap



The specification of the Harvey-Jaeger UC model used here assumed that trend output follows a random walk with a drift that can change over time. The following figure plots the evolution of the estimated drift term across time, indicating a sudden decline from 4.9% in 1974 to 2.1% in 1976. The drift term was fairly steady at this lower level until the mid-1980's when it began increasing, suggesting a structural shift in Luxembourg's economy that may match the move away from the steel sector and towards financial services. The drift term reaches a maximum of 6.5% at the end of the sample. The figure also provides a 95% confidence interval around the estimate of the drift term, calculated as $\pm 2\sigma_{\eta}$. Note that the low point of the series, in 1984, is outside the limits of the confidence interval both at the beginning and at the end of the sample. This is evidence against the hypothesis that the drift term has not changed significantly across the sample.

Figure 4: Slope term from the Harvey-Jaeger univariate UC model



2.3 Kuttner multivariate UC model

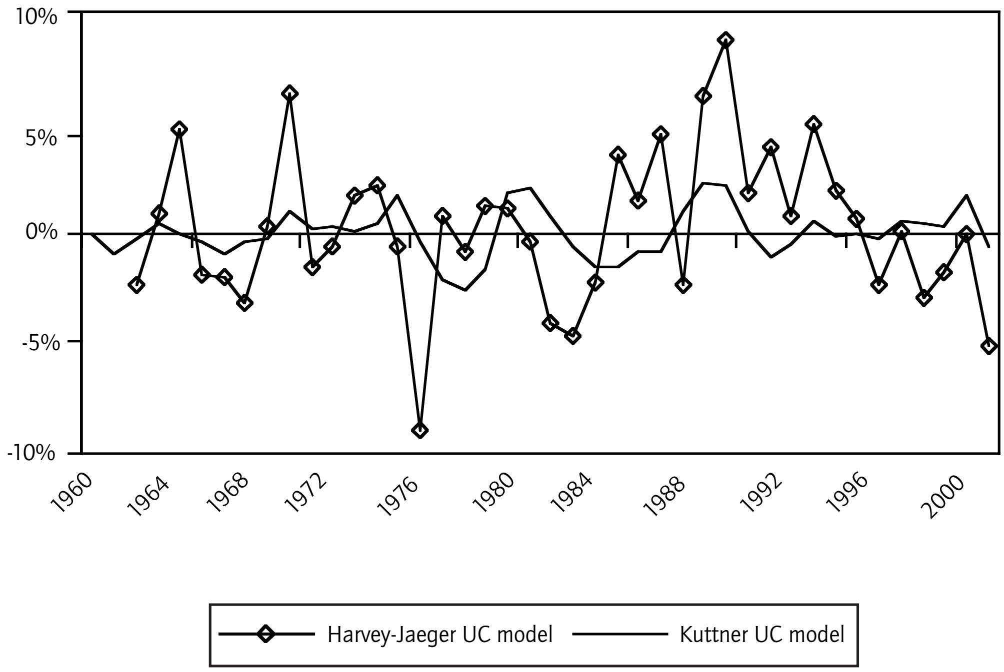
To estimate the Kuttner (1994) model using the Kalman filter, initial estimates of the state vector and covariance matrix were provided by the HP filter measures of output trend and cycle. Coefficients were initialised at the values obtained from OLS regressions using these initial HP estimates of the unobserved state variables. The trend component follows a random walk with drift and the cyclical component follows and AR(2) process. The error process of the Phillips curve was simplified to an MA(2) and the exogenous shocks in the trend and cycle equations were constrained to be orthogonal so that $E(\varepsilon_t \upsilon_t)=0$. Parameter estimates are reported with standard errors in parentheses.

μ_g	σ_ε	ϕ_1	ϕ_2	σ_υ	μ_π	γ	β	σ_ν
0.040 (0.006)	0.031 (0.004)	1.170 (0.207)	-0.759 (0.216)	0.007 (0.004)	1.102 (0.552)	-2.175 (12.10)	89.52 (45.43)	0.631 (2.054)

The estimated value of μ_g suggests an average increase in trend output near 4%, which is consistent with the estimate from the linear trends method. The shocks to trend output are non-negligible, as the estimated value of σ_ε is significantly different from zero. The estimated parameter values for ϕ_1 and ϕ_2 satisfy the stationarity conditions for the autoregressive process driving the cyclical component of output. The low variance of the shocks in this equation suggest the cyclical process is nearly deterministic, although σ_υ is significantly different from zero at the 10% significance level. The estimated positive trend in inflation (μ_π) is not statistically significant. The Phillips curve parameter γ has the wrong sign, but it is not significantly different from zero. In interpreting the parameter estimates for γ and β it is important to note that inflation has been scaled up by a factor of 100. Thus an estimate of β=89 implies that, *ceteris paribus*, a 1% increase in the cyclical component of output leads to a 0.89% increase in inflation the following year.

Figure 5 compares the output gap estimates obtained using the Harvey-Jaeger univariate UC model with those obtained using the Kuttner multivariate UC model. The Kuttner measure is much less erratic, in part because the low variance of the estimated residuals in the equation for the cyclical component. Since the Kuttner measure takes account of price developments, it is more consistent with the surge of inflation in the mid-1970s and the early 1980s. At the end of the sample, the Kuttner measure of the output gap drops from +1.68% of GDP in 2000 to -0.6% in 2001.

Figure 5: Harvey-Jaeger and Kuttner UC measures of output gap



2.4 Apel-Jansson multivariate UC model

The Apel-Jansson (1999) approach to estimating the output gap takes account of Okun's law as well as the Phillips curve. This requires data on unemployment, which in Luxembourg is only available starting in the early 1970's. To implement this approach the sample period was restricted to 1980-2001. Although this means that parameters are estimated with less precision, it avoids the potential structural break implicit in the transformation of Luxembourg's economy away from the steel sector and towards financial services. The last two decades were also dominated by a phenomenon absent in the earlier part of sample, a spectacular rise in the share of employees that commute from outside Luxembourg's borders.

Measuring unemployment in Luxembourg is problematic because one third of employees are non-resident commuters who cannot register as unemployed if they lose their job. Another third of employees is made up of permanently resident immigrants, who may register as unemployed or may leave the country if they lose their job. A regional measure of unemployment was calculated to attempt to account for the fact that so much of the workforce is drawn from the surrounding regions. This is a weighted average of the unemployment rates in Luxembourg and in the immediately surrounding regions within Belgium, France and Germany, where almost all cross-border commuters are resident. The weight attached to each of these unemployment rates reflects the share of Luxembourg's salaried employment that is resident in the respective region[21].

