# IMF Working Paper

# Ghostbusting:
# Which Output Gap Measure
# Really Matters?

*Andreas Billmeier*

INTERNATIONAL MONETARY FUND

**IMF Working Paper**

Middle East and Central Asia Department

**Ghostbusting: Which Output Gap Measure Really Matters?**

Prepared by Andreas Billmeier[1]

Authorized for distribution by Paulo Neuhaus

August 2004

**Abstract**

**This Working Paper should not be reported as representing the views of the IMF.**
The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

This paper investigates various output gap measures in a simple inflation forecasting framework. Reflecting the cyclical position of an economy, an (unobservable) output gap has important implications for economic analysis. I construct and compare common output gap measures for five European countries. Since output above potential reflects domestic inflationary pressures, including a gap could improve the accuracy of autoregressive inflation forecasting. This assertion is tested in a simple simulated out-of-sample forecasting exercise for the period 1990–2002. The main conclusions are that an output gap rarely provides useful information and that there is no single best output gap measure across countries.

Contents                                                                Page

The quantification of potential output—and the accompanying measure of the 'gap' between actual and potential—is at best an uncertain estimate and not a firm, precise measure.
(Arthur M. Okun, 1962)

# I. INTRODUCTION

The output gap—which measures the deviation of GDP from its potential—is a frequently used indicator of the cyclical position and the degree of slack in the economy. Information on the cyclical position is important for a number of analytical reasons. First, variations in output, when assessed relative to potential, have distinct implications for inflationary pressures in the economy. Consequently, assessing the output gap is pivotal in the discussion of monetary policy, such as in the Taylor rule or in the inflation targeting (IT) framework. Second, the cyclical position, as expressed by the size and sign of the output gap, is an important component of calculating the "structural fiscal balance," which aims to gauge the thrust of fiscal policy. Third, the magnitude of the output gap is relevant for assessing economic growth—that is, it helps to evaluate whether variations in actual growth can be attributed to cyclical factors (such as slow growth in trading partner economies) or to a longer-term change in potential growth.

Defined as the difference between actual and unobservable potential output, the output gap is itself an unobserved variable—in a sense, a "ghost" in the equation. In this paper, various measures of the output gap—stemming from different methodological approaches to determining potential output—are first described for a small sample of European countries comprising Finland, France, Greece, Italy, and the United Kingdom.[2] Second, the informational content of these gap measures is evaluated econometrically in an inflation forecasting framework, and a few policy conclusions are drawn.

Descriptively, the mean of the gap measure should be close to zero over longer time horizons. An extended period—say, more than 15 years—of expansion or contraction would run counter to the concept of the business cycle per se. Moreover, the magnitude of the gap should be bracketed in a "reasonable" range, simply because actual GDP at levels too distant from potential are intuitively implausible. Even in a crisis situation, potential output is not likely to grow significantly while actual output contracts sharply. Finally, the measure should capture a number of stylized facts, in line with traditional descriptions of economic activity in each country. In Finland, for example, the

---

[2]This sample of countries represents a relatively heterogenous set of small and large European economies, including the European G-7 economies with the exception of Germany due to data problems.

measure should reflect the low-inflation boom in the late 1980s and the subsequent overheating; the near-collapse of economic activity during the crisis period 1990–93; and, again, the period of strong economic growth during the late 1990s, driven at first by the economic recovery and later by the information and communication technology (ICT) boom. In Italy and the United Kingdom, a major recession preceded the exit from the Exchange Rate Mechanism (ERM) in the early 1990s. While most approaches—albeit to varying degrees—reproduce these stylized facts, the uncertainty stemming chiefly from the volatility in the real economy is reflected, for 2002, in positive as well as negative estimates of the output gap, depending on the extraction technique. This is the case for Finland and the other countries in the sample.

Econometric evaluation is carried out by a simulated out-of-sample forecasting methodology in a Phillips curve framework, which relates the inflation rate to the output gap. If the gap is, in fact, a measure of domestic inflationary pressures, then a naïve inflation forecast (of no change) could be improved by taking into account the information stemming from the output gap. For the forecast period 1990–2002, models are estimated recursively with data up to the forecast period, and the one-year-ahead predicted inflation is compared with actual inflation.[3] Based on this evaluation, a number of conclusions can be drawn for the sample countries. In Finland and France, univariate inflation forecasts cannot be improved upon by adding any measure of the output gap. For Greece, including the output gap as measured by the production function approach yields a consistently better inflation forecast. For Italy, various output gap measures provide useful information in the inflation forecast exercise. Finally, the United Kingdom seems to be the only country that benefits occasionally from the inclusion of the most common output gap measure, based on the Hodrick-Prescott filter (in its "real-time" version). On a more abstract level, this implies that the output gap is not always a useful measure to gauge domestic inflationary pressures, and that no specific measure in this sample consistently dominates all other measures.

From a policy perspective, the output gap plays a prominent role in monetary and fiscal policy, but at a somewhat different policy horizon. For fiscal policy, the output gap is an important building block for the construction of a "structural" balance, mainly at an annual horizon. By contrast, for monetary policy, the analysis of output gaps using quarterly (or even higher frequency) data would be more useful considering the setup of monetary policy decisionmaking. All major monetary policy regimes rely—at least implicitly—on some form of the output gap, and the monetary stance is reviewed, for example, monthly by the European Central Bank and the Bank of England, and eight times per year by the Fed's Federal Open Market Committee. Employing quarterly data often comes at an overlooked cost, however, in that data revisions, in particular to quarterly GDP data, are common and sometimes substantial.[4] To minimize the impact of data revisions,

---

[3]The evaluation period, 1990-2002, was chosen to investigate an economically rather interesting period in Europe: Finland experienced a major crisis after the burst of an asset bubble, and Italy and the United Kingdom exited from the ERM, see above.

[4]See Orphanides and van Norden (2002). Another problem with intra-year data is seasonality and how to eliminate it. See Appendix I.A for a short discussion.

this paper focuses on annual observations.[5] Consequently, the results reported below will be more relevant in the context of fiscal—rather than monetary—policy.

The remainder of this paper is structured as follows. Section II reviews the vast literature on output gaps and anchors the gap in the policy discussion. Section III and Appendix I propose the various output gap methods, whereas Section IV compares the resulting gap measures descriptively and provides the empirical evaluation. Section V concludes.

## II.  A Look at the Literature

General research on the output gap probably started with Okun (1962) and has been abundant ever since.[6] Roughly speaking, two broad approaches have been followed in the literature to estimate potential output and the output gap, a statistical and a more model-oriented approach. The paper will consider two gaps from each category. While this selection of output gap measures is by no means exhaustive, it brings together some of the better-known approaches and—given the results—suffices for the purpose of this paper.[7]

The first set of gap measures is based on the statistical properties of the underlying GDP series. This approach defines potential output as coinciding with the underlying trend of actual output; estimating the output gap amounts to separating longer-run changes in the trend from short-lived temporary movements around potential. Two univariate statistical measures are used below to identify potential output. They derive from the well-known contribution by Hodrick and Prescott (1997), as well as from Corbae and Ouliaris (2002). The latter authors use frequency domain methods to extract information on the business cycle (and underlying trend) properties of GDP.

The second approach, instead, estimates potential output on the basis of an economic model. This approach views business cycle swings and the gap between actual and potential output as the outcome of demand-determined actual output fluctuating around a slowly-moving level of aggregate supply. The corresponding measure of the output gap should account for underemployed resources, in particular in the labor market. The model-based approaches applied below relate to Blanchard and Quah (1989), and to the large strand of literature which focuses on the production function approach to potential output. Contributions to the latter include early research done at the IMF, such as Artus (1977) and, more recently, De Masi (1997). The application of the production function approach follows closely the latent variable approach developed in Kuttner (1994), and further refined by the European Commission.[8]

---

[5]Alternatively, real-time datasets could be analyzed, see Orphanides (2001).

[6]See, for example, Kuttner (1994) and Cotis, Elmeskov, and Mourougame (*forthcoming*) for reviews.

[7]For example, I do not analyze the class of factor-based forecasts; see the series of papers by Stock and Watson (1989,1991,1999). Another common decomposition, pioneered by Beveridge and Nelson (1981), is omitted due a lack of sufficient data.

[8]See, e.g., Denis, Mc Morrow, and Roeger (2002). Proietti, Musso and Westermann (2002)

In Section IV, the various output gap estimates are evaluated econometrically in an inflation forecasting framework similar to Stock and Watson (1999). They investigate forecasts of US inflation at the 12-month horizon, using information from 168 additional indicators of economic activity, including output gap estimates. The present paper, however, focuses exclusively on output gap measures and is, in spirit, related to comparative output gap studies, such as Scacciavillani and Swagel (1999), Cerra and Saxena (2000), and Ross and Ubide (2001), who compare measures for Israel, Sweden, and the Euro area. Recent contributions include Orphanides and van Norden (2002) and Robinson, Stone, and van Zyl (2003), who evaluate inflation forecasts based on real-time measures of the output gap for the United States and Australia. Finally, the statistical test to compare forecast performance used below is based on Diebold and Mariano (1995).

The use and abuse of the output gap in policymaking and related research are manifold. In the field of monetary policy, much of the discussion over the last decade or so has focused on the advantages of rules versus discretion, and the conduct of monetary policy under major regimes relies on some form of the output gap. In a seminal paper, Taylor (1993) proposed a simple instrument rule that tracks US monetary policy surprisingly well during the 1980s and early 1990s. In the simplest form of the rule, interest rates are adjusted according to deviations of inflation from a target level and of output from its trend—that is, the output gap. Since then, uncountable papers have investigated similar issues.[9] It has also been recognized that—notwithstanding the terminology—the output gap plays an important role in Inflation Targeting, at least in its "flexible" form, see Svensson (1999). For the euro area, Gerlach and Svensson (2003) attribute greater importance to the output gap than to money growth as a predictor of future inflation.

Regarding fiscal policy, the concept of output gap has acquired operational—but not legal—status in the Stability and Growth Pact (SGP) of the European Union to calculate annual fiscal balances, and much public attention has been dedicated to the enforcement of the Excessive Deficit Procedure (EDP).[10] In fact, there is a growing discussion on the need to recast the SGP in terms of "structural" instead of "medium-term" fiscal balance.[11] However, it is not clear from the analysis below whether commonly computed output gap measures form a solid basis for calculating structural balances.

_____

evaluate unobserved components models based on the production function approach for the Euro area as a whole.

[9]Svensson (2003) provides an extensive survey. More specifically, the consequences of output gap uncertainty for the Taylor rule are discussed by Smets (1998) for the U.S. and by Eleftheriou (2003) for the Euro area.

[10]The report of the Economic Policy Committee acknowledges the output gap as an essential —but so far only intermediate—input for assessing the progress made by countries towards achieving the goal of medium-term fiscal balance; see European Commission (2001).

[11]See, e.g., Begg and others (2002).

