Finnish Economic Papers – Volume 19 – Number 2 – Autumn 2006

MEASURING A ROLLER COASTER: EVIDENCE ON THE FINNISH OUTPUT GAP*

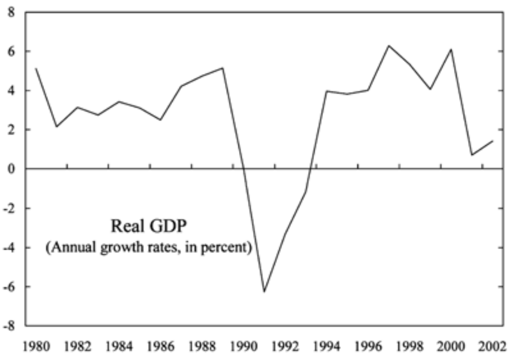
ANDREAS BILLMEIER

Middle East and Central Asia Department, International Monetary Fund, 700 19th Street, NW, Washington, DC 20431, USA

The output gap – which measures the deviation of actual output from its potential – is frequently used as an indicator of slack in an economy. This article estimates the Finnish output gap using various empirical methods. It evaluates these methods against economic history and each other by a simulated out-of-sample forecasting exercise for Finnish CPI inflation. Only two gap measures, stemming from a frequency-domain approach and the Blanchard-Quah decomposition, perform better than the naïve prediction of no change in inflation – but do not improve upon a simple autoregressive forecast. The pronounced volatility of output in Finland makes it particularly difficult to estimate potential output, producing considerable uncertainty about the size (and sign) of the gap. (JEL: E31, E32, E37)

1. Introduction and overview

This article assesses and compares various output gap measures for policymaking purposes in Finland during the turbulent 1990s. Assessing the degree of slack in the economy is important, and the output gap – which measures the deviation of GDP from its potential – is a frequently used indicator for this purpose, for several reasons. First, variations in output – which have been particularly stark in Finland – have distinct implications for inflationary pressures in the economy when assessed relative to potential. Second, the size of the output gap, as an important component of calculating the "structural fiscal balance," helps to gauge the thrust of fiscal policy. Third, the magnitude of the output gap is relevant for assessing economic growth – that is, whether variations in actual growth



Source: IMF

* *Correspondence email: abillmeier@imf.org. The author would like to thank Robert A. Feldman, Lusine Lusinyan, two anonymous referees, and the editor for very helpful comments, as well as Hamid Faruqee, Francisco Nadal de Simone, and especially Helge Berger for extensive discussions. This paper should not be reported as representing the views of the IMF. The views expressed are those of the author and do not necessarily represent those of the IMF or IMF policy.*

can be attributed to cyclical factors (such as slow growth in trading partner economies) or to a longer-term change in potential growth. In this article, we explore the first and, to a lesser extent, the third aspects in the Finnish context. It is not our intention to find the "best" output gap measure – excellence lies in the eye of the beholder. Instead, this article highlights the complexities and pitfalls associated with the output gap concept and its use in policymaking and policy-prescribing.

Defined as the difference between actual and unobservable potential output, the output gap is itself an unobserved variable. This article discusses various measures of the Finnish output gap which fall into two broad sets of methodological approaches to determining potential output – statistical filters and model-based approaches. Both types are discussed against the background of Finland's economic roller coaster ride during the 1990s and evaluated econometrically in an inflation-forecasting framework for the period 1990–2002.[1]

Econometric evaluation is carried out by a simulated out-of-sample forecasting methodology in a Phillips curve framework, which links inflation to a measure of economic slack. If the output gap is to some extent a measure of domestic inflationary pressures, then a naïve forecast (of "no change") for Finnish inflation could be improved by taking into account the information stemming from the output gap. For the period 1960–2002, models are estimated recursively with annual (real-time) data up to the forecast period 1990–2002, and the one-year-ahead predicted inflation is compared with actual inflation.[2] On the basis of this evaluation, only the frequency domain filter and the Blanchard-Quah decomposition help forecasting consumer price index (CPI) inflation, whereas the other measures of the output gap do not improve the naïve forecast. Both measures, however, perform worse than an autoregressive inflation model.

The weak performance of the output gap measures in the forecasting exercise is clearly related to the pronounced volatility of actual output in Finland, which makes it particularly difficult to estimate potential output. Moreover, distinct problems are associated with each of the two groups of approaches to estimating the output gap. Purely statistical measures are often subject to exogenous assumptions on the flexibility of the underlying trend and can, especially after sharp economic turns, misstate potential output. In the case of Finland, the strong expansion in the second half of the 1990s – doubtlessly related to positive structural shifts such as the rise of the ICT sector – would be seen as purely cyclical fluctuations in a statistical sense in some analyses of the statistical properties of the GDP series. On the other hand, model-based approaches, such as the production function approach, avoid the problem of correctly separating trend output from observed output, but they must rely on estimates of labor input at full employment and the stock of effective capital input in production as components of a stable production function. Both estimates are not without problem in the Finnish context. While it is clear that much of the increase in unemployment is structural, it is nevertheless difficult to gauge the extent to which the sudden surge in measured unemployment also caused the natural rate of unemployment to increase; and the collapse of trade with Russia in the early 1990s raises the possibility that a part of the capital stock had become obsolete.

In the remainder of this article, Section 2 reviews briefly the relevant literature and Section 3 contains a description of the various output gap measures. Section 4 compares them in the context of a forecast model for Finnish CPI inflation. Section 5 concludes.

---

[1] During the past 15 years, the Finnish economy has experienced a roller coaster ride including (1) a very sharp recession in the early 1990s, during which the unemployment rate soared from 3.2 percent in 1990 to 16.4 percent only three years later; (2) a strong boom period in the second half of the 1990s – with real GDP growth averaging about 5 percent in 1994–2000, led by the information and communication technology (ICT) sector, especially in the second half of the period; and, lately, (3) a significant slowdown in economic growth. The coefficient of variation of quarterly real GDP growth rates since 1990 is twice as high in Finland as in the euro area as a whole or member countries thereof (e.g., Greece).