Apel and Jansson (1999) considered two different specifications of the Phillips curve. In the unrestricted specification inflation was mean-reverting, whereas in the other it was restricted to follow a random walk. Both variants were estimated and produced similar results, but since the unit-root restriction was not rejected by the Luxembourg data, only the second specification is reported, with the Phillips curve expressed in terms of changes in inflation.

The Phillips curve used by Apel and Jansson is based on Gordon's triangle model of inflation, integrating inflation inertia, excess demand, and supply shocks. Four supply shocks were included: relative import prices, oil prices, labour productivity, and the real effective exchange rate[22]. All four shocks were expressed in growth rates to ensure stationarity and were normalised on their mean in the sample (so that their impact is zero in the absence of a shock). The estimated Phillips curve specification included two lags of inflation and a homogeneity restriction[23]. The unemployment gap was included as a contemporaneous term and a single lag. The supply shocks were only introduced as contemporaneous terms. The unemployment gap entered the Okun's law relationship as a contemporaneous term and a single lag. Two lags of the unemployment gap were retained in the final state equation. Parameter estimates are reported with standard errors in parentheses.

21 See Guarda (1999) for more details.

22 Relative import prices were calculated by splicing the deflators for imports and GDP as published by AMECO and Statec. Oil prices are Brent crude in LUF per barrel (Eurostat Pocketbook and ECB bulletin). Labour productivity is calculated for total employment (AMECO and Statec series). Luxembourg's real effective exchange rate is taken from the IMF's International Financial Statistics.

23 For a meaningful NAIRU to exist the coefficients on lagged inflation must sum to unity, i.e. $\rho(1)=1$.

η_0	η_1	σ_{pc}	ϕ_0	ϕ_1	σ_{ol}	σ_n	α	σ_p
-3.84 (4.38)	1.11 (4.45)	0.62 (0.46)	-9.90 (8.89)	6.81 (6.17)	1.18 (1.46)	0.33 (0.15)	5.23 (0.87)	1.65 (1.12)

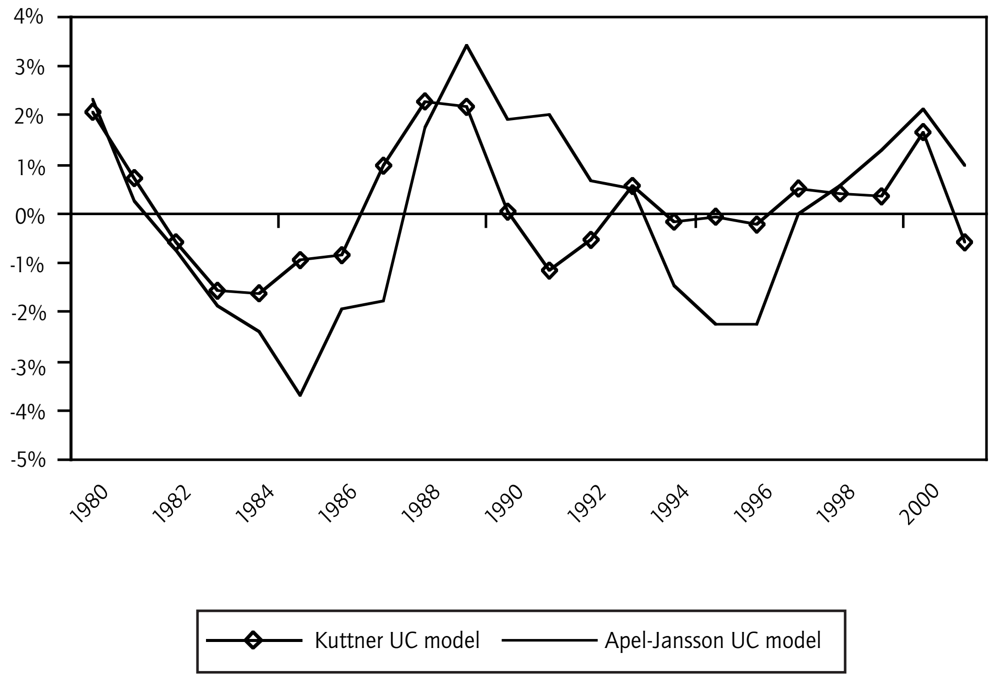
In accord with economic theory, the sum of the coefficients on cyclical unemployment is negative in both the Phillips curve and the Okun's law relationship. Since cyclical unemployment enters the Phillips curve as both a contemporaneous and a lagged term, it will have both level and change effects on inflation. The level effect is captured by the sum of the coefficients and the change (or "speed-limit") effects by the individual coefficients themselves.

To interpret the coefficients correctly, one must note that both inflation and unemployment rates have been scaled up by 100. In the Phillips curve $\eta_1+\eta_2$ = -2.73, suggesting that if the unemployment gap is 1 percentage point higher over a period of two years, inflation will drop 2.73 percentage points. In the Okun's law relationship $\phi_1+\phi_2$ = -3.08, suggesting that an increase in cyclical unemployment by 1 percentage point over two years will lead to a fall in the output gap of more than 3 percentage points. These results may be exaggerated by coefficient uncertainty. In any event, using Swedish data Apel and Jansson found that level effects were much less substantial than change effects.

The estimated value of α suggests a trend rate of potential output growth of 5.23%, which is 1% higher than the estimate obtained using the Kuttner approach. This may reflect the fact that the sample has been restricted to a period of higher growth. The high estimated value of the σ_p parameter, representing the standard error of the innovations in potential output, contrasts with the results found by Apel and Jansson. It suggests that a deterministic trend provides an unsatisfactory description of potential output growth in Luxembourg over this sample. The innovations in the NAIRU are less volatile and the standard deviation of innovations in cyclical unemployment (σ_c, not reported) is practically zero.