### III. Output Gap Measures

The four output gap measures considered below decompose a time series (here output, $y_t$) additively into a cyclical component, $y_t^c$ and a trend component, $y_t^*$:

$$y_t = y_t^c + y_t^*. \tag{1}$$

The trend component is assumed to coincide with potential output, that is, the amount of output that can be achieved under normal capacity utilization and given the constraints in the labor market, in particular given the natural rate of unemployment. Accordingly, the output gap measures the relative distance of actual output from trend (the "smoothed" series), that is, the cyclical component:

$$gap_t = \frac{y_t - y_t^*}{y_t^*} = \frac{y_t^c}{y_t^*}. \tag{2}$$

The measures fall into two broad categories: univariate statistical filters, and more theory-related measures based on an underlying economic model.[12]

Statistical filters can extract information either in the conventional time domain—exemplified by the Hodrick-Prescott (1997, HP) filter—or in the frequency domain. The frequency domain approach (FD) is represented by a filtering method developed by Corbae and Ouliaris (2002), drawing on earlier results by Corbae, Ouliaris, and Phillips (2002).

Among the many measures of potential output that rely to a larger extent on economic theory, the paper evaluates the permanent-transitory decomposition by Blanchard and Quah (1989, BQ) and the production function approach (PF). In the case of the Blanchard-Quah routine, the economic reasoning is tied to the conventional distinction of "demand" versus "supply" shocks, whereas the production function methodology is based on a model of the aggregate production structure of the economy.

A discussion of these measures was provided in a related paper, Billmeier (2004), and is presented in the Appendix.

### IV. Comparing Output Gaps

In this section, the output gap measures are compared descriptively and evaluated with regard to their information content. The data on real GDP for the five sample countries (Finland, France, Greece, Italy, and the United Kingdom) are annual and stem largely from the European Commission's database, as does the unemployment rate representing the demand side-related variable in the Blanchard-Quah decomposition. The same holds true for the variables used to determine the NAWRU and, ultimately, the output gap according to the production function approach. The CPIs used in the inflation forecasting exercise in Sections IV.C and IV.D come

---

[12]The output gap measures differ in the amount of parameters estimated. The issue of parameter instability due to the relatively short estimation period is acknowledged, but not pursued further.

from the International Monetary Fund's International Financial Statistics database.

## A. General Remarks

Several remarks about the empirical strategy are in order. First, predictability tests can be based on the in-sample fit of a model or on the out-of-sample fit obtained from a sequence of recursive or rolling regressions. In the context of inflation forecasting, the latter set-up mimics the data constraints faced by a policy-maker in real time, appears to be the more appropriate evaluation technique. The difference between these two evaluation techniques can be substantial in nested models—see Inoue and Kilian (2002)—but matters much less in non-nested comparisons, such as the ones below.

Second, the tools used by practitioners for ranking models are the same whether the forecasting models are nested or not. In applied work, forecasting models are often chosen or dismissed on the basis of their root (predictive) mean square error (RMSE) compared to a base forecast or a derivative thereof; a well-known example is Meese and Rogoff (1983). However, using relative RMSEs for model selection—as it is also done below—comes at the drawback that there is no straightforward measure of significance.

Third, the output gap is not directly observable—a ghost—and all the standard econometric caveats about two-stage estimation apply. In principle, inference about second-stage inflation forecasts needs to deal with underlying parameter uncertainty of the output gap measures estimates in the first stage; see West (1996). Instead, a test statistic proposed by Diebold and Mariano (1995) has become one of the more common ways to compare alternative forecast paths to the true series—and is also applied further below. This test provides information on whether forecast $i$ is significantly better than forecast $j$ but does not take parameter uncertainty into account. In fact, only two of the gap measures examined, the production function approach and the BQ filter, are affected by first-stage parameter uncertainty. The other two filters (HP and frequency domain) are based on assumptions for the crucial parameters and suffer only from the conventional sampling uncertainty. Furthermore, West (1996) shows that under the assumption that OLS provides consistent estimates of the parameters (such as in the VAR used to estimate the BQ decomposition), one can safely ignore parameter uncertainty when testing for differentiable functions of parametric forecasts and forecast errors such as the mean square error.[13] Consequently, the Diebold-Mariano test statistic is applied in this paper to shed some further light on the usefulness of the output gap measures chosen in predicting inflation.

## B. Descriptive Assessment

Table 1 presents descriptive statistics for the four standard output gap estimates considered (and

---

[13]The same holds true for a limited number of other cases, including the one of a large estimation sample size relative to the prediction sample size; see, e.g., Diebold (2001). McCracken (2000) points out that under the same conditions, parameter uncertainty is not necessarily irrelevant for the moments of non-differentiable functions of parametric forecasts and forecast errors such as the mean absolute error.

discussed in Appendix I). In the correlation matrix, statistics below the main diagonal represent simple correlations of the main gap measures in levels to gauge the similarity of gap measures. The upper triangle, instead, offers correlations of first differences, which indicate whether the gap measures convey the same directional message, that is, whether the economy is improving or not. A number of observations can be made. First, most gaps—with the possible exception of the one based on the BQ decomposition—are centered around zero. Maxima and minima seem in a reasonable range and speak—together with the standard deviation—of the more or less bumpy road the sample countries have traveled down. Particularly interesting in this context is France, where the consistently lowest variation in the sample reflects relatively smooth economic growth, close to potential. The opposite is represented by Finland. Both the extremes and the surprisingly high standard deviation reflect the boom-bust cycle in the early 1990s, when overheating of the economy lead to a burst of the property-price bubble and an economic downturn unrivaled among OECD countries after WWII.[14]

Table 1. Descriptive Output Gap Statistics, 1960–2002

| | Sample statistics | | | | | Correlations | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | Min | Max | SD | Gap[15] | HP100rt | PF | BQ | FD2-8 |
| | | | Finland | | | | | | |
| HP100rt | 0.08 | -7.44 | 5.77 | 3.20 | -0.25 | 1 | 0.94 | 0.44 | 0.83 |
| PF | -0.08 | -6.49 | 6.43 | 3.00 | -1.55 | 0.82 | 1 | 0.48 | 0.87 |
| BQ | -0.38 | -9.52 | 4.00 | 3.27 | 1.06 | 0.09 | 0.27 | 1 | 0.48 |
| FD2-8 | -5.32 | -5.32 | 5.06 | 2.61 | -1.42 | 0.57 | 0.70 | 0.35 | 1 |
| | | | France | | | | | | |
| HP100rt | 0.07 | -2.74 | 2.81 | 1.41 | 0.46 | 1 | 0.90 | 0.18 | 0.88 |
| PF | -0.13 | -2.72 | 2.81 | 1.43 | 0.07 | 0.81 | 1 | 0.01 | 0.79 |
| BQ | 0.22 | -1.11 | 2.70 | 0.99 | 0.66 | 0.36 | 0.26 | 1 | 0.17 |
| FD2-8 | 0.00 | -2.44 | 2.07 | 1.01 | -0.90 | 0.66 | 0.50 | 0.08 | 1 |
| | | | Greece | | | | | | |
| HP100rt | 0.00 | -5.04 | 5.40 | 2.36 | 2.78 | 1 | 0.96 | -0.05 | 0.95 |
| PF | 0.69 | -3.19 | 4.65 | 4.83 | 0.42 | 0.77 | 1 | 0.09 | 0.91 |
| BQ | -0.43 | -8.13 | 8.59 | 3.23 | -0.12 | -0.32 | 0.12 | 1 | -0.09 |
| FD2-8 | -0.04 | -3.89 | 5.19 | 1.84 | 0.61 | 0.68 | 0.62 | -0.12 | 1 |
| | | | Italy | | | | | | |
| HP100rt | -0.07 | -7.12 | 3.33 | 1.95 | -0.28 | 1 | 0.78 | -0.17 | 0.83 |
| PF | -.34 | -4.76 | 3.02 | 1.85 | -0.81 | 0.66 | 1 | 0.32 | 0.72 |
| BQ | 0.96 | -5.51 | 5.54 | 3.98 | 0.22 | 0.11 | 0.20 | 1 | -0.30 |
| FD2-8 | -0.00 | -4.14 | 3.57 | 1.53 | -0.53 | 0.55 | 0.50 | -0.08 | 1 |
| | | | United Kingdom | | | | | | |
| HP100rt | 0.06 | -3.79 | 5.21 | 1.86 | 0.00 | 1 | 0.98 | 0.55 | 0.92 |
| PF | -0.09 | -4.18 | 5.32 | 1.99 | -0.78 | 0.92 | 1 | 0.47 | 0.89 |
| BQ | -0.38 | -9.32 | 4.00 | 3.27 | 1.06 | 0.64 | 0.57 | 1 | 0.54 |
| FD2-8 | 0.06 | -3.09 | 4.57 | 1.47 | 0.44 | 0.70 | 0.64 | 0.43 | 1 |

Notes: HP100rt, real-time Hodrick Prescott filter; PF, production function approach; BQ, Blanchard-Quah decomposition; FD2-8, frequency domain filter. Lower triangle gives level, upper triangle first-difference correlations.

---

[14]See Billmeier (2004) for a more detailed evaluation of the Finnish experience.

[15]Output gap in 2002.

Another striking feature of the table is that the 2002 gap estimates are not consistently positive or negative for any single country. This intrinsic uncertainty regarding the output gap is also reflected in the sometimes surprisingly low correlations between the various measures. Across countries, particularly the BQ decomposition seems to yield an output gap measure that is somewhat distinct from the other three. This is true for correlations in levels and first differences.

Figures A1-A3 in the Appendix present the output gap measures for the five countries in the sample. The above-mentioned boom-and-bust cycle in Finland in the early 1990s is clearly visible, as are the consequences of exiting the ERM for Italy and the United Kingdom. For most countries, the dispersion of gap measures seems to increase toward the end of the sample period, with Greece being a particularly good example. The generally low correlation of the Blanchard-Quah-based measure is clearly visible, especially, again, in Greece. The figures and corresponding cross-country correlations (not presented here) also document the role of Greece and, to a lesser extent, the United Kingdom as outliers from the "European" business cycle.[16]

An additional descriptive statistic—measuring the consistency of gap signals—yields broadly similar results (See Table A1 in the Appendix). This measure is constructed as the share of total observations in which two gap measures give the same cyclical (i.e., boom or bust) signal.[17] Similar to the correlation statistics, the BQ measure again displays the lowest consistency with other measures for most countries.

## C.    Econometric Evaluation

The output gap is often considered a useful instrument to gauge domestic inflationary pressures. Consequently, the information stemming from an output gap measure could increase the precision of inflation forecasts. These forecasts are compared using a simulated out-of-sample methodology. The forecasting model is a variant of the Phillips curve:

$$\pi_{t+1} - \pi_t = \alpha + \beta(L)gap_t + \gamma(L)\Delta\pi_t + e_t \tag{3}$$

where $\pi_{t+1}$ denotes the one-year ahead inflation in the price level at period $t$, $\pi_t$ is the actual inflation in period $t$, $gap_t$ denotes the output gap measure (in levels), $L$ is the lag operator, and $e_t$ an i.i.d. error.[18]

---

[16]For both countries (and across all gap measures), the degree of business cycle integration did not increase substantially after 1990 when compared to pre-1990 (with the exception of the UK-Finnish cycles). While an interesting subject in itself, the issue of business cycle synchronization is beyond the scope of this paper.