[2] Data frequency is constrained to annual observations due to the inclusion of the European Commission's output gap approach based on the production function approach. With this approach, some data – e.g., for the capital stock – are only available at annual frequency.

2. Review of the literature

General research on the output gap has been abundant; see, for example, Kuttner (1994) for a brief review. Broadly speaking, two main approaches have been followed in the literature to estimate potential output and the output gap. The first is based on the statistical properties of the underlying GDP series. Under the second approach, potential output is estimated on the basis of an economic model. As argued in Scacciavillani and Swagel (1999), these different techniques can be viewed as akin to different economic concepts of potential output.

Under the first approach, potential output is driven by productivity movements, and temporary deviations of actual output from potential result from private agents' decisions to reallocate resources in response to these shocks. Given this (neoclassical) reasoning, potential output coincides with the underlying trend of actual output, and the challenge in estimating the output gap is to separate longer-run changes in the trend from short-lived (temporary) movements around potential. The univariate statistical measures in this article can be traced back to the familiar contribution by Hodrick and Prescott (1997) and to a frequency-domain filter proposed by Corbae and Ouliaris (2002).[3]

Under the second approach – somewhat closer to the Keynesian tradition – business cycle swings and, hence, the gap between actual and potential output reflect demand-determined actual output fluctuating around a slowly moving level of aggregate supply. The corresponding measures of the output gap account for underemployed resources – in particular in the labor market – building on models that describe relevant aspects of the economy. The model-based approaches of Section 3 relate to Blanchard and Quah (1989), and to the large strand of literature that 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). In this article, the implementation of the production function approach (in section 3.2.2) instead follows closely the latent variable approach developed in Kuttner (1994), and further refined by the European Commission; see, for example, Denis, Mc Morrow, and Roeger (2002).

There are good accounts of the turbulent period in Finland during the 1990s. A number of observers have linked the economic downturn to fast financial liberalization without an adequate regulatory framework in place. In that context, Ahtiala (2006) and Halme (2002) provide overviews of the crisis and derive some lessons for financial reform. Jonung and Hagberg (2005) compare the cost of the 1990s crisis with that of other crises in Finland and Sweden.

However, relatively little empirical work has been done comparing different output gap estimation methods. Brunila, Hukkinen, and Tujula (1999) briefly describe the approaches used by the Bank of Finland when assessing cyclically adjusted budget measures: a Hodrick-Prescott filter, and the production function approach implicit in the Bank's econometric model BOF5 (which does not incorporate such considerations as the natural rate of unemployment). Other work on potential output and the output gap in Finland includes Gylfason (1998), who uses a broken linear trend to account for a structural shift toward slower economic growth in the early 1970s, and Rasi and Viikari (1998) who apply an unobserved components method developed by Apel and Jansson (1999a,b) to the Finnish data (potential output and the natural rate of unemployment are the unobserved variables estimated simultaneously).

In Section 4, to compare the various output gap estimates in an inflation-forecasting framework, we use a model similar to the one employed by Stock and Watson (1999), who investigate forecasts of U.S. inflation at the 12-month horizon, using information from 168 additional indicators of economic activity, including output gap estimates.[4]

---

[3] *See the working paper version, Billmeier (2004b), for additional evidence on simple arithmetic detrending and the Beveridge-Nelson (1981) decomposition.*

[4] *More recent contributions include Orphanides and van Norden (2003) and Robinson, Stone, and van Zyl (2003); both evaluate the reliability of inflation forecasts based on real-time output gap estimates for the United States and Australia, respectively.*

3. Measuring potential output and the output gap

3.1 Univariate statistical measures of potential output

3.1.1 The Hodrick-Prescott (HP) filter

The HP filter is probably the best known and most widely used statistical filter to obtain a smooth estimate of the long-term trend component of a macroeconomic series. Its prominence 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 official National Bureau of Economic Research (NBER) 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:

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

The penalty parameter, λ, controls the smoothness of the series by setting the ratio of the variance of the cyclical component and the variance of the actual series. The standard value in the literature is $\lambda = 100$ for annual data, which is also assumed here as a base case. 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 cyclicality for integrated series, the neglect of structural breaks and shifts, as well as somewhat arbitrary choice of λ.[5] All these criticisms are certainly of relevance in the case of Finland: real GDP is likely to be integrated; there are clearly structural breaks, which would be removed from the trend component approximating potential output by the filtering process; and the assumption on λ has an impact on the decomposition – for instance, the extent to which the boom in the second half of the 1990s is viewed as having had an effect on the long-run potential. The most important drawback, however, stems from the end-of-sample bias. This bias owes to the symmetric treatment of the trending across the sample and the different constraints that apply within the sample and at its ends.[6] One way to deal with the bias in practice has been to extend the observation period by forecasting. These last two issues (smoothness of the trend, end-sample bias) are considered in turn.