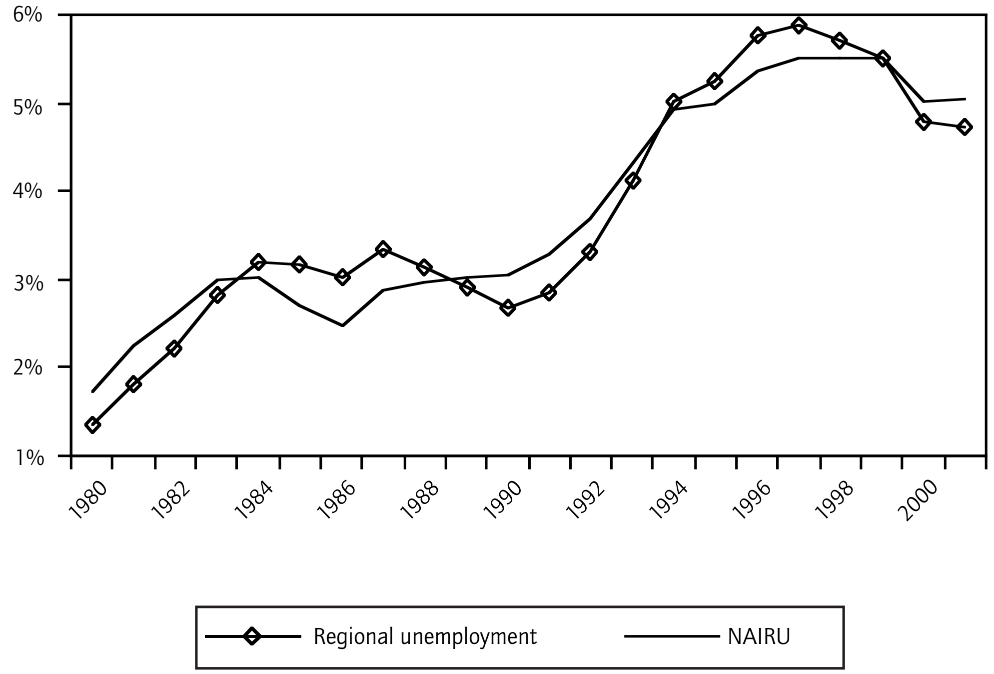
The following figure compares the evolution of the output gap measured according to Kuttner's UC method and the Apel-Jansson UC approach. The two curves are strikingly similar over the first half of the sample but behave rather differently thereafter, with the Kuttner output gap measure displaying more volatility. This difference not only reflects the significant stochastic element estimated in potential output under the Apel-Jansson approach, but also the additional shocks feeding into the system of equations due to movements in the unemployment data. The Apel-Jansson output gap measure appears to lag at turning points and sign changes, possibly reflecting slow adjustment in the labour market that the Kuttner measure does not take into account. The Apel-Jansson output gap also suggests greater cyclical variation in the 1990's, with a trough in 1996 at which output is 2.6% below potential (the Kuttner measure is near zero over this period). At the end of the sample, the Apel-Jansson output gap measure is significantly positive in 2000 and indicates that output is still 1.8% above potential in 2001, whereas at this point in the sample the Kuttner measure indicates an output gap that is –0.6%.

Figure 6: Kuttner and Apel-Jansson UC measures of output gap



The Apel-Jansson approach estimates the output gap simultaneously with the NAIRU. This other latent variable is secondary in the present context but is presented in figure 7 to illustrate that its evolution is complementary to that of the output gap and to better understand developments in the Luxembourg economy over the sample period. The NAIRU is plotted along with the "regional" unemployment rate used in estimation. This was constructed by aggregating unemployment rates across different regions. The regional unemployment series increases markedly over the sample, reflecting the rapidly growing number of cross-border commuters. As more non-residents come to work in Luxembourg, they increase the weight attached to non-resident unemployed in the "regional" measure of unemployment for the Luxembourg labour market. This process reflects the increasing dependence of Luxembourg's economy on the cross-border labour force and the greater integration of the regional labour market. The official unemployment figures (which are limited to residents only) were much lower and more stable throughout the period.

Figure 7: Regional unemployment and Apel-Jansson UC measure of NAIRU



The estimated NAIRU follows the increase in the "regional" unemployment measure fairly closely. In other European economies, such tracking behaviour has been attributed to hysteresis effects, although it is not clear whether they are relevant in the present context. The "regional" unemployment rate is below the NAIRU at the beginning of the sample, when inflation is high following the second oil price shock. During the latter half of the 1980s, when tighter monetary policy drove down inflation across Europe, the regional unemployment rate is above the estimated NAIRU. However, by the end of the decade the two cross again, just as the estimated output gap becomes positive. Regional unemployment is below the estimated NAIRU during the early 1990s, when the economy is operating above potential and inflation rises briefly. By 1994 both the output gap and the unemployment gap have closed. The regional unemployment rate is above the NAIRU and inflation declines until the end of the decade. In 2000, a positive output gap and an estimated NAIRU once again above the regional unemployment rate reflect the recent surge in inflation.

2.5 Production function approach

Finally, the production function approach was also implemented to estimate the output gap in Luxembourg. This was based on a constant-returns-to-scale Cobb-Douglas production function with labour-augmenting Harrod-neutral technical progress. Letting Y_t denote real output, K_t the capital stock at constant prices and L_t the level of employment, this can be written as follows

$$Y_t = A\, K_t^{\beta} L_t^{(1-\beta)} e^{(1-\beta)\gamma t}$$

Direct single equation estimation of production function parameters typically gives implausible results, because the inputs are chosen in some optimal fashion by producers so the exogeneity assumptions required for ordinary least squares will not hold. This was effectively also the case using Luxembourg data, with a labour elasticity of output far below the labour share of gross value added. Therefore, the parameters of the production function were recovered from a system of simultaneous equations based on the first order conditions. Starting from the assumption of competitive factor markets and imperfect competition in the product market, profit maximisation yields a supply-side system of three equations: a price equation, a demand for labour equation, and an investment equation[24].

$$\log(P_t) = \log(\varepsilon) - [\log(1-\beta)+\log(A)/(1-\beta)]+\log(W_t) + \beta/(1-\beta)\log(Y_t/K_t) - \gamma t$$

$$\log(L_t/Y_t) = -\beta\log(K_t/L_t) - \log(A) - (1-\beta)\gamma t$$

$$\log(CC_t) = \log(\beta/(1-\beta)) + \log(W_t) - \log(K_t/L_t)$$

where ε in the price equation is the mark-up (equal to unity under perfect competition) and CC_t is the user cost of capital (calculated as an average of long and short-term interest rates, and accounting for the constant 5% rate of depreciation assumed in constructing the capital stock series by the perpetual inventory method). The three equations were estimated by the seemingly unrelated regression procedure with the cross-equation restrictions on the parameters imposed. For the purposes of estimating the output gap, labour input was later decomposed into resident and non-resident labour. Since data on cross-border labour flows is only available since 1980, estimation was for the more restricted sample period used also under the Apel-Jansson method.

ε	β	A	γ
1.002 (0.043)	0.330 (0.012)	7.865 (0.515)	0.036 (0.002)

According to the standard deviations in parantheses, the parameters are all significant, although there is evidence of autocorrelation in the residuals of the two factor demand equations. The mark-up parameter ε is surprisingly near unity, suggesting that Luxembourg firms were mostly price-takers over this period (although this result may be due to model mis-specification). The parameter β is twice as high as under the direct estimation method, and is consistent with the observed labour share in value added, which averaged 0.68 over the

24 See Slevin (2001) sections 4a and 4b for details.

period. The parameter γ suggests a 3.6% average annual rate of growth of technological progress, which is consistent with estimates found for other countries.