[17]This measure differs from the first difference of the gap considered above in that two measures could signal an improvement ($\Delta gap > 0$) but not necessarily the same cyclical position (negative vs. positive gap).

[18]Proietti, Musso, and Westermann (2002) find that the first difference of the output gap (but not the level) is a significant predictor of inflation in the Euro area. Selected experiments with differenced output gap measures yielded results close to the ones presented and have, hence, been omitted; see also the rather similar correlations between gaps in levels and differences in Table 1.

This specification of the forecast equation mirrors a classic Phillips curve, with the output gap measure substituting for the unemployment rate.[19,20] The evaluation period spans from 1990 to 2002, observations are annual. The simulated out-of-sample procedure consists of the following steps: first, a model is estimated for the period 1960 through 1989 (with data available up to 1989). Lag length selection of each estimated model up to a maximum number of lags is based on minimization of the Akaike information criterion. Due to the low frequency of the data and the limited number of observations, a maximum of two lags is chosen. Section IV.D examines the robustness of this assumption. Second, a one-year-ahead forecast for inflation in 1990 is made. This value is compared to the actual inflation, yielding the forecast error for the first year of the evaluation period. Next, the same procedure is repeated including data until 1990, forecasting inflation in 1991. This exercise is computed recursively, that is, for every year until 2001, the model is re-estimated according to the new information criteria, and the forecast error is computed. This procedure yields a unique series of forecast errors for each output gap measure considered.[21]

Table 2 presents two statistics: in addition to the cumulative root mean square error (RMSE), the so-called U statistic proposed by Theil (1971) is given. The latter consists of the RMSE of a specific inflation forecast standardized by the RMSE of the naïve forecast of "no change" (NC) in inflation. A value smaller than unity stands for a smaller RMSE than under the naïve hypothesis. The obvious advantage of this statistic lies in its ease of comparability across countries. As additional—and often more challenging—benchmarks, the models were also estimated without any output gap measures, that is, as a univariate inflation forecast ("AR") and under the classic Phillips curve specification, which includes the unemployment rate as a RHS variable ("UR"). In the table, specifications that "beat" the univariate model, a very common benchmark in the literature, are in bold.

---

[19]As Stock and Watson (1999) point out, this specification assumes that (i) the inflation rate is integrated of order one (I(1)); (ii) $x_t$ is I(0); and (iii) both are, hence, not cointegrated. Moreover, the constant intercept implies that the "natural rate" of the output gap is constant. In this literature, inflation is commonly modeled as an I(1) process. Results not reported here have confirmed this assumption for wage and CPI inflation in the sample countries; for Finland see also the discussion in the Appendix of Billmeier (2004). While the output gap may seem to behave like an integrated process over limited periods of time, it is clearly mean-reverting from a theoretical perspective.

[20]Given that a main ingredient of the output gap based on the production function approach—the natural rate of unemployment—is derived from a similar framework, the evaluation could expected to be biased in favor of this approach. This, however, does not hold true for at least two reasons: (a) the framework described in Appendix I.D is based on wage inflation whereas the evaluation measures performance in forecasting CPI inflation; and (b) the natural rate of unemployment is only one building block of potential output according to the production function approach—with total factor productivity being quantitatively much more important most of the time.

[21]Parameter estimates of equation (8), while not reported to conserve space, are broadly in line with expectations. In particular, $\widehat{\beta}$ is estimated to be positive and often significant.

Table 2. Evaluation of Forecast Performance I, 1990–2002

|  | HP100rt | PF | BQ | FD2-8 | NC | AR | UR |
|---|---|---|---|---|---|---|---|
| | | | *Finland* | | | | |
| RMSE | 1.90 | 1.71 | 1.18 | 1.52 | | 1.14 | 2.22 |
| Theil's U | 1.24 | 1.11 | 0.77 | 0.99 | 1 | 0.74 | 1.45 |
| | | | *France* | | | | |
| RMSE | 0.64 | 1.24 | 0.55 | 0.54 | | 0.53 | 0.78 |
| Theil's U | 0.93 | 1.81 | 0.80 | 0.78 | 1 | 0.77 | 1.13 |
| | | | *Greece* | | | | |
| RMSE | 2.93 | **1.99** | 3.24 | 2.50 | | 2.23 | **2.07** |
| Theil's U | 1.20 | **0.81** | 1.33 | 1.02 | 1 | 0.91 | **0.85** |
| | | | *Italy* | | | | |
| RMSE | 1.11 | 1.00 | **0.87** | 0.96 | | 0.93 | 0.94 |
| Theil's U | 0.98 | 0.88 | **0.77** | 0.85 | 1 | 0.82 | 0.84 |
| | | | *United Kingdom* | | | | |
| RMSE | **1.53** | 2.06 | 1.99 | 1.84 | | 1.73 | 1.87 |
| Theil's U | **0.67** | 0.91 | 0.87 | 0.81 | 1 | 0.76 | 0.82 |

Notes: see Table 1; NC, assumption of no change; AR, auto-regressive estimate; UR, unemployment rate (Phillips curve specification).

The results can be summarized as follows:

- In Finland and France, no output gap measure yields better results than a univariate inflation forecast. In Finland, the disappointing performance of the output gaps in improving the inflation forecast is most likely due to the high volatility of output itself (see Table 1), hampering the determination of a statistically satisfying measure of potential output.

- Again in Finland and France, the classic Phillips curve featuring the unemployment rate fares much worse than the naïve forecast, whereas this does not hold true to the same extent in the other three countries. This indicates that in these countries, the unemployment rate is not a particularly good indicator for inflationary pressures stemming from the labor market. In fact, large-scale active labor market policies reduce the amount of de facto unemployed in Finland by up to 40 percent; see Feldman and others (2003).

- In Greece, both a production-function-based output gap measure ("PF") and the unemployment rate (i.e., the classic Phillips curve specification) improve the inflation forecast performance.

- In Italy, the gap measure based on the Blanchard-Quah decomposition ("BQ") seems to provide a somewhat better grip on the data than the univariate model.

- The United Kingdom is the only country in our sample where the very popular HP filter in its real-time version for $\lambda = 100$ ("HP100rt") delivers good results, together with the Blanchard-Quah decomposition.

Taken together, these results imply that (i) the univariate inflation forecast is not necessarily improved by adding an output gap measure and that (ii) if a forecast-improving output gap exists, it usually varies by country.

Table 3. Evaluation of Forecast Performance II, 1990–2002

|  | HP100rt | PF | BQ | FD2-8 | UR |
|---|---|---|---|---|---|
| *Finland* | | | | | |
| DM | 2.10 | 1.51 | 0.49 | 2.79 | 2.36 |
| (p-value) | (0.98) | (0.93) | (0.69) | (0.99) | (0.99) |
| *France* | | | | | |
| DM | 0.66 | 2.87 | 0.89 | 0.18 | 1.91 |
| (p-value) | (0.75) | (0.99) | (0.81) | (0.57) | (0.97) |
| *Greece* | | | | | |
| DM | 1.92 | -0.81 | 1.32 | 0.77 | -0.27 |
| (p-value) | (0.97) | (0.21) | (0.91) | (0.78) | (0.39) |
| *Italy* | | | | | |
| DM | 1.89 | 2.13 | -0.90 | 0.26 | 0.15 |
| (p-value) | (0.97) | (0.98) | (0.18) | (0.60) | (0.56) |
| *United Kingdom* | | | | | |
| DM | -0.56 | 1.13 | 0.94 | 0.39 | 0.55 |
| (p-value) | (0.29) | (0.87) | (0.83) | (0.65) | (0.71) |

Note: DM is the Diebold-Mariano test statistic.
For gap measures, see tables 1 and 2.

To explore the statistical significance of these results, the inflation predictions are assessed using a simple version of the test proposed by Diebold and Mariano (1995). This test assesses whether—conditional on the true series—the inflation path predicted by adding one of the output gap measures to the univariate model is significantly better than the benchmark autoregressive forecast itself. Consider the original inflation series, $\{\pi\}_t^T$, two forecasts, $\{\pi_t^i\}_t^T$ and $\{\pi_t^j\}_t^T$, and the corresponding errors, $\{e_t^i\}_t^T$ and $\{e_t^j\}_t^T$. The loss function associated with a specific forecast $i$, $l(\pi_t, \pi_t^i)$, will be in many—but not all—cases a direct function of the forecast error, $l(e_t^i)$. The null hypothesis of equal forecast accuracy for the two forecasts can then be expressed as:

$$E[d_t] = 0, \tag{4}$$

where $d \equiv [l(e_t^i) - l(e_t)]$ is the loss differential. In other words, under the null, the population mean of the loss-differential series is 0. Under the alternative, forecast $i$ is better than forecast $j$. Empirically, the Diebold-Mariano statistic is simply the "t-statistic" of a regression of $d$ on a constant with heteroskedasticity autocorrelation consistent (HAC) standard errors; see Diebold and Mariano (1995) for details.

Table 3 presents the Diebold-Mariano (DM) test statistic (and the corresponding p-value) for the forecasts to be equally accurate. The loss function is specified as

$$d = \frac{1}{T} \sum_{n=t}^{T} \left[ l\left(e_{t,T}^i\right) - l\left(e_{t,T}^j\right) \right]. \tag{5}$$

In other words, $d$ is the sample mean loss differential. As the benchmark forecast, the autoregressive inflation forecast is considered (as opposed to Table 2, in which the base case is "no change"). Failure to reject the null hypothesis implies that the inclusion of the output gap measure does not improve the univariate model significantly.

The results obtained by applying the Diebold-Mariano procedure mirror those anticipated in Table 2—but go beyond in one important aspect. While the RMSE ratios—the Theil statistic—indicated a few output gap-based forecasts that could beat the autoregressive forecast in Greece, Italy, and the United Kingdom, the DM statistic reveals that none of these forecasts is *significantly* better than the AR process. In fact, the null hypothesis of similar forecast prediction accuracy cannot be dismissed for any gap-based inflation forecast at conventional significance levels (of 5 or 10 percent).

