



WP/04/57

IMF Working Paper

Measuring a Roller Coaster: Evidence on the Finnish Output Gap

Andreas Billmeier

IMF Working Paper

European Department

Measuring a Roller Coaster: Evidence on the Finnish Output Gap

Prepared by Andreas Billmeier¹

Authorized for distribution by Subhash Thakur

April 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.

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 paper 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 Classification Numbers: E31, E32, E37

Keywords: Finland, output gap, potential output, inflation forecasting

Author's E-Mail Address: abillmeier@imf.org

¹ The author would like to thank Robert A. Feldman and Lusine Lusinyan for very helpful comments, as well as Hamid Faruqee, Francisco Nadal de Simone, and especially Helge Berger for extensive discussions.

Contents	Page
I. Introduction	3
II. Review of the Literature.....	5
III. Measures of Potential Output and the Output Gap	6
A. Statistical Measures of Potential Output.....	6
Arithmetic Trends	6
Univariate Statistical Filters.....	8
B. Theory-Based Measures.....	17
The Blanchard-Quah Approach	17
The Production Function Approach.....	20
IV. Econometric Evaluation.....	26
V. Summary of Descriptive Results and Concluding Remarks	28
Appendix. Is There a Long-Run Phillips Curve in Finland?	31
References.....	33
Boxes	
1. Finland: Volatility in the Real Economy	3
Tables	
1. Output Gap Measures: Forecast Performance.....	27
2. Output Gap Measures: Descriptive Statistics.....	28
3. Output Gap Measures: Correlations.....	28
Figures	
1. Finland: Linear, Quadratic, and Exponential Detrending, 1960–2002.....	7
2. Finland: Hodrick-Prescott (HP) Detrending, 1980–2002.....	10
3. Finland: HP Filter Endpoint Problem, 1995–2002.....	11
4. Finland: Beveridge-Nelson (BN) Detrending, 1980–2002.....	14
5. Finland: Frequency Domain (FD) Detrending, 1960–2002.....	16
6. Finland: Blanchard-Quah (BQ) Detrending, 1980–2002.....	19
7. Finland: Production Function (PF) Approach, 1965–2002.....	23
8. Finland: Growth Accounting, 1961–2002.....	25
9. Finland: Comparing Output Gap Measures, 1980–2002.....	30

I. INTRODUCTION

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 (see Box 1)—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 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 paper, we explore the first and, to a lesser extent, the third aspects in the Finnish context.

Box 1. Finland: Volatility in the Real Economy

During the past 15 years, the Finnish economy has experienced a roller-coaster ride including (1) one of the sharpest recessions among euro-area countries in the early 1990s; (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. During the crisis, the Finnish unemployment rate soared from 3.2 percent in 1990 to 16.4 percent only three years later. Since 1993, the unemployment rate decreased continuously until it reached about 9.1 percent in 2001, and remained flat in 2002. This pronounced volatility of economic conditions is captured in the table to the right (by the high coefficient of variation of real GDP growth rates) and graphically in Figure A1 in the Appendix.

Growth: Finland vs. other Euro area countries

	Real GDP	
	Growth 1/	Volatility 2/
Finland	1.8	2.1
Euro area 3/	2.3	0.8
Portugal	2.9	0.8
Greece	2.4	1.0
France	1.9	0.9

Sources: IMF, *International Financial Statistics (IFS)* and *World Economic Outlook (WEO)* databases; data for 1990-2002.

1/ Growth refers to the average of quarterly year-on-year growth rates.

2/ Volatility is measured by the coefficient of variation of the quarterly growth rates.

3/ Unweighted average excluding Finland; Ireland is also excluded due to a lack of data over the entire period, while data for Germany start in 1992.

Defined as the difference between actual and unobservable potential output, the output gap is itself an unobserved variable. This paper 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 in the context of a number of “intuitive criteria” and evaluated econometrically in an inflation-forecasting framework.