To estimate potential output and the output gap, labour input was decomposed into resident and non-resident labour. Resident labour was expressed as the product of the population in working age (15-59 years old), a participation rate, and an employment rate (one minus the unemployment rate). The unemployment rate was calculated using resident employment and the DENS series (*demandes d'emploi non-satisfaites*) spliced over the change of definition in 1997. The HP filter was then used to extract a trend unemployment rate, a trend participation rate and the trend in the number of non-resident employees. These were combined to calculate a trend level of labour input. The deterministic trend in technical progress was then replaced by calculating the Solow residual and using the HP filter to estimate the trend growth rate in Total Factor Productivity (TFP). Potential output was calculated by combining the level of capital input, the trend level of labour input and trend TFP using the estimated parameters of the production function.

Figure 8: Apel-Jansson and Production Function measures of output gap

Apart from the beginning of the sample, the Production Function approach yields a measure of the output gap that is remarkably close to that of the Apel-Jansson UC model. This is all the more surprising as the Production Function approach considered only resident unemployment while the Apel-Jansson method was implemented using a regional measure of unemployment. However, the similarity of the two output gap measures is encouraging in so far as it suggests that the different methods are measuring essentially the same concept.

3. Comparing and Evaluating Different Output Gap Measures

In the previous section, the different measures of the output gap were only compared graphically, giving, at best, a subjective indication of how similar or dissimilar they are. This section provides a more formal basis of comparison for the six measures considered (linear trend, HP filter, Harvey-Jaeger univariate UC model, Kuttner multivariate UC model, Apel-Jansson multivariate UC model and Production Function approach). The comparison is limited to the period 1980-2001. To give the linear trend approach the benefit of the doubt, this particular output gap measure was re-estimated on the shorter sample period to allow for the possibility that the constant rate of trend output growth over this period changed from that in the 1960s and 1970s. The different output gap measures are also assessed in terms of their inflation forecasting performance and through pair-wise encompassing tests performed within the "triangle" model of inflation.

Table 3: Comparing output gap measures - descriptive statistics

	linear trend	HP filter	Harvey-Jaeger	Kuttner	Apel-Jansson	Production Function
mean	0.00	0.04	0.22	-0.02	-0.02	-0.82
mean absolute value	2.30	2.03	2.50	0.86	1.65	2.17
standard deviation	3.25	2.82	3.37	1.13	1.94	2.54
σ^2_i/σ^2_j distributed F(21,21) under H_0: equal variances						
Linear trend						
HP filter	1.32					
Harvey-Jaeger	0.93	0.70				
Kuttner	8.23	6.21	8.85			
Apel-Jansson	2.80	2.11	3.01	0.34		
Production Function	1.64	1.24	1.76	0.20	0.59	

The upper part of Table 3 reports the mean, mean absolute value and standard deviation of the different output gap measures. For the linear trend method the mean is naturally zero as the output gap is measured simply by OLS residuals over the sample. For the other measures the mean is also close to zero and in all but one case is less than 0.25% in absolute value, suggesting that the sample period covers two whole cycles. The mean absolute value of the gap measures provides a better indication of how far on average actual output is from its potential level. This is close to 2% for five of the six measures, with the Harvey-Jaeger univariate UC model leading to slightly larger deviations on average and the Kuttner multivariate UC model leading to smaller deviations closer to 1% on average. The standard deviation of the output gap gives an indication of the volatility of the business cycle. The Harvey-Jaeger univariate UC model leads to the most volatile output gap of the six and the Kuttner multivariate UC model to the least volatile measure of the output gap.

The lower part of Table 3 reports the ratios of the squares of the standard deviations of the different measures of the output gap. Under the null hypothesis of equal variances, this statistic follows an *F* distribution with degrees of freedom equal to the number of observations (21 for each output gap measure). Asymptotically, this ratio must exceed 2.08 or be less than 0.48 to reject the null hypothesis of equal variances at the 5% significance level. The variance of the output gap measure on the Kuttner multivariate UC method is significantly lower than that of all the other methods. Although the Harvey-Jaeger univariate UC method produced the output gap measure with the highest variance, the test cannot reject equality of variance with the output gap obtained with the HP filter, linear trends or the production function approach.

Table 4: Comparing output gap measures - correlation coefficients and sign test

	Linear trend	HP filter	Harvey-Jaeger	Kuttner	Apel-Jansson
linear correlation coefficients					
Linear trend					
HP filter	0.61				
Harvey-Jaeger	0.05	0.67			
Kuttner	0.43	0.54	0.27		
Apel-Jansson	0.70	0.75	0.26	0.59	
Production F.	0.56	0.84	0.50	0.49	0.81
Pesaran-Timmermann test distributed N(0,1) under H_0: signs independent					
linear trend					
HP filter	2.69				
Harvey-Jaeger	-0.89	1.9			
Kuttner	2.24	2.57	-0.32		
Apel-Jansson	2.71	4.23	-0.61	3.72	
Production F.	2.15	3.65	0.32	2.77	3.71

Table 4 compares the different output gap measures in terms of their pair-wise linear correlation and also reports the Pesaran-Timmermann (1992) predictive accuracy test of directional change. The latter tests whether the signs on two variables are independent. Rejecting this hypothesis for two different measures of the output gap would indicate that they systematically carry the same sign. This would make it seem likely that they are effectively measuring the same output gap although they may differ in size and occasionally give conflicting policy messages.

Using the *t* distribution with 22 degrees of freedom, the linear correlation coefficient must exceed 0.38 to reject the null hypothesis of no correlation at the 5% level. This is the case for most of the pair-wise combinations. Only the output gap on the Harvey-Jaeger method is not significantly correlated with three of the five alternative measures. It *is* significantly correlated with the output gap calculated by the HP filter, which is to be expected, and also with the output gap according to the Production Function approach. The Production Function measure yields the two highest linear correlation coefficients, respectively with the HP filter and the Apel-Jansson measures of the output gap. The linear correlation coefficient is lowest between the linear trend measure and the Harvey-Jaeger measure of the output gap.

Under the null hypothesis that the signs of the two series are independent, the Pesaran-Timmermann test follows a standard normal distribution. The results indicate that the test can reject the null hypothesis for most of the pairwise comparisons at all the conventional levels of significance. All the exceptions involve the Harvey-Jaeger measure of the output gap, whose sign only seems to be systematically related to that of the HP filter (again, not surprisingly).