There are at least two simple explanations for the lack of significance of the estimates. First, domestic inflationary pressures (as captured by the output gap) may not be the main driver of CPI inflation. This observation is particularly relevant for relatively open economies that do not peg their exchange rate to their major trading partners and are, therefore, more directly affected by fluctuations in the exchange rate.[22] In the sample, this holds true for the United Kingdom (belonging more to the US cycle), and, to some extent, also for Finland and Greece. While Finland's trade patterns shifted towards Europe only after the collapse of the Soviet Union, Greece is slowly becoming a hub for the countries in the Balkans—trading less and less with EU countries. For Italy and France, however, this explanation does not prove satisfactory because both countries shared a stable peg with their major trading partners, at least for most of the sample period. Second, the analysis suffers from a lack of observations. Basing the forecasts on annual observations is beneficial, in that some of the data revision issues are avoided, but comes at the cost of a significantly reduced sample.[23] While the alternative approach—basing the forecasts on quarterly observations—would probably be more relevant from a monetary policymaker's perspective, the lack of significance using annual observations clearly shows that the output gaps used in the prediction exercise do not fully capture the business cycle—at least not from the perspective of inflationary pressures arising in an overheating economy.

As a corollary, these results also cast doubt more broadly on using output gaps to calculate a cyclically-adjusted fiscal balance. First, none of the gap measures seems to capture domestic inflationary pressures well, which are arguably an indicator of the cyclical position of the economy. Second, it is not clear a priori which measure of the output gap should be used, since there are many and they would result in widely different estimates of an adjusted deficit or surplus.

### D.    Robustness Checks

In this section, the previous results are checked for robustness in two different dimensions. First, for those gap measures that are subject to a crucial single assumption in the first stage, this assumption is modified. This concerns the penalty parameter $\lambda$ used in the HP filter and the assumption on the business cycle length in the frequency domain filter. Second, the results are tested with respect to a modelling choice in the second step, namely the maximum number of lags

---

[22]This is true under the assumptions that the relevant price index, e.g., the CPI, contains imported goods and that the exchange rate pass-through is positive.

[23]This effect would be even more noticeable if the forecast took parameter uncertainty into account; see Section IV.A.

in the prediction model.

Generally speaking, HP filters based on differing assumptions on $\lambda$ find a qualitatively similar pattern of the output gap, that is their turning points coincide chronologically. The assumption on the smoothness of the trend has, however, strong implications for the magnitude of the gap, in particular at the end of the observation period. Since the "real-time" constructs in the previous section consist in the last part (from 1990) of a series of "last observations," the choice of $\lambda$ becomes even more critical. Here, we redo the exercise for $\lambda = 20$ ("HP20rt")and $\lambda = 200$ ("HP200rt"), generating a more volatile (HP200rt) and a less volatile (HP20rt) gap measure. Moreover, we flag for comparison the results for the traditional (two-sided) HP filter ("HP100," "HP20," "HP200") to gauge the importance of the "real time" construct. A similar argument applies to the frequency domain filter. To corroborate the above results, the filter has been re-estimated assuming business cycle durations between 2 and 6 years ("FD2-6") and between 2 and 10 years ("FD2-10"). To increase comparability with the results described above, the maximum lag length has been held constant (at 2).

Results for the alternative parametrization at the first stage (reported in table A2 in the Appendix) confirm to a large extent the outcome obtained above. Again, the estimated output gaps do not improve on the univariate prediction for Finland and France. For Greece, only one parametrization of the two-sided HP filter (HP200) produces a slightly but insignificantly smaller RMSE than the autoregressive forecast. For Italy and the United Kingdom, the results from varying the parameters are somewhat more encouraging. For Italy, the output gap measure improves the inflation forecast if a shorter cycle frequency (between two and six years) is assumed in the frequency domain approach. More importantly, this version of the frequency domain is the only forecast that is statistically significant at the 5 percent level in the Diebold-Mariano sense, that is, with at test statistic of $d = -1.97$, it significantly improves upon the autoregressive forecast (p-value 0.02). For the United Kingdom, instead, the real-time HP filter seems to be the best way to capture inflationary pressures—almost independently of the assumption regarding the trend smoothness ($\lambda$). In fact, both alternative real-time specifications produce smaller RMSEs that the autoregressive forecast. In the Diebold-Mariano test, however, none of the United Kingdom real-time measure is statistically significant at conventional levels.

An interesting corollary derives from the comparison of the real-time HP filter with its common, two-sided version. The additional information on the relative position in the cycle stemming from future observations could be expected to improve the quality of the gap measure, and hence the forecast performance. This conjecture is rejected for the countries in the sample. For all countries but Greece, the real-time versions of the HP-filtered gaps provide better forecasts. This observation is both intriguing and reassuring. Consider the following situation: a country has a consistently positive output gap, coinciding with elevated inflation for an extended period with the exception of one year in the middle of the sample. In that year, due to an exogenous shock, the output gap is negative, resulting in somewhat lower-than-average inflation in that period and the following one. The two-sided HP filter, in principle, smooths over the outlier, and consequently predicts relatively high inflation. With the forward-looking information missing, instead, the

real-time filter picks up the drop in inflation to a larger extent, resulting in a reduced smoothing effect and a more accurate prediction of inflation. From a policy making perspective, this is reassuring since the two-sided HP filter is, of course, not available (or based on forecasts). In Greece, though, the two-sided filter yields better inflation forecasts than the real-time variant. Both are dominated, however, by another gap measure, namely the one stemming from the production function (in addition to the Phillips curve prediction using the unemployment rate).

The second-stage robustness check is related to the model parameters. Determination of the optimal lag length for the recursive inflation prediction models requires the specification of a maximum number of lags of the RHS variables in equation (8), that is, the output gap and inflation. With the total number of observations being rather small due to the annual frequency of the data, this limit was set at two in the previous section. Changing this assumption can have a strong impact on the degrees of freedom, and implicitly, on the precision of the estimates.

Most of the results obtained above for a maximum of two lags are robust to a higher or lower limit, see tables A3, A4, and A5 in the Appendix. In particular, for all models evaluated (between one and four lags), no output gap measure yields better results than a univariate inflation forecast in Finland and France.[24] In Greece, if a maximum of one lag is allowed, the BQ filter produces an inflation forecast that is significantly better than the autoregressive estimate at the 10-percent level. In addition, both the production-function-based output gap measure and the classic Phillips curve specification improve the inflation forecast performance at all levels—but not significantly so.[25] In all countries but Greece, instead, the unemployment rate does not provide any useful information when compared to the univariate specification, and hence, is not a good indicator for domestic inflationary pressures. Again, the performance of the unemployment rate is particularly disappointing in Finland and France. The results are the least robust for Italy. While for short lag limits (one and two), the frequency domain and the Blanchard-Quah measure yield the best results, the production function approach (and the real-time HP filter) prevail with a maximum of three and four lags. At three lags, the gap measure based on the production function provides a significant improvement over the autoregressive inflation forecast. For UK data, the HP filter performs well for three and four lags—but remains insignificant.

## V. CONCLUDING REMARKS

In this paper, I compared a number of commonly used output gap measures in an inflation forecasting exercise for a small set of European countries. The measures evaluated included variants of the HP filter, the Blanchard-Quah decomposition, the production function approach, and a frequency domain filter. Reflecting domestic inflationary pressures, the unobservable output

---

[24]For a maximum of one lag, the frequency domain approach is just as good as the univariate forecast.

[25]The production function approach is almost significant for a maximum of one and three lags, indicating some scope for a better performance of a refined estimate of gap based on the production function.

gap should, at least in theory, provide some information for one-year-ahead actual inflation and, hence, improve a univariate forecast.

So which ghost should one go after? The economic answer, judging from the sample countries, is that *it depends*. In Finland and France, the widely used output gap measures evaluated in this paper do not provide any improvement over a univariate inflation forecast. In Greece, Italy, and the United Kingdom, however, inflation forecasts are better if some measure of the gap is included. The best measure to be included varies by country, but, where applicable, it is robust to alternative ways of computing the measure (Italy: frequency domain method; United Kingdom: HP filter), and modeling assumptions in the inflation forecast exercise (with the possible exception of Italy). Unfortunately, few of these forecasts perform significantly better than the autoregressive forecast in a statistical sense as documented by the Diebold-Mariano test. From a policymaking perspective, the conclusion is that (i) various measures of the output gap should be taken into account when assessing the cyclical position of the economy; (ii) it is hard to significantly improve on an autoregressive forecast using the simple output gap measures presented above; (iii) a broader set of indicators may be needed to capture domestic inflationary pressures well; and (iv) assessing the fiscal stance based on a structural balance is difficult if the gap measure is to reflect the business cycle well.

The failure to identify output gap measures that improve consistently and significantly upon the univariate forecast in all countries—but, in particular, Finland and France—could also be due to the limited number of observations, given the choice of annual frequency. In principle, this choice is well-motivated by seasonality issues, data revision considerations, and, to some extent, the availability of data (measures of the capital stock and the natural rate of unemployment in the production function approach). Nevertheless, a quarterly assessment may yield additional insights. Thus, directions for future research include: analyzing real-time data series at a quarterly frequency, also to better mimic the (monetary) policymaker's decisionmaking process; extending the analysis to a larger sample of countries; and including "optimized" versions of the HP and the frequency domain filters—as opposed to the "conventional" robustness analysis with regard to the crucial parameters undertaken above.

## I. Measures of the Output Gap

In this appendix, a brief discussion of the four output gap measures is provided.

### A.    The Hodrick-Prescott Filter

The Hodrick-Prescott (HP) filter is probably the most well-known and most widely used statistical filter to obtain a smooth estimate of the long-term trend component of a macroeconomic series. This is chiefly due to its simplicity, but also to the fact that, for the United States, business cycle movements can be extracted that resemble the NBER-backed definitions (see Canova (1999)). The HP filter is a linear, two-sided filter that computes the smoothed series by minimizing the squared distance between trend ($y_t^*$) and the actual series ($y_t$), subject to a penalty on the second difference of the smoothed series:

$$\underset{y_t^*}{Min} \left\{ \sum_{t=1}^{T} (y_t - y_t^*)^2 + \lambda \sum_{t=2}^{T-1} \left[ \left( y_{t+1}^* - y_t^* \right) - \left( y_t^* - y_{t-1}^* \right) \right]^2 \right\} \tag{A-1}$$

The penalty parameter, $\lambda$, controls the smoothness of the series by setting the ratio of the variance of the cyclical component and the variation in the second difference of the actual series. A higher value for $\lambda$ implies a smoother trend (and, hence, more volatile gaps). In the extreme case of $\lambda \to \infty$, the trend is a straight line. The standard value in the literature is $\lambda = 100$ for annual data, which is also assumed as a base case in the analysis.[26]

In a policy-related context, the traditional Hodrick-Prescott measure poses a substantial problem: the filter as described above is fundamentally a two-sided filter, that is, computation of the underlying trend at time $t$ is based on observations *before* and *after* period $t$. Economic policy makers, instead, will—at the time of decision-making—only dispose of an estimate of the output gap that is based on a purely backward-looking evaluation of potential output. Hence, a "real-time" output gap series based on the HP filter is constructed, $HP.rt$. This new series consists of "last observations," that is, real-time estimates of the underlying trend in the last observation period $t$ given the information set in period $t$.[27]

This way to proceed is subject to two important caveats. First, an observation for output produced in period $t$ has to be a prediction while the economy is still in period $t$ and will be finally observed only in $t + 1$ (for example, data on 1999 GDP is only issued (at best) in the course of 2000) Second, data may be revised in later periods. I abstract from these important details, since the empirical analysis will build on yearly observations, implying that by the end of a given year, the first three quarters of the yearly figure for output have already been observed and provide a sound

---

[26]This follows Burns and Mitchell (1946) and Hodrick and Prescott (1997); see Section IV.D for a robustness check regarding the value of $\lambda$. Ross and Ubide (2001) discuss alternative approaches to determine the parameter $\lambda$ endogenously.