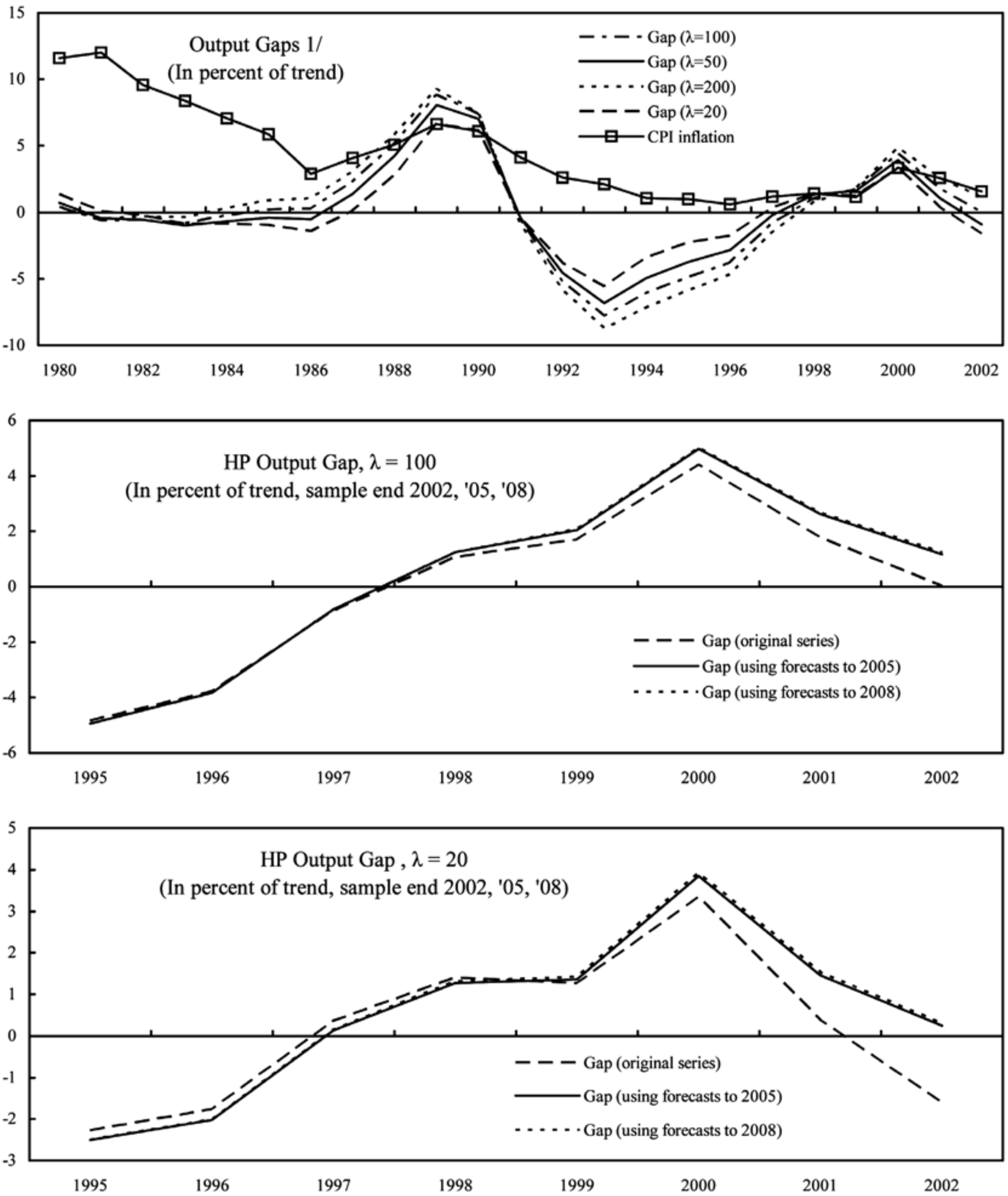
While various HP filters find a qualitatively similar pattern of the output gap, the assumption on the smoothness of the trend has strong implications for the magnitude of the gap, in particular at the end of the observation period. In Figure 1, top panel, the results of HP-filtering Finnish real GDP – for various parameter values for λ – are shown for the period 1980–2002 (but estimated over 1960–2002). Trend estimates are provided for λ = 20, 50, 100, and 200, with a higher parameter value indicating a smoother trend. All estimated gaps are close to zero on average, and – unsurprisingly – agree on the peak and trough dates. However, the size of the gap varies significantly, in particular at peaks/troughs: for instance, estimates for the gap in the trough year 1993 vary between approximately –5 percent and –9 percent of trend GDP. All filters indicate that the gap was closed as early as 1997. However, the estimates indicate also a larger (positive) output gap during the boom in the late 1980s than during the more recent expansion. The estimates of the gap in 2002, the last data point, vary widely: the smoothest trend indicates that actual GDP was above trend by 1.1 percent of potential GDP, whereas the least smooth trend results in actual output being 1.6 percent below potential. The standard assumption of $\lambda = 100$ yields a gap of approximately zero. The fact that the HP filters do not provide a set of estimates of the output gap which is uniformly above (or below) zero – independently of the assumed trend smoothness – is clearly unsatisfactory.

The severity of the endpoint bias of the HP filter, instead, is underscored by simply adding

[5] *See, e.g., Harvey and Jaeger (1993) for an overview of the shortcomings. Ross and Ubide (2001) discuss alternative approaches to determine the parameter λ endogenously.*

[6] *In equation (1), the second difference of the trend is not defined around the first and the final observation, implying different summation bounds for value function and punishment term.*

Figure 1. Finland: Hodrick-Prescott (HP) Detrending and Endpoint Problem, 1980–2002 1/



Sources: IMF, *International Financial Statistics (IFS), World Economic Outlook (WEO);* and author's calculations.
1/ The last observation of the original time series is 2002.

a few forecasted values at the end of the observation period of the variable to be filtered. Values for real GDP were added using the IMF's *World Economic Outlook (WEO)* database forecast, in the first step until 2005, and until 2008 in the second step. The premise of a medium-term recovery of the Finnish economy has a remarkable effect on the estimates of the output gap in 2002. For all parameter values, gap estimates based on the extended series are higher than for the original series (middle and bottom panel of Figure 1). In particular, for $\lambda = 100$, the output gap is positive, indicating real GDP above potential by 1.2 percent of trend, as opposed to a closed gap for the series ending in 2002. The swing of the output gap as a result of the two extensions of the estimation period is even more pronounced for the less smooth trend, $\lambda = 20$.[7] In this latter case, both estimates based on the extended series indicate a positive output gap on the order of 0.3 percent of GDP, whereas the gap using the original series is equal to –1.6 percent of trend GDP. In sum, the HP estimation approach creates a trade-off between the endpoint bias problem – using the original series – and the reliability of out-of-sample forecasts of the underlying variable.[8]

3.1.2 A frequency domain (FD) filter

Economic fluctuations occur at different frequencies (displaying, for instance, seasonal or business cycle duration). Starting from the classical finding contained in Burns and Mitchell (1946) that the duration of business cycles takes between 1.5 and 8 years, the approach to extracting those cycles from a *stationary* time series is straightforward from the frequency domain perspective. Using an exact band-pass filter, the original series can be filtered such that fluctuations below or above a certain frequency (the business cycle) are eliminated.