Table 5: Dynamic Correlation Analysis (output gap and inflation)

	lag/lead on inflation						
	-3	-2	-1	0	1	2	3
Linear trend	-0.53	-0.44	-0.13	0.28	0.49	0.57	0.49
HP filter	-0.24	-0.22	-0.12	0.08	0.25	0.36	0.33
Harvey-Jeager	-0.03	-0.14	-0.25	-0.26	-0.29	-0.21	-0.04
Kuttner	-0.32	-0.63	-0.61	-0.15	0.15	0.31	0.25
Apel-Jansson	-0.76	-0.74	-0.45	-0.02	0.28	0.42	0.38
Production Fn.	-0.48	-0.57	-0.43	-0.30	-0.17	-0.01	0.14

Table 5 reports a dynamic correlation analysis between the estimated measures of the output gap and inflation. Each column represents a lag (negative) or a lead (positive) on inflation. The central column (zero) represents the contemporaneous correlation between the output gap measure and inflation. None of these is statistically significant. Correlation with lagged inflation is generally negative, suggesting that following episodes of high inflation the output gap tends to fall as policy measures are taken in response. Correlation with future inflation is generally positive instead, suggesting that increasing output gaps are followed by inflation. The linear trend output gap measure seems to do exceedingly well as a leading indicator of inflation, but this is likely to be a statistical artefact due to the short sample size. There is little reason to believe that this result would hold out of sample. The Apel-Jansson measure is the next best performer, with the HP filter and the Kuttner measure of the output gap following close behind. This would suggest that, with the exception of the Harvey-Jaeger method and the Production Function approach, most measures of the output gap have some leading indicator properties for inflation.

Table 6: Inflation Forecast Performance using "triangle" Phillips Curve

	linear trend	HP filter	Harvey-Jaeger	Kuttner	Apel-Jansson	Production Function
R^2	0.30	0.29	0.28	0.92	0.44	0.22
F-test (*p*-value)	0.71	0.82	0.91	0.00	0.11	0.89
Theil U-statistic	0.75	0.76	0.76	0.25	0.67	0.78
Diebold-Mariano Test distributed N(0,1) under H_0: equal predictive accuracy						
Linear trend						
HP filter	-0.32					
Harvey-Jaeger	-0.44	-0.30				
Kuttner	3.51	3.26	3.16			
Apel-Jansson	1.71	1.55	1.57	-3.72		
Production Fn.	-1.29	-0.62	0.03	-3.30	-2.06	

The different measures of the output gap were also compared in terms of their inflation forecasting performance. For this purpose, the triangle inflation Phillips curve presented under the Apel-Jansson approach was estimated with each of the output gap measures serving in turn as the excess demand indicator. Full results are not reported, but Table 6 includes the R^2 from these equations as well as the *p*-value associated with the F-test of the hypothesis that the contemporaneous and lagged output gap contribute no explanatory power beyond that already in lagged inflation and the supply shock variables. Theil's U-statistic is also reported. This compares the forecast performance of each model to that of a "naive" method based on the random walk (the change in inflation this period will be the same as the change observed last period). Theil's U-statistic is calculated as the root mean squared error (RMSE) divided by the RMSE of the "naive" forecast. A U-statistic below unity indicates that the forecasting method represents an improvement over the "naive" method. If the U-statistic is above unity, a random walk could forecast just as well. While these statistics are used routinely in forecast comparisons, comparing U-ratios (or R^2 for that matter) does not represent a formal test. For this purpose, the Diebold-Mariano (1995) test of predictive accuracy is also applied to perform pair-wise comparisons of the forecast errors generated using different measures of the output gap. Under the null hypothesis of equal forecast accuracy, this test follows a standard normal distribution.

Of the different output gap measures, the Kuttner method provides by far the highest R^2 in the inflation regression and the Production Function approach the lowest R^2. On the F-test, only the output gap on the Kuttner method has a p-value below 1%, indicating that it contributes explanatory power in forecasting inflation beyond that in the inertia terms or the supply shock. The output gap on the Apel-Jansson method is borderline significant at the 10% level. Theil's U-statistic is below unity for all the output gap measures, but much closer to zero when using the Kuttner estimate of the output gap. The overwhelming superiority of the Kuttner measure is not surprising in this context, as it concentrates on price fluctuations while most of the other methods simply ignore them. It is also not surprising that the Apel-Jansson method comes in second-best as it is the only other method to include a Phillips curve in estimation.

The Diebold-Mariano test is negative when the output gap measure in the row generates weaker forecasts than the measure in the column and vice versa. On this criterion, the Kuttner method is significantly better than all others, and the Apel-Jansson method is significantly better than either the Production Function approach or the linear trends method. For the other pair-wise comparisons the Diebold-Mariano test cannot reject the null hypothesis of equal predictive accuracy at conventional significance levels.

Table 7: Inflation forecasting - Simplification Encompassing Test (SET)

distributed F(2,10) under H_0: $M_i \, \varepsilon \, M_j$						
	Linear trend	HP filter	Harvey-Jaeger	Kuttner	Apel-Jansson	Production F.
Linear trend		0.04	0.00	52.93	3.92	0.51
HP filter	0.14		0.02	56.76	3.42	0.22
Harvey-Jeager	0.20	0.10		45.83	2.59	0.07
Kuttner	1.07	1.37	0.08		0.21	1.22
Apel-Jansson	1.94	1.41	0.68	34.62		1.71
Production Fn.	1.23	0.27	0.02	51.25	4.21	

The Mizon-Richard (1986) Simplification Encompassing Test (SET) was applied to perform pair-wise comparisons of the predictive content of the different output gap measures. This is based on extending the Phillips curve equation estimated for each measure of the output gap by adding the contemporaneous and lagged term of one of the other gap measures. The SET is then calculated as an F-test of the restriction that the coefficients on the additional terms are zero. If this restriction can be rejected then the terms from the alternative output gap measure make a significant additional contribution and the original model cannot encompass forecasts using the alternative output gap measure. Thus failure to reject indicates that the model based on the output gap measure in the row may encompass inflation forecasts based on the output gap in the column.

Note that encompassing is not symmetric, so that two models based on different measures of the output gap may both fail to encompass each other. This would suggest that the two measures are based on different sets of information that contribute to different aspects of inflation forecasts. For an F(2, 10) distribution, the SET must exceed 4.1 at the 5% significance level to reject the null hypothesis that the model in the row can encompass the model in the column. The only significant test statistics are in the columns relating to the multivariate UC models. None of the other models can encompass the model using the Kuttner measure. The Production Function approach is unable to encompass the Apel-Jansson VC model at the 5% level, as are the linear trends and the HP filter methods at the 10% level. Since the other test statistics are not significant, this would suggest that the Kuttner or Apel-Jansson models are capable of encompassing the other models.