With regard to the intuitive criteria, the mean of the gap measure should be close to zero over longer time horizons. Furthermore, the gap measure should (in terms of magnitude) produce “reasonable” maxima and minima. Finally, the measure should capture a number of “stylized

facts,” in line with traditional descriptions of economic activity in Finland: the closing of the gap in the boom period in the late 1980s and the subsequent overheating; the swing in the gap during/after the crisis period 1990–93; and, again, the narrowing of the gap during the late 1990s, driven, at first, by the economic recovery and, later, by the ICT boom. While most approaches, albeit to varying degrees, reproduce these stylized facts, the uncertainty stemming chiefly from the volatility in the real economy is reflected in positive as well as negative estimates for the output gap for 2002, depending on the estimation technique.

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 (real-time) data up to the forecast period 1990–2002, and the one-year-ahead predicted inflation is compared with actual inflation. 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 a simple autoregressive model of inflation.

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. The observed volatility is due, at least in part, to the development and swings in the performance of the ICT sector. At the same time, 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 particular, 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. Furthermore, the ICT revolution draws attention to technological progress and growth in total factor productivity as a crucial element in the production function approach.

The remainder of this paper is structured as follows. Section II reviews briefly the relevant literature. Section III contains a description of the various output gap measures for the

Finnish economy, whereas Section IV compares them in the context of a forecast model for Finnish CPI inflation. Section V reviews the main results and concludes.

II. 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. In addition to simple arithmetic detrending, the univariate statistical measures pertaining to this group and employed in Section III.A to identify potential output can be traced back to the familiar contributions by Hodrick and Prescott (1997) and Beveridge and Nelson (1981). In addition, the paper explores a method proposed by Corbae and Ouliaris (2002), who use frequency domain methods to extract information on the business cycle (and hence the underlying trend) properties of GDP.

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. Thus, any measure of the output gap should account for underemployed resources, in particular in the labor market. This can be done by using an underlying model that describes relevant aspects of the economy. The model-based approaches of Section III.B relate to Blanchard and Quah (1989), and to the large strand of literature that focuses on the production function approach to potential output. Contributions include early research done at the IMF such as Artus (1977), and, more recently, De Masi (1997). The implementation of the production function approach 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).

On Finland, 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 (1997) to the Finnish data (potential output and the natural rate of unemployment are the unobserved variables estimated simultaneously).

In Section IV, 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.²

III. MEASURES OF POTENTIAL OUTPUT AND THE OUTPUT GAP

A. Statistical Measures of Potential Output

Statistical measures of potential output aim at identifying potential output by decomposing actual output into a trend and a cyclical component.³ The most commonly used methods include simple detrending and the application of statistically motivated filters, such as the Beveridge-Nelson decomposition, or band pass filters in the time (Hodrick-Prescott) and frequency domain.

Arithmetic Trends

With linear, quadratic, or exponential detrending, potential output is assumed to follow a deterministic process, which can be approximated by a polynomial or exponential function of time. Furthermore, the trend and cycle are assumed to be uncorrelated. While these measures lack theoretical underpinning, they provide a glimpse of the information contained in the data.⁴