For estimation purposes, however, these filters are usually spelled out in the time domain, since *nonstationary* series – such as real GDP – could not be handled by traditional frequency domain approaches, and therefore require pre-filtering to render them stationary.[9] Corbae and Ouliaris (2002) indicate a way to avoid pre-filtering of the series in the time domain and provide a consistent band-pass filter for nonstationary data in the frequency domain that does not involve a loss of observations at either end – a property highly relevant for policymaking.[10]

Figure 2 presents all calculated output gap measures, including the result from the FD approach. The FD method reproduces the Finnish boom and bust period around 1990 well. In particular, the overheating in the late 1980s is associated with a substantial positive output gap, which reaches 5 percent of (potential) GDP. After the trough in 1993, the gap closed quite rapidly according to this method, turning (marginally) positive over the period 1995–2000. The ensuing growth slowdown resulted in a negative output gap in the last two years, reaching –1.4 percent of GDP in 2002. A striking feature of this filter is the period between 1984 and 1988.

The trough attributed to these years is just as deep as the one during the crisis of the early 1990s – a result not mirrored by any other gap measure. Moreover, the FD approach associates the period of overheating at the end of the 1980s with a higher output gap peak (in absolute terms) of 5.1 percent of trend GDP than the trough in 1993 (–3.5 percent of trend GDP). Regarding the duration of the economic downturn in the early 1990s, the FD filter results in a negative gap limited to the period 1991–94, contrary to, for example, the results stemming from the HP filter (where the duration is much longer). Moreover, according to the FD filter, the output gap stayed small during the second half of the 1990s, attributing much of the high actual growth rates during the late 1990s to the underlying trend, and little to cyclical factors.

While the major advantage of the frequency domain approach – and, indeed, other statistical methods – is their simplicity, they are subject to the criticism of lacking foundation in economic theory. Thus, the next section turns to theory-based models of trend GDP.

---

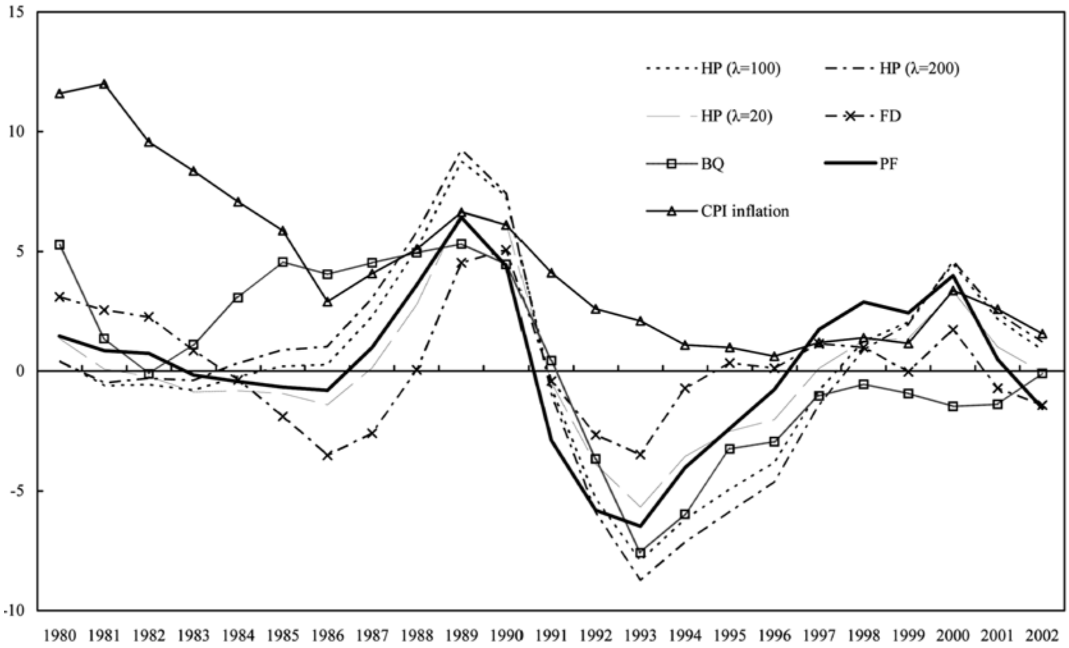
[7] No particular meaning is attached to the value $\lambda = 20$, other than in terms of a simple sensitivity check. For the United States, Ravn and Uhlig (2002) suggest a value of 6.25.

[8] Given the very similar results for the time series ending in 2005 and 2008, the 2005 endpoint results are reported in what follows – with the exception of the evaluation exercise, where a real-time measure is constructed to reflect data availability at the time the inflation forecast is made.

[9] See, e.g., Baxter and King (1999).

[10] See Corbae, Ouliaris, and Phillips (2002) for the analysis of the asymptotic case. With annual data, a periodicity for the business cycle between two and eight years has been assumed. Somewhat longer and shorter cycles yield broadly similar results; see Billmeier (2004a).

Figure 2. Finland: Comparing Measures of the Output Gap (In percent of potential GDP) 1/ 2/



Sources: IMF; *IFS, WEO;* and author's calculations.
1/ HP, Hodrick-Prescott; PF, production function; FP, frequency domain; BQ, Blanchard-Quah.
2/ The results from the Hodrick-Prescott filter are based on the extended data series until 2005.