4. Conclusions

In summary, the different methods of calculating the output gap give fairly different results, although they are not wildly dissimilar. In particular, the sign of the output gap on the different measures seems to be systematically related, suggesting that the methods are at least measuring a related concept. Most of the output gap measures considered seemed to have some leading indicator properties for inflation. Within the triangle model of inflation, at least two of the output gap measures considered seemed to contribute something to inflation forecasting beyond information contained in lagged inflation or supply shock variables. On several criteria, unobserved components models seemed to perform better than linear trends, the HP filter or the production function approach. Multivariate methods also seemed preferable to univariate methods. In terms of inflation forecasts, the Kuttner method could encompass rival measures but could not be encompassed by them.

The dangers of using some of the simpler methods for calculating output gaps (linear time trends or the HP filter) have been extensively documented. However, there still remain several alternative methods to choose from. These generally have different advantages, including the ability to incorporate structural relationships based on economic theory, to allow for flexible specifications for the dynamics of output, and to calculate confidence intervals for their estimates of the output gap. This last point is important, as any use of output gap estimates for the purpose of policy formulation must take account of a measure of the uncertainty attached to the estimate. Different sources of this uncertainty are discussed below.

4.1 Sources of Uncertainty

In its October 2000 bulletin, the ECB emphasised three sources of uncertainty affecting any estimate of the output gap: model uncertainty, parameter uncertainty, and data uncertainty[25]. *Model uncertainty* is inevitable given the lack of consensus on the appropriate formal framework. It may affect univariate methods more strongly as these usually incorporate more explicit restrictions on the dynamics of output. In the case of multivariate methods, the inclusion of several macroeconomic time series gives the system more flexibility and the restrictions applied are usually drawn from economic theory, giving them a less arbitrary character. To the extent that economic relationships are consistent with observed series, these multivariate approaches are potentially more robust.

The parameters of the model are always unknown and must be estimated. *Parameter uncertainty* is associated with the possibility that the given sample produces parameter estimates that are far from the "true" values. This uncertainty may be partially quantified if the estimation method allows for the calculation of standard deviations for the parameter estimates. However, these depend on a variety of regularity assumptions. Parameter

25 In fact, these sources of uncertainty affect any econometric exercise (see Fair 1986).

uncertainty is severe when the parameter values change over the sample but the model assumes they are fixed. In this case, the conclusions of the analysis can be completely invalidated. The effects of such structural change only become obvious after several periods. However, in many modelling contexts there are some simple tests that can be used to assess the risk of parameter instability over the observed sample.

The final sort of uncertainty is associated with the data. *Data uncertainty* is inevitable as the exact value of most economic series is not known immediately and is subject to a series of revisions as more data becomes available. Output data are particularly prone to significant revisions in the subsequent months or even years, as they are dependent on a host of other variables. Unfortunately, data uncertainty affects observations towards the end of the sample that are often critical in constructing the output gap measures that are of interest for policymaking. The highest degree of data uncertainty is associated with forecasts of the output gap in the future, which require assumptions regarding the future paths of exogenous variables in the model.

The ECB provided an assessment of the "reliability" of euro area output gap estimates based on the HP filter. This evaluated the combined effects of parameter and data uncertainty, by sequentially recalculating the output gap as more observations were added to the sample. Although the "model" used was unchanged (the HP filtering framework), the "parameter estimates" changed with the additional observations of data. The discrepancy between real-time[26] and ex post estimates as measured by the difference between the final and sequential output gap indicated that the sign of the gap in a given period could change when data for later periods became available. This is a sobering exercise that is easy to perform with the HP filter, but which would be valuable with all the methods considered above. Orphanides and van Norden (1999) and Rünstler (2001) perform similar exercises for several different methods of calculating output gaps.

4.2 Future Work

Business cycle analysis is generally conducted on quarterly (or even monthly) data to reflect developments that can take place within the span of a year. Unfortunately, quarterly national accounts are still unavailable in Luxembourg and the present study had to follow the unorthodox route of calculating output gap estimates using annual data. If quarterly national accounts are released soon, they should make it possible to provide a better analysis of the business cycle in Luxembourg.

For those models that incorporated a Phillips curve, inflation was measured here using the index of consumer prices. Apel and Jansson (1999) also considered the private consumption deflator and the GDP deflator as alternative indicators of inflation on a wider basis. This route could also be explored in future work.

26 This was actually "quasi real-time" as a strict definition of real time analysis is based on the data as released before revisions.

Different measures of the output gap were assessed in terms of their contribution to inflation forecasting. It should be noted that the analysis here was only in terms of *in-sample* forecasting. A more rigorous procedure would focus on *out-of-sample* inflation forecasting at different horizons (Ross and Ubide 2001, and Rünstler 2001).

Finally, in the context of the UC models, the present study limited itself to the specifications of the dynamics used in the original articles by Kuttner (1994) and Apel and Jansson (1999). These assumed that potential output followed a random walk with drift and that the NAIRU followed a pure random walk. More recent work by Fabiani and Mestre (2001) and Ross and Ubide (2001) found that other dynamic specifications were more suited to the euro area. For example, it would be straightforward to allow for a nonzero drift in the NAIRU.

In conclusion, UC models provide a fruitful framework for estimating output gaps, combining flexible specifications with economic theory. In particular, these models can quantify the uncertainty surrounding estimates and can provide out-of-sample forecasts. Their inflation forecasting performance is superior to that of simpler methods such as linear trends or the HP filter. It remains to be established whether the relative performance of different methods remains unchanged once data at quarterly frequency is available for Luxembourg.

Bibliography

Adams, C. and D.T. Coe (1990) "A Systems Approach to Estimating the Natural Rate of Unemployment and Potential Output for the United States," *International Monetary Fund Staff Papers*, 37(2):232-93.

Apel, M. and P. Jansson (1999a) "System Estimates of Potential Output and the NAIRU," *Empirical Economics*, 24:373-388.

Apel, M. and P. Jansson (1999b) "A Theory-Consistent System Approach for Estimating Potential Output and the NAIRU," *Economics Letters*, 64:271-275.

Baxter, M. and R.G. King (1995) "Measuring Business Cycles: approximate band-pass filters for economic time series," National Bureau of Economic Research Working Paper 5022.

Beveridge, S. and C.R. Nelson (1981) "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle," *Journal of Monetary Economics*, 7(2):151-74.

Blanchard, O. and D. Quah (1989) "The Dynamic Effects of Demand and Supply Disturbances," *American Economic Review*, 79(4):655-73.

Bolt, W. and P.J.A. van Els (2000) "Output Gap and Inflation in the EU" De Nederlandsche Bank Staff Report 44.

Botas, S., C.R. Marques and P.D. Neves (1998) "Estimation of the Potential Output for the Portuguese Economy," *Banco de Portugal Quarterly Bulletin*, December.