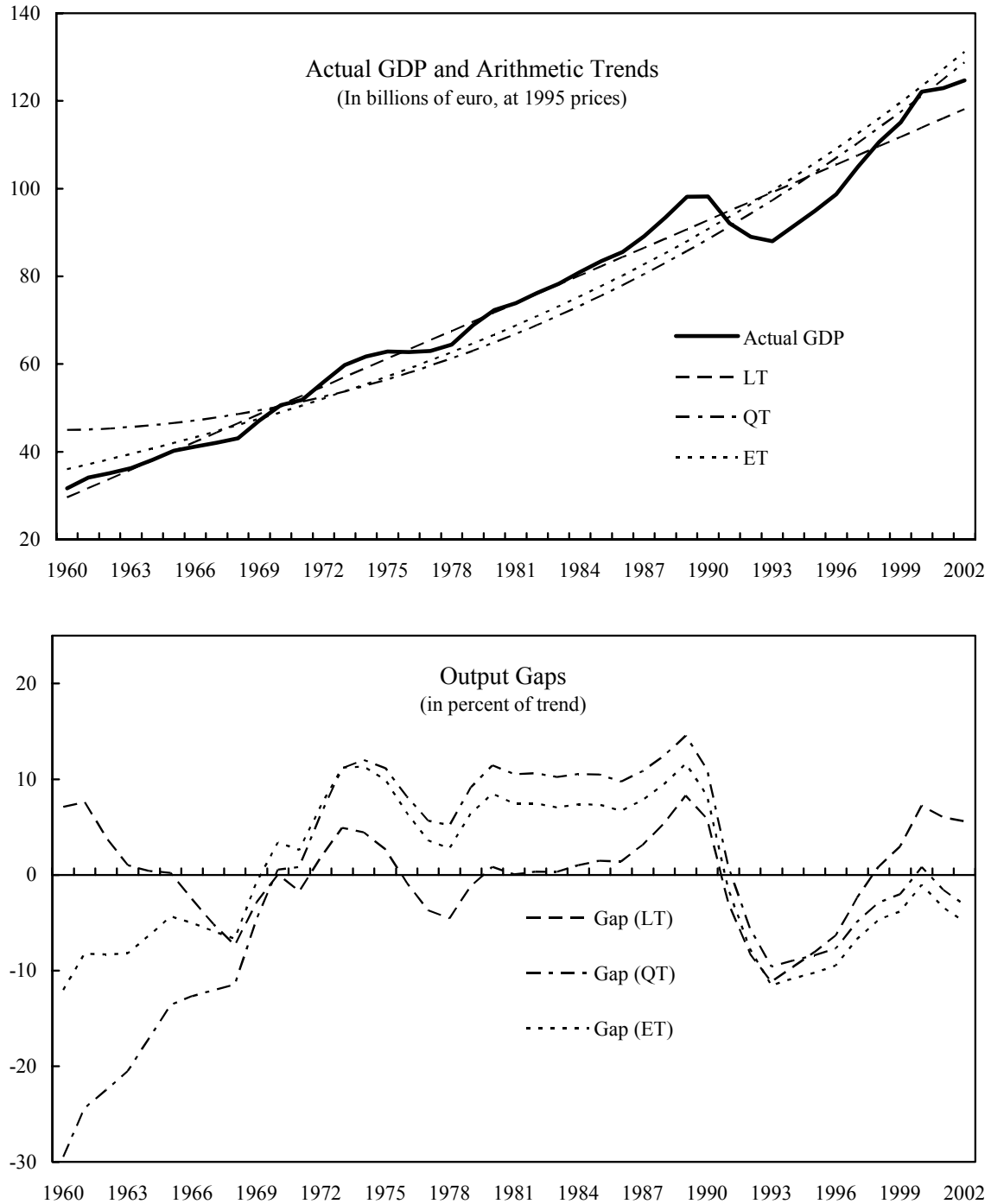
Results for potential output in Finland from linear (LT), quadratic (QT) and exponential (ET) detrending are somewhat similar in shape but indicate unrealistic swings in the output gap (Figure 1). The upper panel contains the raw and the smoothed series, whereas the lower panel shows the resulting output gap. All measures are rather similar in that they clearly identify the trough during the early 1990s, with the estimated gap in excess of 10 percent of GDP. However, the measures differ considerably with regard to the closing gap in the recent

² 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 In what follows, the estimated underlying trend is assumed to coincide with potential output, and the two expressions are used interchangeably.

⁴ In fact, Ross and Ubide (2001) identify the quadratic trend as the best methodology (out of many) to forecast both business cycle turning points and inflation in the euro area.