3.2 Theory-based measures

3.2.1 The Blanchard-Quah (BQ) approach

The appeal of the approach by Blanchard and Quah (1989) to the identification of structural shocks in a vector autoregression (VAR) stems from its compatibility with a wide array of theoretical models. 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. To identify these disturbances, the following assumptions are made: no disturbance has long-run effects on the time series employed in the estimation. Furthermore, disturbances to (the growth rate of) real GDP might have long-run effects on the (cumulative) *level* of both series, while shocks to the unemployment rate do not have long-run effects on the *level* of output. These assumptions technically identify the shocks and it seems natural to label them as supply and demand shocks.[11]

Here, potential output is associated with cumulated supply shocks, whereas the output gap reflects cyclical (temporary) swings in aggregate demand. This approach, hence, 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 (unemployment rate) is not a good indicator of the cyclical behavior of output. Finally, while the amount of variables determines – under the orthogonality assumption – the number of shocks present in the system,

[11] Blanchard and Quah (1989) show that small violations of the identification scheme (e.g. lasting nominal effects on output stemming from a wealth effect) are of minor consequence.

there are clearly shocks that have a supply as well as a demand component, for instance, public infrastructure investment.

The estimated output gap based on the BQ decomposition is – on some but not all accounts – similar to those stemming from other approaches (see Figure 2).[12] For example, the measure also echoes the overheating of the Finnish economy in the late 1980s. However, according to this method, actual output remained markedly above potential in the five years between 1985 and 1990 – that is, starting several years earlier than indicated by most other gap measures. The crisis period in the early 1990s is characterized by a large trough, which amounts to about eight percent of potential GDP in 1993. The striking feature, however, is the subsequent path. Unlike any other gap measure, the output gap based on the BQ decomposition remains negative until the end of the sample period – implying that the Finnish economy was able to achieve strong growth almost without inflationary pressure during the second half of the 1990s as the output gap had not yet been closed. In other words, the BQ decomposition attributes much of the strong real growth over that period to high growth of potential output.

3.2.2 The production function approach

Generic set up. The production function (PF) approach describes a functional relationship between output and factor inputs. Output is at its potential if the rates of capacity utilization are normal; that is, labor input is consistent with the natural rate of unemployment, and 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:

$$(2) \quad Y_t = TFP_t K_t^{\alpha} L_t^{\beta}.$$

Constant returns to scale imply $\alpha + \beta = 1$. Under perfect competition, α 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:

$$(3) \quad tfp_t = y_t - \alpha k_t - (1-\alpha) l_t,$$

where variables in lowercase are in logs. Log trend *TFP, tfp**, is then obtained by appropriately smoothing this residual series. Potential labor input, L_t^*, is taken to be the level of employment consistent with the (time-varying) natural rate of unemployment UR_t^*:

$$(4) \quad L_t^* = LF_t\left(1 - UR_t^*\right),$$

where LF_t is the labor force. The natural rate of unemployment can be derived in a number of ways – for example, by HP-filtering the observed unemployment rate, or by assessing it as a latent variable. Potential output can hence be expressed as:[13]

$$(5) \quad y_t^* = \alpha k_t + (1-\alpha) l_t^* + tfp_t^*.$$

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, a feature particularly relevant in a country like Finland. For instance, the dynamic growth of the ICT sector during the second half of the 1990s has added substantially to potential growth from a productivity point of view.[14] Moreover, the strong movements of the unemployment rate since the crisis in the early 1990s convey valu-

[12] *The VAR model underlying the estimation includes, in addition to a constant, three lags of the endogenous variables, as indicated by information criteria. No residual autocorrelation was present in the specification chosen. Other specifications were tested, but dismissed on statistical grounds. Among these, a system including the real effective exchange rate instead of the unemployment rate to reflect foreign demand for exports.*

[13] *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).*

[14] *See Wagner (2001) and Jalava and Pohjola (2001), who find that the importance of "multifactor productivity" – a concept similar to total factor productivity in the present paper – almost doubled in the second half of the 1990s. Van Ark (2001) provides an international comparison of the ICT sector's contribution to output and productivity growth during the 1990s.*

able information on labor market conditions. An important feature of this approach is the reliance on filtered series, such as the trend total factor productivity and the natural rate of unemployment, which adds an element of discretion.[15]

An application to Finland. The first step in the calculation of the output gap using the production function approach emphasizes the derivation of the NAWRU – nonaccelerating wage inflation rate of unemployment – as a latent variable using the framework adopted recently by the European Commission, following Kuttner (1994). In a second step, the NAWRU serves as input in estimating potential output using a simple Cobb-Douglas specification.[16]

In the first step, the Finnish natural rate of unemployment – here the NAWRU – is computed applying a Kalman filter to the observable unemployment rate to extract the cyclical component. The procedure employs a bivariate state space model, where the observables "unemployment rate" and "change in wage inflation" (the second differences of wages) are the 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 (autoregressive moving average) process. The trend unemployment rate, in turn, serves to determine the (full-employment) stock of labor entering the production function.[17]

Figure 3 (top panel) presents the estimated trend unemployment rate tracking the observed unemployment rate with a small lag. In particular, the highest rate of trend unemployment is achieved in 1996, at 14.7 percent. Since then, and until 2001, the trend unemployment rate remained above the observed rate. Under the optimal model structure, wage inflation follows an ARMA(2,3) process, and no additional regressors are employed. The key implications of these assumptions are (1) a negative and significant coefficient to the contemporaneous cyclical unemployment component in the Phillips curve, reflecting the dampening effect of an adverse economic environment on the size of wage increases; and (2) an estimate of the NAWRU in 2002 of 8.3 percent – compared with the official unemployment rate of 9.1 percent.[18]

In the second step, the natural rate of unemployment serves as input in a production function as described above. The resulting measure of the output gap combines many of the features of other approaches (Figure 3, lower panel). A gap of almost equal magnitude (approximately 6.4 percent of trend GDP) is associated with the peak in 1989 and the trough in 1993 – a result somewhat more intuitive in terms of size and symmetry than that provided by the Blanchard-Quah decomposition, which attributes a rather small gap to the overheated economy in 1989. Furthermore, the gap – according to the PF measure – was closed by 1997, turned positive, but slipped back into negative territory in 2002.