Brouwer, G. de (1998) "Estimating Output Gaps," Reserve Bank of Australia Research Discussion Paper 9809

Butler, L. (1996) "A Semi-Structural Method to Estimate Potential Output: Combining Economic Theory with a Time-Series Filter," Bank of Canada Technical Report 77.

Camba-Mendez, G. and D. Rodriguez-Palenzuela (2001) "Assessment Criteria for Output Gap Estimates," European Central Bank Working Paper 54.

Canova, F. (1993) "Detrending and Business Cycle Facts," *Journal of Monetary Economics*, 41(3):475-512.

Canova, F. (1995a) "The Economics of VAR Models," chapter 3 in K.D. Hoover (ed.) *Macroeconometrics: Developments, Tensions and Prospects*, Boston: Kluwer Academic Publishers.

Canova, F. (1995b) "Vector Autoregressive Models: Specification, Estimation, Inference and Forecasting," chapter 2 in M.H. Pesaran and M. Wickens (eds.) *Handbook of Applied Econometrics: Macroeconomics*, Oxford: Blackwell Publishers.

Campbell, J.Y. and N.G. Mankiw (1987) "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics*, 102(4):857-80.

Cerra, V. and S. C. Saxena (2000) "Alternative Methods of Estimating Potential Output and the Output Gap: An Application to Sweden," International Monetary Fund Working Paper 00/59.

Claus, I. (2000) "Estimating Potential Output: A Structural VAR Approach," Reserve Bank of New Zealand Discussion Paper 00/3.

Clark, P.K. (1987) "The Cyclical Component of U.S. Economic Activity," *Quarterly Journal of Economics*, 102(4):797-814.

Cochrane, J.H. (1994) "Permanent and Transitory Components of GNP and Stock Prices," *Quarterly Journal of Economics*, 109(1):421-465.

Cogley, T. and J. Nason (1995) "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series," *Journal of Economic Dynamics and Control*, 19:253-78.

Conway, P. and B. Hunt (1997) "Estimating Potential Output: A Semi-Structural Approach," Reserve Bank of New Zealand Discussion Paper 97/9.

Cooley, T.F. and M. Dwyer (1998) "Business Cycle Analysis Without Much Theory: A Look at Structural VARs," *Journal of Econometrics*, 83:57-88.

De Masi, P.R. (1997) "IMF Estimates of Potential Output: Theory and Practice," International Monetary Fund Working Paper 97/177.

Diebold, F.X. and R.S. Mariano (1995) "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13:253-63.

Dimitz, M.-A. (2001) "Output Gaps and Technological Progress in European Monetary Union," Bank of Finland Discussion Paper 20/01.

Dupasquier, C., A. Guy and P. St.-Amant (1999) "A Survey of Alternative Methodologies for Estimating Potential Output and the Output Gap," *Journal of Macroeconomics*, 21(3):577-595.

ECB (2000) "Potential Output Growth and Output Gaps: Concept, Uses and Estimates," European Central Bank Monthly Bulletin, October.

Engle, R.F. (1978) "Estimating Structural Models of Seasonality," in Zellner, A. (ed.) *Seasonal Analysis of Economic Time Series*, Washington, DC: Department of Commerce, Bureau of the Census, reprinted as ch. 11 in Hylleberg, S. (ed.) (1992) *Modelling Seasonality*, Oxford, Oxford University Press.

Evans, G. and L. Reichlin (1994) "Information, Forecasts and Measurement of the Business Cycle," *Journal of Monetary Economics*, 33(2):233-54.

Fair, R.C. (1986) "Evaluating the Predictive Accuracy of Models," chapter 33 in Z. Griliches and M.D. Intriligator (eds.) *Handbook of Econometrics*, vol. III, Elsevier Science.

Fabiani, S. and R. Mestre (2001) "A System Approach for Measuring the Euro Area NAIRU," European Central Bank Working Paper 65.

Fisher, P.G., L. Mahadeva and J.D. Whitley (1997) "The Output Gap and Inflation: Experience at the Bank of England," in *Monetary Policy and the Inflation Process*, Bank for International Settlements Conference Papers, vol. 4.

Gerlach, S. and F. Smets (1997) "Output Gaps and Inflation: Unobservable Component Estimates for the G7 Economies," unpublished manuscript, Bank for International Settlements.

Gerlach, S. and F. Smets (1999) "Output Gaps and Monetary Policy in the EMU Area," *European Economic Review*, 43(4-6):801-812.

Giorno, C., P. Richardson, D. Rosevaere and P. van den Noord (1995) "Estimating Potential Output, Output Gaps and Structural Budget Balances," Organisation for Economic Co-operation and Development Economics Department Working Paper 152.

Gordon, R.J. (1997) "The Time-Varying NAIRU and its Implications for Economic Policy," *Journal of Economic Perspectives*, 11(1):11-32.

Guarda, P. (1999) "Wages, Prices and Employment: the Luxembourg Supply Side" *Cahiers d'économie du Centre Universitaire de Luxembourg*, fascicule XIV, p. 29-62.

Guay, A. and P. St-Amant (1996) "Do Mechanical Filters Provide a Good Approximation of Business Cycles?" Bank of Canada Technical Report 78.

Hall, S.G and N.G. Zonzilos (1997), "The Output Gap and Inflation in Greece," *Bank of Greece Economic Bulletin*, 9.

Hamilton, J.D. (1994) "State-Space Models," chapter 50 in R.F. Engle and D.L. McFadden (eds.) *Handbook of Econometrics*, vol. IV, Elsevier Science.

Harvey, A.C. (1985) "Trends and Cycles in Macroeconomic Time Series," *Journal of Business and Economic Statistics*, 3(3):216-27.

Harvey, A.C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

Harvey, A.C. and A. Jaeger (1993) "Detrending, Stylized Facts and the Business Cycle," *Journal of Applied Econometrics*, 8(3):231-47.

Harvey, A.C. and P. H. J. Todd (1983) "Forecasting Economic Times Series with Structural and Box-Jenkins Models: A Case Study," *Journal of Business and Economic Statistics*, 1:299-306.

Hodrick, R.J. and E.C. Prescott (1997) "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29(1):1-16.

Kenny, G. (1995) "Some Estimates of Potential Output and the Output Gap for Ireland," Central Bank of Ireland Technical Paper 5/RT/95.

Kaiser, R. and A. Maravall (2001) *Measuring Business Cycles in Economic Statistics*, Lecture Notes in Statistics 154, New York: Springer-Verlag.

Kichian, M. (1999) "Measuring Potential Output within a State-Space Framework," Bank of Canada Working Paper 99-9.