Figure 1. Finland: Linear, Quadratic, and Exponential Detrending, 1960-2002



Sources: IMF, *International Financial Statistics (IFS)* and *World Economic Outlook (WEO)* databases; and IMF staff calculations.

Note: LT, linear trend; QT, quadratic trend; ET, exponential trend.

past. Using the linear trend, the output gap turned positive in 1998 and remained so. The quadratic trend estimates the gap to have closed only briefly in 2000. The exponential detrending method—notwithstanding the boom in the second half of the 1990s—indicates that real GDP did not pass potential, suggesting strong growth in the latter. In terms of the intuitive criteria introduced above, the quadratic trend fails to fulfill the criterion of an average output gap of approximately zero over longer horizons. In fact, the average (positive) gap between 1980 and 2002 amounts to 3.0 percent of potential GDP. With respect to the max/min criterion, all gap measures record very large extremes of more than 10 percent of potential GDP in both directions. These results seem to be a by-product of the rigidity inherent in the trend assumptions, and, while qualitatively not implausible, raise concerns about the underlying methodology. Thus, in the next section, methods are used that attempt to glean more information from the original series by applying statistical filtering techniques.

Univariate Statistical Filters

Statistical filters can extract information either in the conventional time domain or in the frequency domain. Examples of the former include the filters by Hodrick and Prescott (1997), and the Beveridge-Nelson decomposition (1981). The frequency domain approach is represented by a filter recently developed by Corbae and Ouliaris (2002), drawing on earlier results by Corbae, Ouliaris, and Phillips (2002).

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:

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

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. Following Burns and Mitchell (1946), the standard value in the literature is $\lambda = 100$ for annual data, which is also assumed as a base case in what follows.

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 cyclicity for integrated series, the somewhat arbitrary choice of λ , as well as the neglect of structural

breaks and shifts.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 ICT 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 two issues (smoothness of the trend, end-sample bias) are considered in turn.