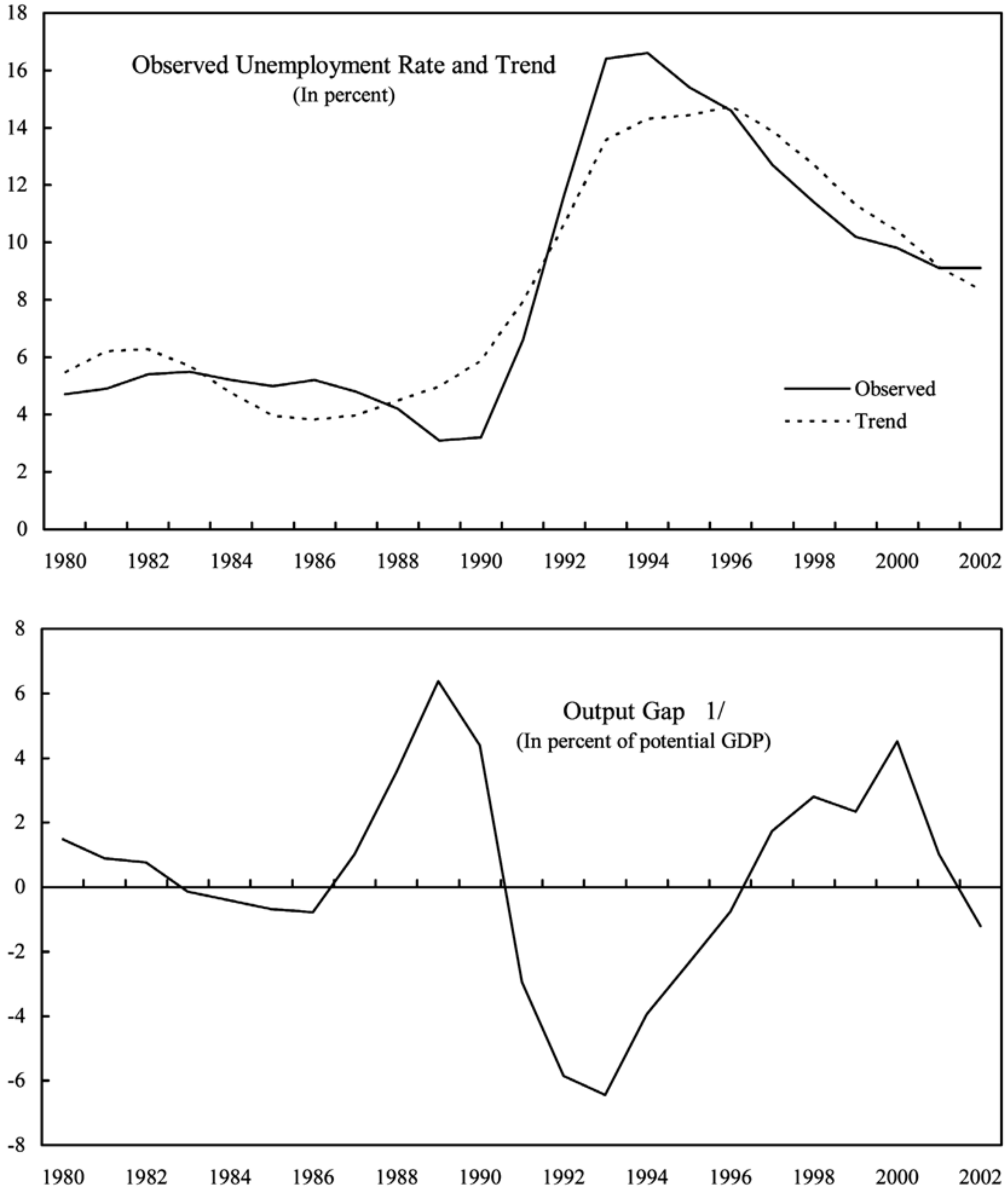
[15] *In the simplest case, potential output is a linear combination of HP-filtered series, but the approach can also be implemented flexibly by using more sophisticated filtering procedures. Other important shortcomings of the approach include the dependence on a number of crucial assumptions – for example, (constant) shares of capital and labor, and the functional form of the production relationship in a more general sense (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.*

[16] *From a conceptual point of view, this approach rests on the assumption that a natural rate of unemployment exists – in other words, that the Phillips curve is partly vertical. Supportive evidence for this assumption is omitted here due to space constraints; see the appendix of the working paper version, Billmeier (2004b).*

[17] *See Denis, Mc Morrow, and Roeger (2002), and Planas and Rossi (2003). In the European Commission's work, this methodology substitutes for more "traditional" approaches – such as the HP filter – and, at the same time, unifies the Commission's efforts toward a consistent representation of business cycles across member countries. For a more detailed description of the model set-up for the Finnish case, see the appendix of Berger and Billmeier (2003).*

[18] *Recourse to additional explanatory variables – capturing either the unemployment surge in the early 1990s or particularities of the Finnish wage negotiation process – could potentially increase the level of significance of the cyclical component further. However, experiments with various variables, including labor productivity and measures of the terms of trade resulted in a deterioration of the system's performance from a statistical point of view.*

Figure 3. Finland: Production Function Approach, 1980–2002



Sources: IMF, *IFS,* European Commission; and author's calculations.
1/ Based on the ARMA(2,3) model.

Table 1. Output Gap Measures: Descriptive Statistics (1980–2002)

	Gap measure					
	HP100	HP200	HP20	FD	BQ	PF
Mean	0.1	0.1	0.1	0.2	0.4	0.2
Min	–7.9	–8.7	–5.7	–3.5	–7.6	–6.5
Max	8.8	9.2	6.7	5.1	5.3	6.4
Standard deviation	4.0	4.4	2.8	2.3	3.6	3.1
Output gap in 2002	0.9	1.2	0.0	–1.4	–0.1	–1.5

Table 2. Output Gap Measures: Correlations 1/

	Gap measure						
	HP100	HP200	HP20	FD	BN	BQ	PF
HP100	1	0.999	0.985	0.756	0.105	0.883	0.888
HP200	0.996	1	0.975	0.725	0.088	0.877	0.870
HP200	0.963	0.938	1	0.858	0.157	0.891	0.922
FD	0.562	0.516	0.745	1	0.258	0.764	0.840
BQ	0.751	0.788	0.662	0.348	0.004	1	0.648
PF	0.906	0.886	0.934	0.715	0.398	0.648	1

1/ The lower triangle of the correlation matrix gives correlations over the sample period 1980–2002. The upper triangle for the period 1990–2002.