King, G.K., G.H. Plosser, J.H. Stock and M.W. Watson (1991) "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81(4):819-40.

King, G.K. and S. Rebelo (1993) "Low Frequency Filtering and Real Business Cycles," *Journal of Economic Dynamics and Control*, 17:207-33.

Kuttner, K.N. (1994) "Estimating Potential Output as a Latent Variable," *Journal of Business and Economic Statistics*, 12(3):361-367.

Lalonde, R., J. Page and P. St-Amant (1998) "Une Nouvelle Méthode d'Estimation de l'Écart de Production et son application aux États-Unis, au Canada et à l'Allemagne," Bank of Canada Working Paper 98-21.

Laxton, D. and R. Tetlow (1992) "A Simple Multivariate Filter of the Measurement of Potential Output," Bank of Canada Technical Report 59.

Layard, R., S. Nickell and R. Jackman (1991) *Unemployment, Macroeconomic Performance and the Labour Market*, Oxford: Oxford University Press.

Lippi, M. and L. Reichlin (1994) "Diffusion of Technical Change and the Decomposition of Output into Trend and Cycle," *Review of Economic Studies*, 61:19-30

Maravall, A. (1995) "Unobserved Components in Economic Time Series," chapter 1 in M.H. Pesaran and M. Wickens (eds.) *Handbook of Applied Econometrics: Macroeconomics*, Oxford: Blackwell Publishers.

McMorrow, K. and W. Röger (2001) "Potential Output: Measurement Methods, New Economy Influences and Scenarios for 2001-2010," *European Commission DG2 Economics Papers no. 150*.

Mizon, G.E. and J.-F. Richard (1986) "The Encompassing Principle and its Application to Testing Non-nested Hypotheses," *Econometrica*, 54:657-78.

Nelson, C.R. and H. Kang (1981) "Spurious Periodicity in Inappropriately Detrended Time Series," *Econometrica*, 49(3):741-51.

Nelson, C.R. and C. Plosser (1982) "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 10:139-67.

Ongena, H. and W. Röger (1997) "Les Estimations de l'Écart de Production de la Commission Européenne," *Économie Internationale*, 69(1):77-96.

Orphanides, A. and S. van Norden (1999) "The Reliability of Output Gap Estimates in Real Time," Board of Governors of the Federal Reserve System, Finance and Economics Discussion Paper 99/38.

Pesaran, M.H. and A. Timmermann (1992) "A Simple Nonparametric Test of Predictive Performance," *Journal of Business and Economic Statistics*, 10(4):461-65.

Pinheiro, M. (1998) "Estimation of the Output Gap: a Bivariate Approach," *Banco de Portugal Economic Bulletin*, December.

Quah, D. (1992) "The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds," *Econometrica*, 60(1):107-18.

Rasi, C-M. and J-M. Viikari (1998) "The Time-Varying NAIRU and Potential Output in Finland," Bank of Finland Discussion Paper 6/98.

Ravn, M.O. and H. Uhlig (2002) "On Adjusting the Hodrick-Prescott Filter for Frequency of Observations," *Review of Economics and Statistics*, 84(1):371-380.

Ross, K. and A. Ubide (2001) "Mind the Gap: What is the Best Measure of Slack in the Euro Area?" International Monetary Fund Working Paper 01/203.

Rünstler, G. (2001) "Are Real Time Estimates of the Output Gap Reliable? An Application to the Euro Area," presented at Banca d'Italia/CEPR Conference on Monitoring the Euro Area Business Cycle, Rome, September.

Shapiro, M.D. and M.W. Watson (1988) "Sources of Business Cycle Fluctuations," National Bureau of Economic Research Working Paper 2589.

Sims, C. (1980) "Macroeconomics and Reality," *Econometrica*, 48(1):1-48.

Slevin, G. (2001) "Potential Output and the Output Gap in Ireland," Central Bank of Ireland Technical Working Paper RT/5/2001.

Stock, J.H. and M.W. Watson (1988a) "Testing for Common Trends," *Journal of the American Statistical Association*, 83(404):1097-1107.

Stock, J.H. and M.W. Watson (1988b) "Variable Trends in Economic Time Series," *Journal of Economic Perspectives*, 2(3):147-74.

St-Amant, P. and S. van Norden (1997) "Measurement of the Output Gap: A Discussion of Recent Research at the Bank of Canada," Bank of Canada Technical Report 79, also in *Monetary Policy and the Inflation Process*, Bank for International Settlements Conference Papers, vol. 4.

Watson, M.W. (1986) "Univariate Detrending Methods with Stochastic Trends," *Journal of Monetary Economics*, 18(1):49-75.

7. Annex: estimated output gaps

	GDP	Linear trend	HP filter	Harvey-Jaeger	Kuttner	Apel-Jansson	Production function
1980	6558.64	9.78	0.24	-0.39	2.05	2.36	NA
1981	6522.55	3.84	-2.69	-4.08	0.71	0.23	-2.71
1982	6596.21	-0.42	-4.28	-4.67	-0.58	-0.75	-4.34
1983	6793.36	-2.87	-4.54	-2.24	-1.56	-1.89	-4.23
1984	7213.7	-2.25	-2.32	3.49	-1.60	-2.40	-2.02
1985	7424.36	-4.76	-3.59	1.41	-0.92	-3.70	-2.85
1986	7999.62	-2.68	-0.82	4.40	-0.82	-1.91	-0.04
1987	8186.21	-5.77	-3.48	-2.35	0.99	-1.76	-3.14
1988	9036.7	-1.27	1.01	6.14	2.29	1.76	0.96
1989	9926.48	2.73	4.97	8.78	2.19	3.42	4.45
1990	10140.8	-0.52	1.35	1.76	0.06	1.94	1.70
1991	10764.2	0.06	1.63	3.91	-1.14	2.02	0.77
1992	11244.4	-0.96	0.26	0.78	-0.54	0.68	-1.11
1993	12224.3	2.01	2.95	4.88	0.54	0.51	0.47
1994	12737.5	0.73	1.34	1.88	-0.18	-1.48	-1.28
1995	13220.2	-0.94	-0.62	0.69	-0.05	-2.26	-3.00
1996	13691.1	-2.82	-2.77	-2.32	-0.24	-2.27	-4.63
1997	14925.3	0.42	0.11	0.08	0.53	0.00	-1.36
1998	15794.3	0.69	0.04	-2.96	0.39	0.55	-0.65
1999	16736.2	1.10	0.09	-1.78	0.34	1.30	0.10
2000	17990.8	2.94	1.58	-0.01	1.68	2.13	3.01
2001	18613.5	0.95	-0.76	-5.20	-0.60	1.00	2.65