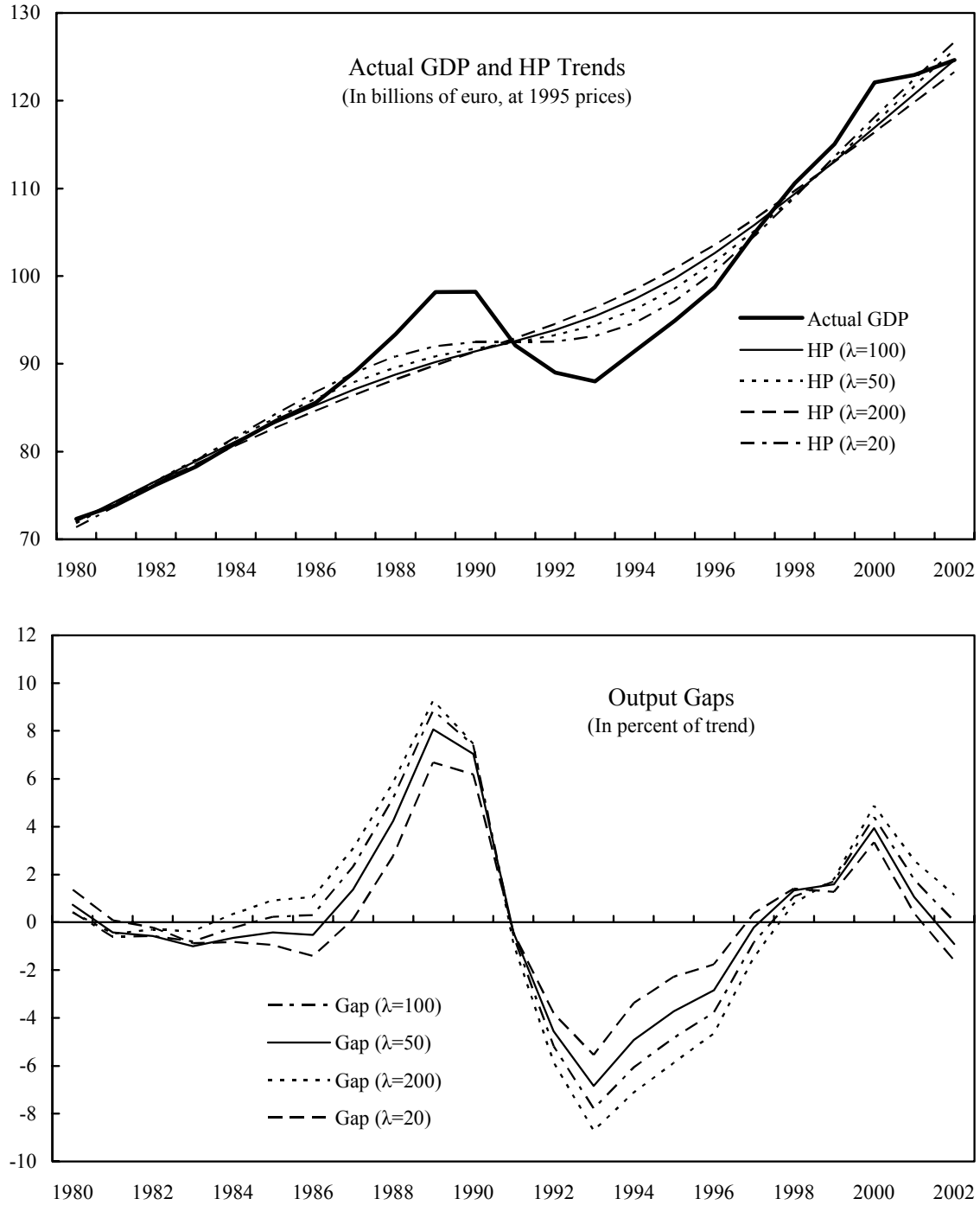
While various HP filters find a qualitatively similar pattern of the output gap, the assumption on the smoothness of the trend has strong implications for the magnitude of the gap, in particular at the end of the observation period. In Figure 2, the results of HP filtering of Finnish real GDP—for various parameter values for λ —are shown for the period 1980–2002.⁷ Trend estimates are provided for $\lambda = 20, 50, 100,$ and $200,$ with a higher parameter value indicating a smoother trend. As far as the intuitive criteria are concerned, 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. An interesting point to note is that all estimates indicate a larger (positive) output gap during the boom in the late 1980s than during the more recent ICT-related expansion. The estimates of the gap in 2002 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. To some extent, this is related to the endpoint problem.

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, hence the different summation bounds between the value function and punishment term.