Note: HP…, Hodrick-Prescott filter for λ = 100, 200, 20; FD, frequency domain filter; BQ, Blanchard-Quah decomposition; PF, production function approach.

While the positive gap over the period 1997–2001 seems somewhat high, given other statistics pointing to an essentially closed gap, the negative gap in 2002 is in line with general economic indicators.[19]

Tables 1 and 2 summarize descriptive statistics for the output gap estimates considered above (and reproduced in Figure 2). Overall, there are significant differences across the various methods – visible, for instance, in the 2002 gap. Most measures' averages are close to zero, with the exception of the Blanchard-Quah approach. All filters display standard deviations in the range of 2 to 4 percent and the rather high correlations between the output gaps stemming from the HP filter and the PFA are noteworthy. However, the associated estimates of the output gap in 2002 vary between –1.5 percent and 1.2 percent of potential. In other words, there is considerable uncertainty about the size of the output gap in Finland. Given prospects of low inflation and high unemployment, estimates suggesting actual output above potential seem somewhat counterintuitive.[20]

4. *Econometric evaluation of the output gap measures*

Since the output gap is often considered a useful instrument for gauging domestic inflationary pressures, the information stemming from an output gap measure should, in principle, increase the precision of inflation forecasts. To test this conjecture, inflation forecasts based on the output gap estimates are compared using a simulated out-of-sample methodology. The forecasting model is a variant of the Phillips curve:

$$\pi_{t+1} - \pi_t = \alpha + \beta(L)x_t + \gamma(L)\Delta\pi_t + e_t, \tag{6}$$

[19] *See IMF (2003). Potential TFP is derived by HP-filtering the residual. We set λ = 100; the results are rather robust to the choice of the smoothness parameter.*

[20] *See IMF (2003).*

Table 3. Output Gap Measures: Forecast Performance

	Gap measure								
	HP100	HP20	HP100rt	HP20rt	FD	BQ	PF	No change	No OG/AR
RMSE 1/	1.94	2.12	1.90	1.97	1.52	1.18	1.71		1.14
Theil's U 2/	1.27	1.38	1.24	1.29	0.99	0.77	1.11	1	0.74

1/ Root mean square error (multiplied by 100).
2/ Theil's (1971) U statistic.

Note: HP...(rt), Hodrick-Prescott (real-time) filter for λ = 100, 20; FD, frequency domain filter; BQ, Blanchard-Quah decomposition; PF, production function approach; No OG/AR, no output gap/autoregressive inflation model.

where π_{t+1} denotes the one-year-ahead inflation in the consumer price index (CPI) level at period t, π_t is the actual inflation in period t, x_t denotes the output gap measure, L is the lag operator, and e_t is an i.i.d. error term. This specification of the forecast equation mirrors a classic Phillips curve, with the output gap measure substituting for the unemployment rate.[21] 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. Second, a one-year-ahead forecast for inflation in 1990 is made. This value is compared with the actual inflation, yielding the forecast error. This exercise is computed recursively adding one year at a time – that is, for every year until 2001, the model is reestimated 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.

In order to meaningfully evaluate the HP measure, an adjustment is needed: the filter as described in the previous section 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. Hence, a new "realtime" series, HP..rt, is constructed for the forecasting evaluation assuming two parameter values – λ = 20,100. This new series consists of "last observations" – that is, real-time estimates of the underlying trend in the period when the inflation forecast is made – and reflects more appropriately the information set available at that time.

Table 3 presents two statistics: in addition to the cumulative root mean square error (RMSE), the so-called U-statistic proposed by Theil (1971) is displayed. The latter, added for ease of comparison, is defined as the RMSE of a specific inflation forecast standardized by the RMSE of the naïve forecast of "no change" in inflation. Hence, a value smaller than unity stands for a smaller RSME than under the naïve hypothesis.

From the table, a clear ranking of the various output gap measures emerges. Only the Blanchard-Quah and, to a limited extent, the frequency domain gap measures provide forecasts that are better than the naïve forecast. The addition of any other output gap measure does not improve, in a statistical sense, the forecast of Finnish inflation over the period 1990–2002. The univariate autoregressive model of Finnish inflation ("no OG/AR"), however, performs even better than both the FD and the BQ filter. The relatively low RMSE of the inflation fore-

[21] *As Stock and Watson (1999) point out, this specification assumes that (1) the inflation rate is integrated of order one (I(1)); (2) x_t is I(0); and (3) 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 Finland; see the appendix of Billmeier (2004b). While the output gap may seem to behave like an integrated process over limited periods of time, it is mean-reverting by definition.*

cast based on the BQ output gap is noteworthy. While at odds with most other gap measures in the second half of the 1990s, the (negative) BQ gap seems to capture best the domestic inflationary pressures – or better there lack thereof. Surprisingly, the RMSEs associated with the real-time measures of the HP filter are slightly smaller than those stemming from the ex-post, two-sided measures discussed in detail in the previous section. Both versions, however, cannot provide useful information for the forecast, nor can the PF approach. These results resemble what has been found in the literature. While we are not aware of a similar, Finland-specific, exercise, examples of papers that document the weak out-of-sample forecast performance of Phillips curves using an output gap measure abound.[22]

The disappointing performance of the output gap measures in improving the inflation forecast can be attributed to a variety of factors, including: (1) the high volatility of output itself, hampering the determination of a statistically satisfying measure of potential output; (2) annual data frequency. While quarterly data would have resulted in a substantially higher number of observations, and hence better model fit, other complications would have surfaced, such as the lack of intra-year observations on the capital stock (used in the production function approach), and seasonality issues; and (3) the fact that the consumer basket as measured by the consumer price index contains imported goods. With at least some pass-through, prices of these goods are likely to change to some extent if the exchange rate fluctuates – an effect not related to domestic price pressures and, hence, not captured by measures of the output gap. A similar case can be made for indirect taxation. While not caused by domestic economic developments, changes in these variables have an impact on the CPI level. To isolate these effects, the forecasting exercise was repeated with the GDP deflator instead of the CPI. Results for the GDP deflator are equally disappointing – no output gap measure produces an inflation forecasts that is better than the univariate forecast. [23]