[27]See Stock and Watson (1999) for a similar argument.

footing for an end-year estimate. Annual data are also less likely to be affected by substantial data revisions since these revisions usually occur in the periods immediately following the (quarterly) observation, hence, mostly before the end of the calendar year.  In addition, revisions due to seasonal factors are limited for annual data.[28]

Other prominent drawbacks of the HP filter (in the version described above) have been well documented in the literature and include the possibility of finding spurious cycles for integrated series, the somewhat arbitrary choice of $\lambda$, as well as the neglect of structural breaks and shifts.[29] Section IV.D elaborates on the robustness of the results with regard to the choice of $\lambda$.

## B.        The Frequency Domain Approach

Economic fluctuations occur at different frequencies (displaying, for instance, seasonal, or business cycle duration). Starting from the classical assumption contained in Burns and Mitchell (1946) that the duration of business cycles takes between 6 and 32 quarters, the approach to extracting those cycles from a stationary time series is relatively straightforward from the frequency domain perspective. The original series should be filtered in such a way that fluctuations below or above a certain frequency are eliminated.

This can be achieved with the help of a so-called exact band-pass filter (BPF). An exact BPF acts in principle as a double filter: it eliminates frequencies outside a range, here the business cycle frequency.  For estimation purposes, however, these filters are usually spelled out in the time domain, since integrated series—such as real GDP—could traditionally not be handled by frequency domain approaches.[30]  However, transformation of the exact band-pass filter back into the time domain results in a moving average process of infinite order. For this reason, Baxter and King (1999) and others have provided time domain approximations to the exact band pass filter capable of dealing with integrated series.  Their method involves a trade-off between the quality of approximation and the ability to smooth the series at the extreme points of the sample, since every additional lag employed in the estimation process improves the filter but translates

---

[28]Orphanides and van Norden (2002) show that ex-post revisions of output gap estimates for quarterly US GNP data are of the same order of magnitude than the gap itself. The bulk of the revisions, however, is attributed to unreliable end-sample estimates, not revisions of published data.

[29]See, e.g., Harvey and Jaeger (1993), King and Rebelo (1993) and Cogley and Nason (1995) for overviews of the shortcomings.  Billmeier (2004) provides an illustration of another problem of the HP filter, the end-sample bias. The discussion of the optimal $\lambda$ is circumvented here by comparing 3 values. While Ravn and Uhlig (2002) argue a value of 6.25 for annual observations, they base their argument on the assumption that $\lambda = 1600$ is the optimal value for quarterly data (which is the common assumption for the United States, but not necessarily true for other countries). Artis, Marcellino and Proietti (2002) argue the superiority of the band-pass version of the HP-filter.

[30]As Corbae and Ouliaris (2002) explain, this is due to a "leakage problem." The frequency responses generated by the discrete Fourier transform of an I(1) process are dependent across fundamental frequencies.

into one lost observation at either end of the series.  This, in turn, substantially diminishes the attractiveness of this class of filters for policy-related analysis.  Alternatively, the estimation can take place directly in the frequency domain.  According to Baxter and King (1999), pre-filtering of the non-stationary series is required to remove stochastic trends and, hence, avoid the leakage problem with integrated series.  They argue that upfront detrending of the series in order to apply discrete Fourier transforms involves a discretionary choice of the detrending method, whereas the symmetric moving average approximation would successfully remove any deterministic or stochastic trends up to second order.  In this context, Corbae and Ouliaris (2002) provide a frequency domain fix for the leakage problem and, hence, a consistent band pass filter for non-stationary data.  In addition, the filter does not involve a loss of observations at either end—a property highly relevant for policymaking.[31]

In the base case econometric evaluation in Section IV.C, a business cycle duration between 2 and 8 years is assumed.  Robustness checks are contained in Section IV.D.

While the major advantage of the frequency domain approach and, indeed, other statistical methods not mentioned here such as simple arithmetic detrending is their simplicity, they are subject to the criticism of lacking foundation in economic theory.  Thus, the next two sections turn to theory-based models of trend GDP and the output gap.

## C.      The Blanchard-Quah Decomposition

The appeal of the approach by Blanchard and Quah (1989) to the identification of structural shocks in a VAR stems from its compatibility with a wide array of theoretical models.  In a bivariate model, structural supply and demand shocks are identified by assuming that the former have a permanent impact on output, while the latter can only have a temporary effect.  In particular, two types of (uncorrelated) structural disturbances are postulated, which possibly affect two time series, (log) real GDP and the unemployment rate.  The following assumptions identify these disturbances: no disturbance has long-run effects on the time series employed in the estimation, more precisely on the first differences of the original time series (i.e., growth rates are stationary).  Furthermore, disturbances to (the growth rate of) real GDP may have long-run effects on the level of both series, while disturbances to the unemployment rate are restricted to not having long-run effects on the level of output. These assumptions technically identify the shocks. Given the chosen structure, it seems natural to label the shocks as supply and demand shocks.[32]

In the present context, potential output is associated with cumulated supply shocks, whereas the output gap reflects cyclical (temporary) swings in aggregate demand.  This approach, hence,

---

[31]See Corbae and Ouliaris (2002) for a technical description of the filter and its small sample properties, and Corbae, Ouliaris, and Phillips (2002) for the analysis of the asymptotic case.
[32]Blanchard and Quah (1989) also show that small violations of the identification scheme (e.g. lasting effects on output stemming from nominal shocks through a wealth effect) are of minor consequence.  The empirical set up has been employed and documented numerous times in the literature; see, e.g., Clarida and Gali (1994) for a more detailed description of the approach in the context of a three-variable model.

benefits from explicit economic foundations. Furthermore, the gap—identified as the demand component of output—is not subject to any end sample bias. However, the identification scheme employed may not be appropriate under all circumstances, in particular if the variable representing demand (here the unemployment rate) does not provide a good indication of the cyclical behavior of output. Finally, given the orthogonality assumption on the structural shocks, the amount of variables also determines the number of shocks present in the system. Conversely, there are clearly shocks that have a supply as well as a demand component, for instance, public infrastructure investment.

The VAR models estimated include, in addition to a constant, up to four lags of the endogenous variables, as indicated by information criteria. No residual autocorrelation was present in the specifications chosen.

### D.    The Production Function Approach

Another way to avoid the problem of assigning shocks to demand or supply origins is to start from a growth-accounting perspective. The production function approach describes a functional relationship between output and factor inputs. While the method as such is not new, recent focus on the input factors, in particular labor, has triggered new interest in the subject.[33] I describe both issues in turn.

**The Input-Output Relationship**

Output is at its potential if the rates of capacity utilization are normal, i.e. labor input is consistent with the natural rate of unemployment and technological progress/total factor productivity is at its trend level. A convenient functional form is the Cobb-Douglas type, where output $Y_t$ depends on labor $L_t$ and capital $K_t$, as well as the level of total factor productivity $TFP_t$:

$$Y_t = TFP_t K_t^\alpha L_t^\beta. \tag{A-2}$$

Assuming constant returns to scale implies that $\alpha + \beta = 1$; under perfect competition, $\alpha$ corresponds to the share of capital income, and $\beta = 1 - \alpha$ to the share of labor. Since total factor productivity is not observable, it is usually derived as a residual from the above equation:

$$tfp_t = y_t - \alpha k_t - (1 - \alpha)l_t \tag{A-3}$$

where variables in small caps are in logs. Log trend $TFP$, $tfp^*$, is then obtained by appropriately smoothing this residual series, for instance by a Hodrick-Prescott filter. Potential labor input $L^*$ is taken to be the level of employment consistent with the (time varying) natural rate of unemployment $UR^*$:

$$L_t^* = LF_t (1 - UR_t^*) \tag{A-4}$$

---

[33]Early work on the production function approach includes Artus (1977). Subsequent research has refined the approach in various directions, see, e.g., De Masi (1997) for a recent overview of work done at the IMF. Proietti, Musso and Westermann (2002) use unobserved components models based on a production function to determine potential output in the euro area.

where $LF_t$ is the labor force. Potential output can be written (in logs) as:

$$y_t^* = \alpha k_t + (1 - \alpha)l_t^* + tfp_t^*. \tag{A-5}$$

The most important advantage of the production function approach lies in its tractability together with the possibility to account explicitly for different sources of growth. For instance, the dynamic growth of the Finnish ICT sector during the second half of the 1990s had been mostly driven by potential growth from a productivity point of view and, hence, had resulted in a rather small output gap (see Section IV.B). Moreover, the strong movements of the unemployment rate since the crisis in the early 1990s convey valuable information on labor market conditions. Important shortcomings of the approach include the dependence on a number of crucial assumptions, e.g. (constant) shares of capital and labor, and the functional form of the production relationship (number of input factors, returns to scale). In addition, data requirements can pose significant problems to any production function approach: in particular, the capital stock is difficult to measure consistently at an intra-year frequency.

**Factor Inputs and the NAWRU**

A crucial feature of this approach is the reliance on filtered factor input series, in particular the trend total factor productivity and the natural rate of unemployment. Given the assumption that capital is always employed at full potential, however, no capacity adjustment is usually made to the capital stock.[34] The natural rate of unemployment can be derived in a number of ways, for example by HP-filtering the observed unemployment rate.[35] However, the approach can also be implemented more flexibly by using more sophisticated filtering procedures, including those that incorporate themselves structural assumptions based on economic theory. Here, the calculation of the output gap using the production function approach emphasizes the derivation of the NAWRU (non-accelerating wage inflation rate of unemployment) as a latent variable following Kuttner (1994).[36] From a conceptual point of view, however, this approach rests on the premise that a natural rate of unemployment exists, in other words that the Phillips curve is vertical at said natural rate.

Under the latent variable approach, the natural rate of unemployment—defined here as the NAWRU—is computed using a Kalman filtering process on the observable unemployment rate, to extract the cyclical component. The procedure employs a bivariate model, where the observables

---

[34]In other words, the full-capacity stock of capital is usually approximated by the actual stock of capital. For a more elaborate approach, using French data on capital operating time, see Everaert and Nadal De Simone (2003).

[35]Of course, the choice of a filter to detrend the unemployment rate and TFP adds an element of discretion.

[36]This approach—known as the "GAP model"—was recently adopted by the European Commission; see Denis, Mc Morrow, and Roeger (2002) and Planas and Rossi (2003). In the Commission's work, the new methodology substitutes for more "traditional" approaches—such as the Hodrick-Prescott filter—and, at the same time, unifies the Commission's efforts toward a consistent representation of business cycles in the member countries.