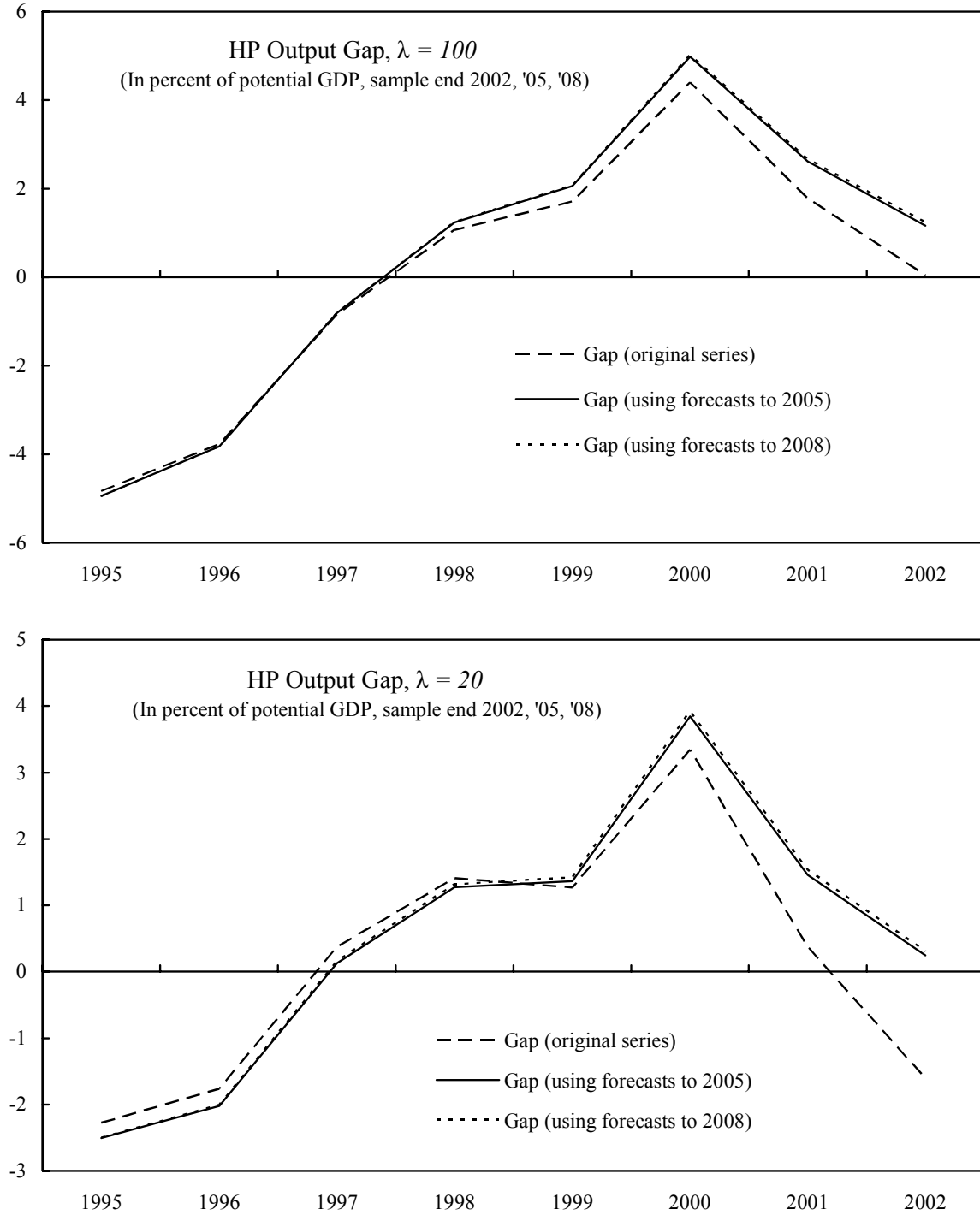
⁷ Estimation spans the period 1960–2002.

Figure 2. Finland: Hodrick-Prescott (HP) Detrending, 1980-2002 1/



Source: IMF, *IFS*; and IMF staff calculations.
1/ The last observation of the original time series is 2002.

Figure 3. Finland: HP Filter Endpoint Problem, 1995-2002



Sources: IMF, *IFS*, *WEO*; and IMF staff calculations.

The severity of the endpoint bias of the HP filter is underscored by simply adding 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 *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 (see Figure 3). For all parameter values, gap estimates based on the extended series are higher than for the original series (see Figure 3 for $\lambda = 20, 100$).⁸ In particular, for $\lambda = 100$, the output gap is clearly 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, $\lambda = 20$. 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.⁹ In fact, the negative output gap according to the original series in 2002 was replaced—due to the specific forecasted path of GDP—by a closed gap, which is less consistent with the general economic background of low inflation and slowly rising unemployment. Hence, one way to strengthen further the conclusions derived from the HP filter approach could be to undertake a sensitivity analysis, using different growth scenarios for the forecast period. Alternatively, the empirical focus could move to the nonstationary character of the underlying data, treated only implicitly by the HP filter but more technically by other models, for instance the Beveridge-Nelson decomposition.