5. *Concluding remarks*

In this paper, we have considered the potential role for output gap measures in economic policymaking and in predicting inflation in Finland. While all output gap measures document the economic roller coaster in the 1990s, the measures have very different implications for potential output, and multiple estimates should therefore be assessed – instead of basing any policy decision on a single gap measure. Moreover, while many gap measures are attractive from a descriptive point of view, none of the measures considered above proves useful in the inflation-forecasting exercise. Instead, the simple univariate model for Finnish inflation fares best.

Consequently, further refinement of the output gap measures could improve their performance for Finnish data. Generally speaking, the production function method has much appeal due to specific attention given to the derivation of full-capacity labor input via the NAWRU estimation, and the potential association between the ICT sector and TFP. Nevertheless, future work could take into account long-run demographics and the ICT sector even more explicitly along the lines of Jalava and Pohjola (2001) to enhance the forecasting performance. For the frequency domain measure, the definition of the business cycle duration is crucial, and further research could be dedicated to determine the "optimal" frequency. A similar argument applies to the Hodrick-Prescott filter, which displays a high correlation with the production function and frequency domain approaches.

[22] *Examples include Clark and McCracken (2003), who explain the weak out-of-sample forecast performance of their Phillips curve model of U.S. inflation with instabilities of the (real-time) output gap estimates and Orphanides and van Norden (2003).*

[23] *These results have been omitted here due to space constraints, but are available from the author. Exploring a larger sample of countries, Billmeier (2004a) shows that the result for Finland is robust to a change in the maximum number of lags. The analysis of French data yields a similar outcome, whereas for other European countries (Greece, Italy, United Kingdom), some measures produce inflation forecasts that consistently outperform the univariate model.*

Thus, despite differences in levels, the suggested dynamics are fairly similar, implying that this filter may be a reasonably good way to estimate changes in the gap even if the level is uncertain. Finally, the output gap measure in our sample with the lowest RMSE (Blanchard-Quah) was obtained from a bivariate model. Larger VARs may be able to single out cyclical components in a more precise way, but the lack of observations (at annual frequency) would probably dictate switching to quarterly data.

References

Ahtiala, P. (2006). "Lessons From Finland's Depression of the 1990s: What Went Wrong in Financial Reform?" *Journal of Policy Reform* 9(1), 25–54.

Apel, M., and P. Jansson (1999a). "A Theory-Consistent System Approach for Estimating Potential Output and the NAIRU." *Economic Letters* 64, 271–75.

– (1999b). "System Estimates of Potential Output and the NAIRU." *Empirical Economics* 24, 373–388.

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

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

Berger, H., and A. Billmeier (2003). "Estimating the Output Gap in Finland." In *Finland: Selected Issues,* IMF Country Report 03/326. Washington, DC: 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* 7, 151–74.

Billmeier, A. (2004a). Ghostbusting: Which Output Gap Measure Really Matters? IMF Working Paper 04/146, Washington, DC.

– (2004b). Measuring a Roller Coaster: Evidence on the Finnish Output Gap. IMF Working Paper 04/057, Washington, DC.

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

Brunila, A., J. Hukkinen, and M. Tujula (1999). Indicators of the Cyclically Adjusted Balance: The Bank of Finland's Experience. Bank of Finland Discussion Paper 1/99, Helsinki.

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* 109, 126–50.

Clark, T.E., and M.W. McCracken (2003). The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence. Federal Reserve Bank of Kansas City Research Working Paper 03-06, Kansas City, MO.

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

Corbae, D., S. Ouliaris, and P.C.B. Phillips (2002). "Band Spectral Regression with Trending Data." *Econometrica* 70, 1067–109.

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

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.

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

Gylfason, T. (1998). "Unemployment, Efficiency, and Economic Growth: The Case of Finland." *Ekonomiska Samfundets Tidskrift* 3, 193–203.

Halme, L. (2002). "The 1990s Banking Crisis in Finland: Main Causes and Consequences, Lessons for the Future." Paper presented at the Seminar on Financial Crises – Ten Years After the Nordic Banking Crisis, Oslo: Kredittilsynet.

Harvey, A.C., and A. Jaeger (1993). "Detrending, Stylized Facts and the Business Cycle." *Journal of Applied Econometrics* 8, 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* 29, 1–16.

IMF (International Monetary Fund) (2003). *Finland: Staff Report for the 2003 Article IV Consultation.* IMF Country Staff Report No. 03/325. Washington, DC: International Monetary Fund.

Jalava, J., and M. Pohjola (2001). Economic Growth in the New Economy. World Institute for Development Economics Research Discussion Paper No. 2001/5, Helsinki.

Jonung, L., and T. Hagberg (2005). How Costly Was the Crisis of the 1990s? A Comparative Analysis of the Deepest Crises in Finland and Sweden Over the Last 130 Years. European Economy Economic Papers Number 224. Brussels: European Commission.

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

Orphanides, A., and S. van Norden (2003). The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time. CIRANO Scientific Series Working Paper 2003s-01 Montreal.

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

Rasi, C-M., and J-M. Viikari (1998). The Time-Varying NAIRU and Potential Output in Finland. Bank of Finland Discussion Paper No. 6/98, Helsinki.

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

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, DC.

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

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

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

van Ark, B. (2001). The Renewal of the Old Economy: An International Comparative Perspective. OECD Science, Technology and Industry Working Papers 2001/5, Paris.

Wagner, N. (2001). "A Note on Finland's New Economy." In *Finland: Selected Issues.* IMF Country Staff Report No. 01/215 Washington: International Monetary Fund.