"unemployment rate" and "change in wage inflation" (that is, second differences of wages) play the role of endogenous variables. While the first equation contains a simple decomposition of the observed unemployment rate in trend and cyclical component, the second equation—in principle a Phillips curve—relates the wage inflation to a number of regressors, including lags of wage inflation and the cyclical component of unemployment. Given the error term, wage inflation is assumed to follow an ARMA process. The trend unemployment rate, in turn, serves to determine the (full-employment) stock of labor entering the production function. Estimation takes place in the state-space form, some exogenous regressors (such as a variable reflecting terms of trade) are added for some countries to (marginally) improve the statistical fit.[37]

---

[37]See Berger and Billmeier (2003) for a description of the Finnish case. There, it is shown that both the assumed representation of wage inflation and the inclusion of additional regressors can have substantial impact on trend unemployment.

## II. TABLES

Table A1. Gap Signal Consistency, 1960–2002

|          | HP100rt | PF   | BQ   | FD2-8 |
|----------|---------|------|------|-------|
| Finland  |         |      |      |       |
| HP100rt  | 1       |      |      |       |
| PF       | 0.81    | 1    |      |       |
| BQ       | 0.46    | 0.51 | 1    |       |
| FD2-8    | 0.79    | 0.71 | 0.59 | 1     |
| France   |         |      |      |       |
| HP100rt  | 1       |      |      |       |
| PF       | 0.86    | 1    |      |       |
| BQ       | 0.62    | 0.57 | 1    |       |
| FD2-8    | 0.79    | 0.71 | 0.54 | 1     |
| Greece   |         |      |      |       |
| HP100rt  | 1       |      |      |       |
| PF       | 0.67    | 1    |      |       |
| BQ       | 0.46    | 0.46 | 1    |       |
| FD2-8    | 0.71    | 0.64 | 0.38 | 1     |
| Italy    |         |      |      |       |
| HP100rt  | 1       |      |      |       |
| PF       | 0.83    | 1    |      |       |
| BQ       | 0.73    | 0.65 | 1    |       |
| FD2-8    | 0.76    | 0.76 | 0.65 | 1     |
| United Kingdom |   |      |      |       |
| HP100rt  | 1       |      |      |       |
| PF       | 0.81    | 1    |      |       |
| BQ       | 0.73    | 0.78 | 1    |       |
| FD2-8    | 0.74    | 0.62 | 0.68 | 1     |

Note: Statistic gives ratio of "same signs";

see section IV.B (also for gap abbreviations).

Table A2. Forecast Performance for Alternative Gap Measures, 1990–2002

|  | HP20rt | HP200rt | HP100 | HP20 | HP200 | FD2-6 | FD2-10 |
|---|---|---|---|---|---|---|---|
| | | | *Finland* | | | | |
| RMSE | 1.97 | 1.94 | 1.94 | 2.12 | 1.57 | 1.37 | 1.36 |
| Theil's U | 1.29 | 1.26 | 1.27 | 1.38 | 1.03 | 0.89 | 0.89 |
| DM | 2.34 | 2.12 | 1.82 | 1.55 | 1.70 | 1.35 | 1.99 |
| (p-value) | (0.99) | (0.98) | (0.97) | (0.94) | (0.96) | (0.91) | (0.98) |
| | | | *France* | | | | |
| RMSE | 0.60 | 0.74 | 0.85 | 0.72 | 0.87 | 0.54 | 0.52 |
| Theil's U | 0.87 | 1.07 | 1.23 | 1.04 | 1.26 | 0.78 | 0.75 |
| DM | 0.67 | 1.15 | 1.46 | 1.31 | 1.68 | 0.30 | -0.12 |
| (p-value) | (0.75) | (0.87) | (0.93) | (0.91) | (0.95) | (0.62) | (0.45) |
| | | | *Greece* | | | | |
| RMSE | 3.08 | 2.73 | 2.26 | 3.00 | **2.17** | 2.39 | 2.72 |
| Theil's U | 1.26 | 1.11 | 0.92 | 1.23 | **0.89** | 0.98 | 1.11 |
| DM | 1.77 | 1.04 | 0.04 | 0.87 | -0.11 | 0.25 | 0.97 |
| (p-value) | (0.96) | (0.85) | (0.52) | (0.81) | (0.46) | (0.60) | (0.83) |
| | | | *Italy* | | | | |
| RMSE | 1.07 | 1.24 | 1.29 | 1.25 | 1.24 | **0.69** | 1.11 |
| Theil's U | 0.95 | 1.09 | 1.14 | 1.11 | 1.10 | **0.61** | 0.97 |
| DM | 2.23 | 1.48 | 1.89 | 2.14 | 1.48 | **-1.97** | 1.21 |
| (p-value) | (0.99) | (0.93) | (0.97) | (0.98) | (0.93) | **(0.02)** | (0.89) |
| | | | *United Kingdom* | | | | |
| RMSE | **1.51** | **1.61** | 1.84 | 2.14 | 1.87 | 1.86 | 1.97 |
| Theil's U | **0.66** | **0.71** | 0.81 | 0.94 | 0.82 | 0.82 | 0.86 |
| DM | -0.64 | -0.37 | 0.51 | 1.01 | 0.61 | 0.44 | 0.73 |
| (p-value) | (0.26) | (0.36) | (0.70) | (0.84) | (0.73) | (0.67) | (0.77) |

Note: Bold estimates indicate a better performance than the univariate

forecasting model (see Section IV.B, including for abbreviations).

Table A3. Forecast Performance (max. one lag), 1990–2002

|            | HP100rt | PF     | BQ      | FD2-8   | NC | AR   | UR      |
|------------|---------|--------|---------|---------|----|------|---------|
| *Finland*  |         |        |         |         |    |      |         |
| RMSE       | 1.91    | 1.87   | 1.19    | 1.48    |    | 1.14 | 2.62    |
| Theil's U  | 1.24    | 1.22   | 0.78    | 0.97    | 1  | 0.74 | 1.71    |
| DM         | 2.33    | 1.99   | 0.60    | 2.60    |    |      | 1.71    |
| (p-value)  | (0.99)  | (0.98) | (0.72)  | (0.99)  |    |      | (0.96)  |
| *France*   |         |        |         |         |    |      |         |
| RMSE       | 0.61    | 0.92   | 0.55    | **0.53** |   | 0.53 | 0.66    |
| Theil's U  | 0.88    | 1.34   | 0.80    | **0.76** | 1 | 0.77 | 0.96    |
| DM         | 0.67    | 1.59   | 0.62    | -0.33   |    |      | 0.96    |
| (p-value)  | (0.75)  | (0.94) | (0.73)  | (0.37)  |    |      | (0.83)  |
| *Greece*   |         |        |         |         |    |      |         |
| RMSE       | 2.96    | **1.78** | **1.58** | 2.47  |    | 2.23 | **1.94** |
| Theil's U  | 1.21    | **0.73** | **0.67** | 1.01  | 1  | 0.91 | **0.79** |
| DM         | 2.06    | -1.25  | **-1.47** | 0.85   |   |      | -0.47   |
| (p-value)  | (0.98)  | (0.11) | **(0.07)** | (0.80) |  |      | (0.32)  |
| *Italy*    |         |        |         |         |    |      |         |
| RMSE       | 0.97    | 1.00   | **0.84** | **0.81** |  | 0.90 | 0.91    |
| Theil's U  | 0.86    | 0.88   | **0.75** | **0.71** | 1 | 0.80 | 0.80    |
| DM         | 1.18    | 1.04   | -0.72   | -0.77   |    |      | 0.02    |
| (p-value)  | (0.88)  | (0.85) | (0.24)  | (0.22)  |    |      | (0.51)  |
| *United Kingdom* |   |        |         |         |    |      |         |
| RMSE       | 1.81    | 2.00   | 1.99    | 1.84    |    | 1.65 | 1.87    |
| Theil's U  | 0.80    | 0.88   | 0.87    | 0.81    | 1  | 0.73 | 0.82    |
| DM         | 0.57    | 1.12   | 1.19    | 0.63    |    |      | 0.79    |
| (p-value)  | (0.72)  | (0.87) | (0.88)  | (0.74)  |    |      | (0.79)  |

Note: Bold estimates indicate a better performance than the univariate

forecasting model (see Section IV.B, including for abbreviations).

Table A4. Forecast Performance (max. three lags), 1990–2002

|  | HP100rt | PF | BQ | FD2-8 | NC | AR | UR |
|---|---|---|---|---|---|---|---|
| | | | *Finland* | | | | |
| RMSE | 2.12 | 2.36 | 1.18 | 1.56 | | 1.14 | 2.20 |
| Theil's U | 1.38 | 1.54 | 0.77 | 1.01 | 1 | 0.74 | 1.43 |
| DM | 2.54 | 2.33 | 0.49 | 2.67 | | | 2.35 |
| (p-value) | (0.99) | (0.99) | (0.69) | (0.99) | | | (0.99) |
| | | | *France* | | | | |
| RMSE | 0.64 | 1.31 | 0.57 | 0.54 | | 0.51 | 0.78 |
| Theil's U | 0.93 | 1.91 | 0.82 | 0.78 | 1 | 0.74 | 1.13 |
| DM | 0.85 | 3.04 | 0.58 | 0.82 | | | 1.86 |
| (p-value) | (0.80) | (1.00) | (0.72) | (0.79) | | | (0.97) |
| | | | *Greece* | | | | |
| RMSE | 2.93 | **1.65** | 3.46 | 2.50 | | 2.23 | **2.13** |
| Theil's U | 1.20 | **0.67** | 1.41 | 1.02 | 1 | 0.91 | **0.87** |
| DM | 1.92 | -0.94 | 1.47 | 0.77 | | | -0.14 |
| (p-value) | (0.97) | (0.17) | (0.93) | (0.78) | | | (0.44) |
| | | | *Italy* | | | | |
| RMSE | 1.09 | **0.86** | **0.95** | 1.10 | | 1.01 | 1.05 |
| Theil's U | 0.96 | **0.76** | **0.84** | 0.98 | 1 | 0.89 | 0.92 |
| DM | 1.23 | **-1.32** | -0.95 | 0.54 | | | 0.18 |
| (p-value) | (0.89) | **(0.09)** | (0.17) | (0.71) | | | (0.57) |
| | | | *United Kingdom* | | | | |
| RMSE | **1.53** | 2.07 | 1.99 | 1.84 | | 1.73 | 1.88 |
| Theil's U | **0.67** | 0.91 | 0.87 | 0.81 | 1 | 0.76 | 0.83 |
| DM | -0.56 | 1.15 | 0.94 | 0.39 | | | 0.59 |
| (p-value) | (0.29) | (0.88) | (0.83) | (0.65) | | | (0.72) |

Note: Bold estimates indicate a better performance than the univariate

forecasting model (see Section IV.B, including for abbreviations).