The Beveridge-Nelson (BN) Decomposition

This approach, pioneered by Beveridge and Nelson (1981), takes a different road to decomposing a nonstationary time series such as real GDP into a nonstationary trend and a cyclical component. This filter tackles the issue by applying the Box-Jenkins (1976) method—that is, fitting an ARIMA (p, d, q) model to the real GDP series.¹⁰ The decomposition rests on the crucial assumption that innovations in the permanent (trend) and

⁸ No particular meaning is attached to the value $\lambda = 20$, other than in terms of a simple sensitivity check. For the US, Ravn and Uhlig (2002) suggest a value of 6.25.

⁹ 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.

¹⁰ In this ARIMA (autoregressive integrated moving average) formulation, p refers to the number of autoregressive lags, d refers to the order of integration, and the third parameter, q , gives the number of moving average lags; the series is assumed to be integrated of order one, that is, $d = 1$.

the transitory (cyclical) component are perfectly negatively correlated.¹¹ Based on inspection of the resulting trend component, an ARIMA(1,1,1) model is chosen to represent the structure of the underlying series.¹²

The decomposition yields a trend, which closely tracks actual real GDP (see Figure 4). When judged against the intuitive criteria, the decomposition correctly identifies the slump during the early 1990s, with the maximum output gap reached in 1991 at about -7 percent of (trend) GDP. According to the decomposition, the output gap becomes positive for one year in 1994, and then more consistently between 1996 and 2000. The growth slowdown in 2001 (from 5.6 percent in 2000 to 0.7 percent) translates into a swing of the output gap from almost 4 percent to -2.7 percent of potential GDP. The high volatility of the underlying trend is reflected in the fact that the moderate pickup in economic activity in 2002—with real GDP growth amounting to 1.6 percent—causes the output gap to again turn positive to 0.6 percent of trend GDP.

Overall, the rather pronounced volatility of the gap—together with the facts that the overheating of the Finnish economy during the late 1980s is almost completely ignored and that the gap after the crisis in the early 1990s closed immediately according to the estimates—raises doubts about the reliability of this decomposition. One way to overcome the restrictive assumption on the correlation between trend and cycle could be to focus on the latter and limit the cycle to specific frequencies, similar to the real business cycle literature. This approach is taken in the frequency domain literature.