Table A5. Forecast Performance (max. four lags), 1990–2002

| | HP100rt | PF | BQ | FD2-8 | NC | AR | UR |
|---|---|---|---|---|---|---|---|
| | | | *Finland* | | | | |
| RMSE | 2.27 | 1.86 | 1.47 | 1.52 | | 1.36 | 2.24 |
| Theil's U | 1.48 | 1.21 | 0.96 | 0.99 | 1 | 0.89 | 1.46 |
| DM | 2.14 | 1.72 | 1.15 | 1.31 | | | 2.14 |
| (p-value) | (0.98) | (0.96) | (0.87) | (0.91) | | | (0.98) |
| | | | *France* | | | | |
| RMSE | 0.64 | 1.36 | 0.57 | 0.54 | | 0.51 | 0.76 |
| Theil's U | 0.93 | 1.97 | 0.82 | 0.78 | 1 | 0.74 | 1.11 |
| DM | 0.85 | 2.99 | 0.56 | 0.82 | | | 1.72 |
| (p-value) | (0.80) | (1.00) | (0.71) | (0.79) | | | (0.96) |
| | | | *Greece* | | | | |
| RMSE | 2.93 | **2.21** | 4.49 | 2.50 | | 2.23 | **2.13** |
| Theil's U | 1.20 | **0.90** | 1.84 | 1.02 | 1 | 0.91 | **0.87** |
| DM | 1.92 | -0.03 | 2.23 | 0.77 | | | -0.14 |
| (p-value) | (0.97) | (0.49) | (0.99) | (0.78) | | | (0.44) |
| | | | *Italy* | | | | |
| RMSE | **0.97** | **0.96** | **0.95** | 1.39 | | 1.01 | 1.05 |
| Theil's U | **0.86** | **0.85** | **0.84** | 1.22 | 1 | 0.89 | 0.92 |
| DM | -0.45 | -0.56 | -0.56 | 1.58 | | | 0.18 |
| (p-value) | (0.33) | (0.29) | (0.29) | (0.94) | | | (0.57) |
| | | | *United Kingdom* | | | | |
| RMSE | **1.53** | 2.06 | 1.99 | 1.84 | | 1.73 | 1.88 |
| Theil's U | **0.67** | 0.90 | 0.87 | 0.81 | 1 | 0.76 | 0.83 |
| DM | -0.56 | 1.13 | 0.94 | 0.39 | | | 0.59 |
| (p-value) | (0.29) | (0.87) | (0.83) | (0.65) | | | (0.72) |

Note: Bold estimates indicate a better performance than the univariate

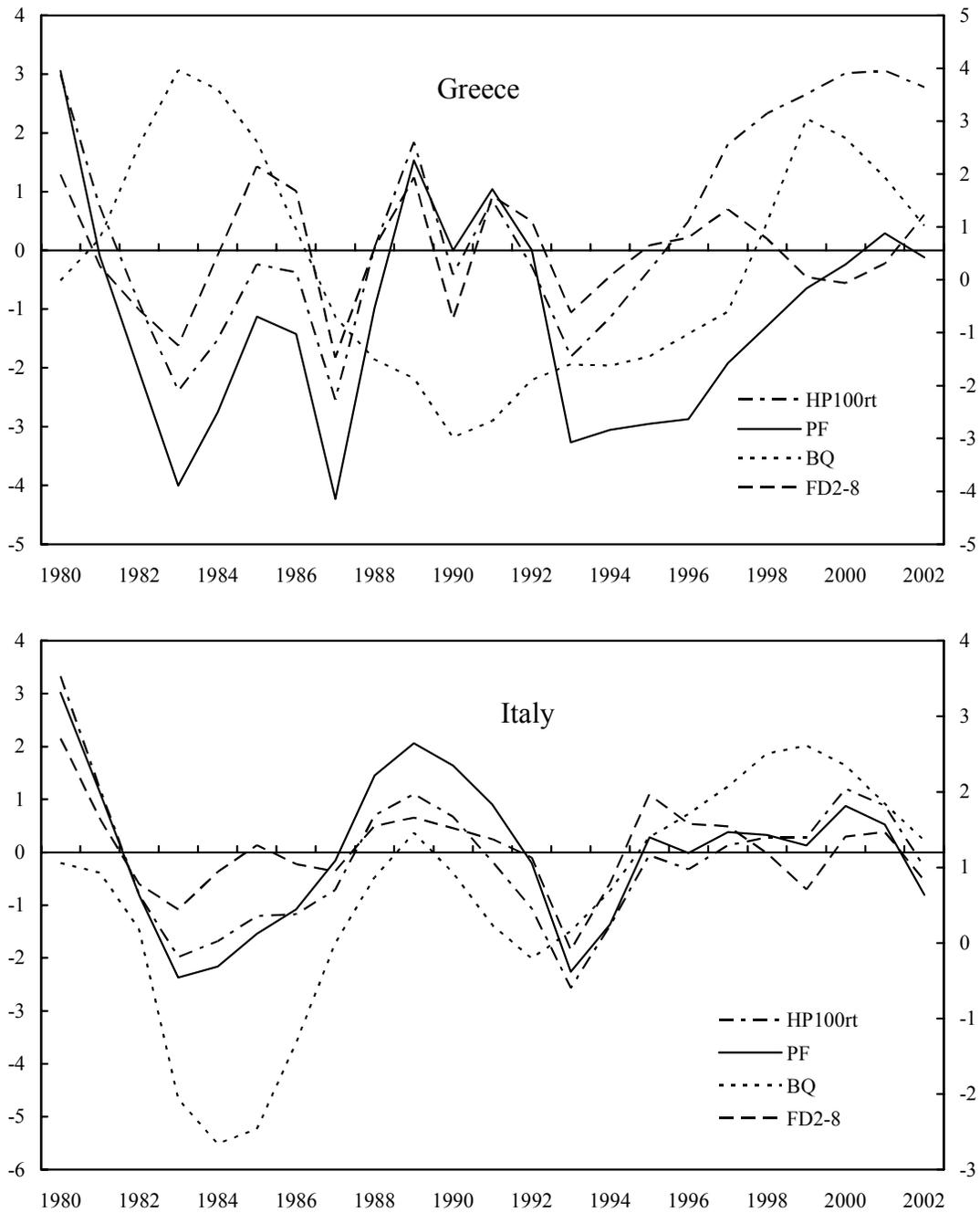forecasting model (see Section IV.B, including for abbreviations).

## III. FIGURES

### Figure A1. Output Gap Measures, 1980-2002
(In percent of potential)

Finland

HP100rt
PF
BQ
FD2-8

HP20
PF
BQ
FD2-8

France
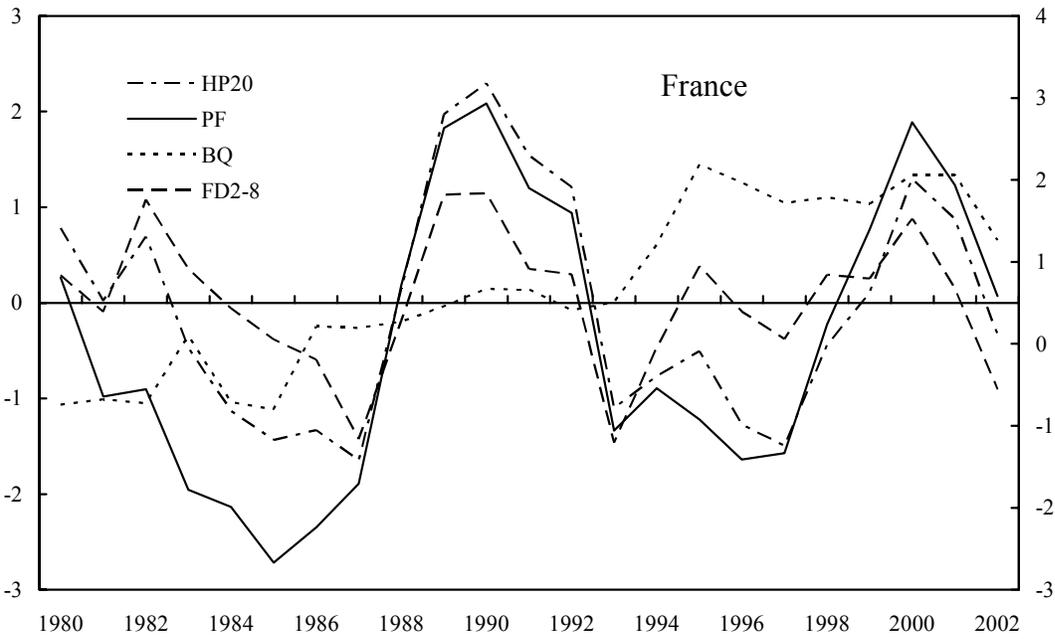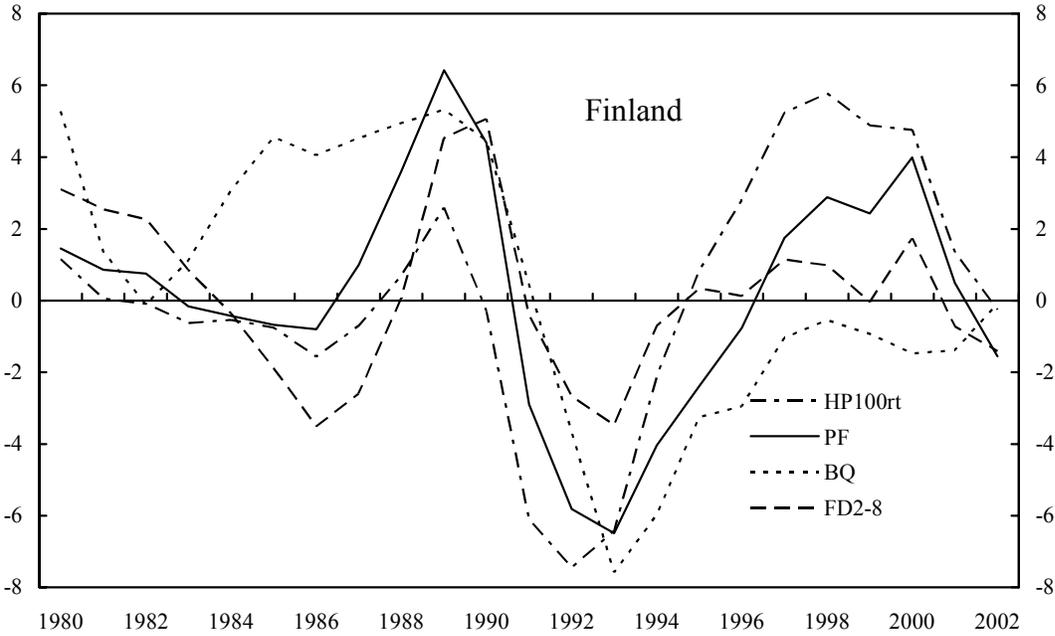
Sources: IMF *International Financial Statistics (IFS)*, European Comission; and author's calculations.

Figure A2. Output Gap Measures, 1980-2002
(In percent of potential)



Sources: IMF *IFS*, European Comission; and author's calculations.

Figure A3. Output Gap Measures, 1980-2002
(In percent of potential)

United Kingdom

HP100rt

PF

BQ

FD2-8
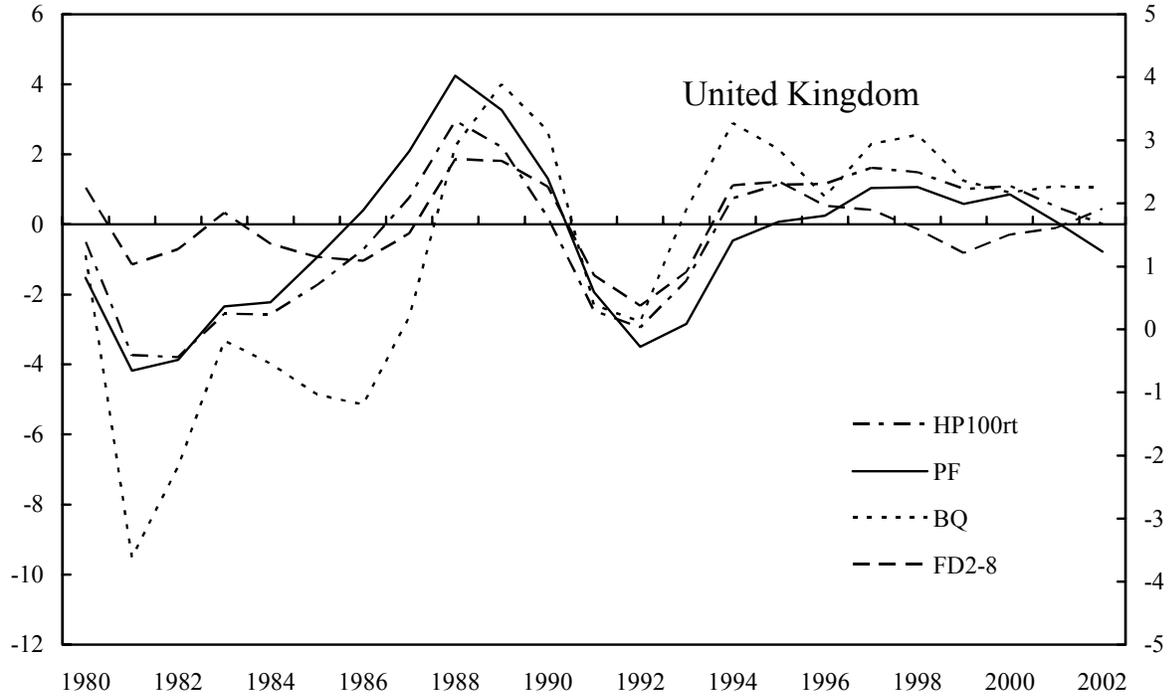
Sources:  IMF *IFS* , European Comission; and author's calculations.

**REFERENCES**

Artis, M., M. Marcellino, and T. Proietti, 2002, "Dating the Euro Area Business Cycle," EUI Working Paper ECO No. 2002/24 (Florence: European University Institute).

Artus, J.R., 1977, "Measures of Potential Output in Manufacturing for Eight Industrial Countries, 1955–78," *Staff Papers*, International Monetary Fund, Vol. 24, pp. 1–35.

Baxter, M., and R. King, 1999, "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics*, Vol. 81, pp. 575–93.

Begg, D., F. Canova, P. De Grauwe, A. Fatás, and P. Lane, 2002, *MECB Update December 2002* (London: Centre for Economic Policy Research).

Berger, H., and A. Billmeier, 2003, "Estimating the Output Gap in Finland" in *Finland—Selected Issues*, IMF Country Report 03/326 (Washington: International Monetary Fund).

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*, Vol. 7, pp. 151–74.

Billmeier, A., 2004, "Measuring a Roller-Coaster: Evidence on the Finnish Output Gap," IMF Working Paper 04/57 (Washington: International Monetary Fund).

Blanchard, O., and D. Quah, 1989, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, Vol. 79, 655–73.

Burns, A.F., and W.C. Mitchell, 1946, *Measuring Business Cycles* (New York: National Bureau of Economic Research).

Canova, F., 1999, "Does Detrending Matter for the Determination of the Reference Cycles and Selection of Turning Points," *Economic Journal*, Vol. 109, pp. 126–50.

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

Clarida, R.H., and J. Gali, 1994, "Sources of Real Exchange rate Fluctuations: How Important are Nominal Shocks?" *The Carnegie-Rochester Conference Series on Public Policy*, Vol. 41, pp. 1–56.

Cogley, T., and J. Nason, 1995, "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implication for Business Cycle Research," *Journal of Economic Dynamics and Control*, Vol. 19, pp. 253–78.

Corbae, D., and S. Ouliaris, 2002, "Extracting Cycles from Non-Stationary Data," (unpublished; Austin and Singapore: University of Texas and National University of Singapore).

———, and P.C.B. Phillips, 2002, "Band Spectral Regression with Trending Data," *Econometrica*, Vol. 70, pp. 1067–109.

Cotis, J.-P., J. Elmeskov, and A. Mourougame, *forthcoming*, "Estimates of Potential Output: Benefits and Pitfalls from a Policy Perspective," in *The Euro Area Business Cycle: Stylized Facts and Measurement Issues*, ed. by L. Reichlin (London: Centre for Economic Policy Research).

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

Denis, C., K. Mc Morrow, and W. Roeger, 2002, "Production function approach to calculating potential growth and output gaps—estimates for the EU Member States and the US," European Commission Economic Papers No. 176 (Brussels: European Commission).

Diebold, F.X., 2001, *Elements of Forecasting*, (Cincinnati: South-Western, 2nd ed.).

———, and R.S. Mariano, 1995, "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, Vol. 13(3), pp. 134–44.

Eleftheriou, M., 2003, "On the Robustness of the 'Taylor rule' in the EMU," EUI Working Paper ECO No. 2003/17 (San Domenico: European University Institute).

European Commission, 2001, *Report on Potential Output and the Output Gap*, available at http://europa.eu.int/comm/ economy_finance/epc/documents/finaloutput_en.pdf.

Everaert, L., and F. Nadal De Simone, 2003, "Capital Operating Time and Total Factor Productivity," IMF Working Paper 03/128 (Washington: International Monetary Fund).

Feldman, R. A., and others, 2003, *Finland—Staff Report for the 2003 Article IV Consultation*, IMF Staff Country Report No. 03/325 (Washington: International Monetary Fund).

Gerlach, S., and L.E.O. Svensson, 2003, "Money and inflation in the euro area: A case for monetary indicators?" *Journal of Monetary Economics*, Vol. 50, pp. 1649–72.

Harvey, A.C., and A. Jaeger, 1993, "Detrending, Stylized Facts and the Business Cycle," *Journal of Applied Econometrics*, Vol. 8, pp. 231–47.

Hodrick, R.J., and E.C. Prescott, 1997, "Post-war U.S. business cycles: An empirical Investigation," *Journal of Money, Credit, and Banking*, Vol. 29, pp. 1–16.

Inoue, A., and L. Kilian, 2002, "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Chose?" CEPR Discussion Paper No. 3671 (London: Centre for Economic Policy Research).

King, R.G., and S. Rebelo, 1993, "Low Frequency Filtering and Real Business Cycles," *Journal of Economic Dynamics and Control*, Vol. 17, pp. 207–31.

Kuttner, K.N., 1994, "Estimating Potential Output as a Latent Variable," *Journal of Business & Economic Statistics*, Vol. 12, pp. 361–8.

McCracken, M.W., 2000, "Robust out-of-sample inference", *Journal of Econometrics*, Vol. 99, pp. 195–223.

Meese, R.A., and K. Rogoff, 1983, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out-of-Sample?" *Journal of International Economics*, Vol. 14, pp. 3–24.

Okun, A. M., 1962, "Potential GDP: Its Measurement and Significations," Cowles Foundation Paper 190, *1962 Proceedings of the Business and Economic Statistics Section of the American Statistical Association*. Reprinted in A. M. Okun, 1970, *The Political Economy of Prosperity*, pp. 132–45 (New York: Norton).

Orphanides, A., 2001, "Monetary Policy Rules Based on Real-Time Data," *American Economic Review*, Vol. 91, pp. 964–985.

———, and S. van Norden, 2002, "The Unreliability of Output Gap Estimates in Real Time," *Review of Economics and Statistics*, Vol. 84, pp. 569–83.

Planas, C., and A. Rossi, 2003, "Program GAP - Version 2.2, Technical Appendix" (unpublished; Ispra: European Commission, Joint Research Centre).

Proietti, T., A. Musso, and T. Westermann, "Estimating Potential Output and the Output Gap for the Euro Area: a Model-Based Production Function Approach," EUI Working Paper ECO No. 2002/9 (Florence: European University Institute).

Ravn, M.O., and H. Uhlig, 2002, "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics*, Vol. 84, pp. 371–80.

Robinson, T., A. Stone, and M. van Zyl, 2003, "The Real-Time Forecasting Performance of Phillips Curves," Reserve Bank of Australia Research Discussion Paper 2003-12 (Sidney: Reserve Bank of Australia).

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

Scacciavillani, F., and P. Swagel, 1999, "Measures of Potential Output: An Application to Israel," IMF Working Paper 99/96 (Washington, DC: International Monetary Fund).

Smets, F., 1998, "Output Gap Uncertainty: Does It Matter for the Taylor Rule?" In *Monetary Policy under Uncertainty*, ed. by B. Hunt and A. Orr (Wellington: Reserve Bank of New Zealand).

Stock, J.H., and M.W. Watson, 1999, "Forecasting Inflation", *Journal of Monetary Economics*, Vol. 44, pp. 293–335.

———, 1991, "A probability model of the coincident economic indicators," in *Leasing Economic Indicators: New Approaches and Forecasting Records,* ed. by K. Lahiri and G.H. Moore (New York: Cambridge University Press).

———, 1989, "New indexes of coincident and leading indicators," in: *NBER Macroeconomics Annual,* ed. by O.J. Blanchard and S. Fischer (Cambridge: MIT Press).

Svensson, L.E.O., 2003, "What is Wrong with Taylor Rules? Using Judgement in Monetary Policy Rules through Targeting Rules," *Journal of Economic Literature*, Vol. 41, pp. 426–77.

———, 1999, "Inflation Targeting as a Monetary Policy Rule," *Journal of Monetary Economics*, Vol. 43, pp. 607–54.

Taylor, J.B., 1993, "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Series on Public Policy*, Vol. 39, pp. 195–214.

Theil, H., 1971, *Principles of Econometrics*, (New York: Wiley).

West, K.D., 1996, "Asymptotic Inference about Predictive Ability," *Econometrica*, Vol. 64, pp. 1067–84.