A Frequency Domain (FD) Filter

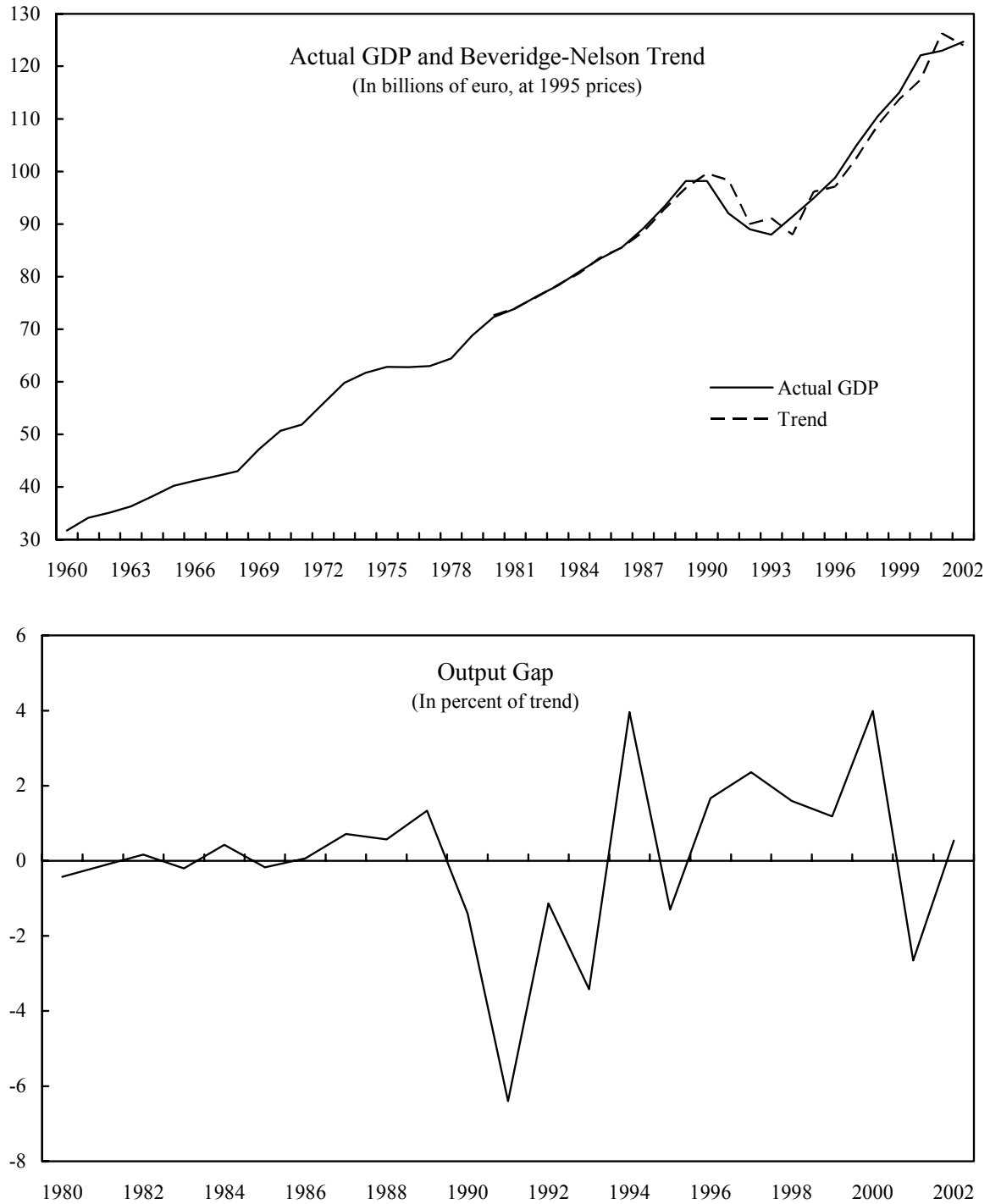
Economic fluctuations occur at different frequencies (displaying, for instance, seasonal or business cycle duration). Starting from the classical assumption contained in Burns and Mitchell (1946) that the duration of business cycles takes between 1.5 and 8 years, the approach to extracting those cycles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

A so-called exact band-pass filter 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 not be handled by traditional frequency domain approaches. Transformation of the exact band-pass filter back into the time domain results in a moving average process of

¹¹ This implies in particular for higher order models that the filtered trend can be more volatile than the original series.

¹² Various information criteria (Akaike, Schwarz, Hannan-Quinn, forecast prediction error) pointed to an ARIMA(2,1,2) specification (based on Box-Jenkins estimation). This specification, however, results in unrealistic volatility of the trend component around the peak-trough period in the late 1980s and early 1990s.

Figure 4. Finland: Beveridge-Nelson (BN) Detrending, 1980-2002 1/



Source: IMF, *IFS*; and IMF staff calculations.

1/ Due to the assumptions regarding initial conditions in the estimation procedure, the gap measure starts only 1980.

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 approximation involves a trade-off between the quality of approximation and the ability to smooth the series at the extreme points of the sample: every additional lag improves the filter but translates into one lost observation at either end of the series.

Alternatively, the estimation can take place directly in the frequency domain, where pre-filtering of the nonstationary series is required to remove stochastic trends according to Baxter and King (1999). 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 a recent paper, Corbae and Ouliaris (2002) indicate a way to avoid pre-filtering of the series in the time domain. They 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.¹³

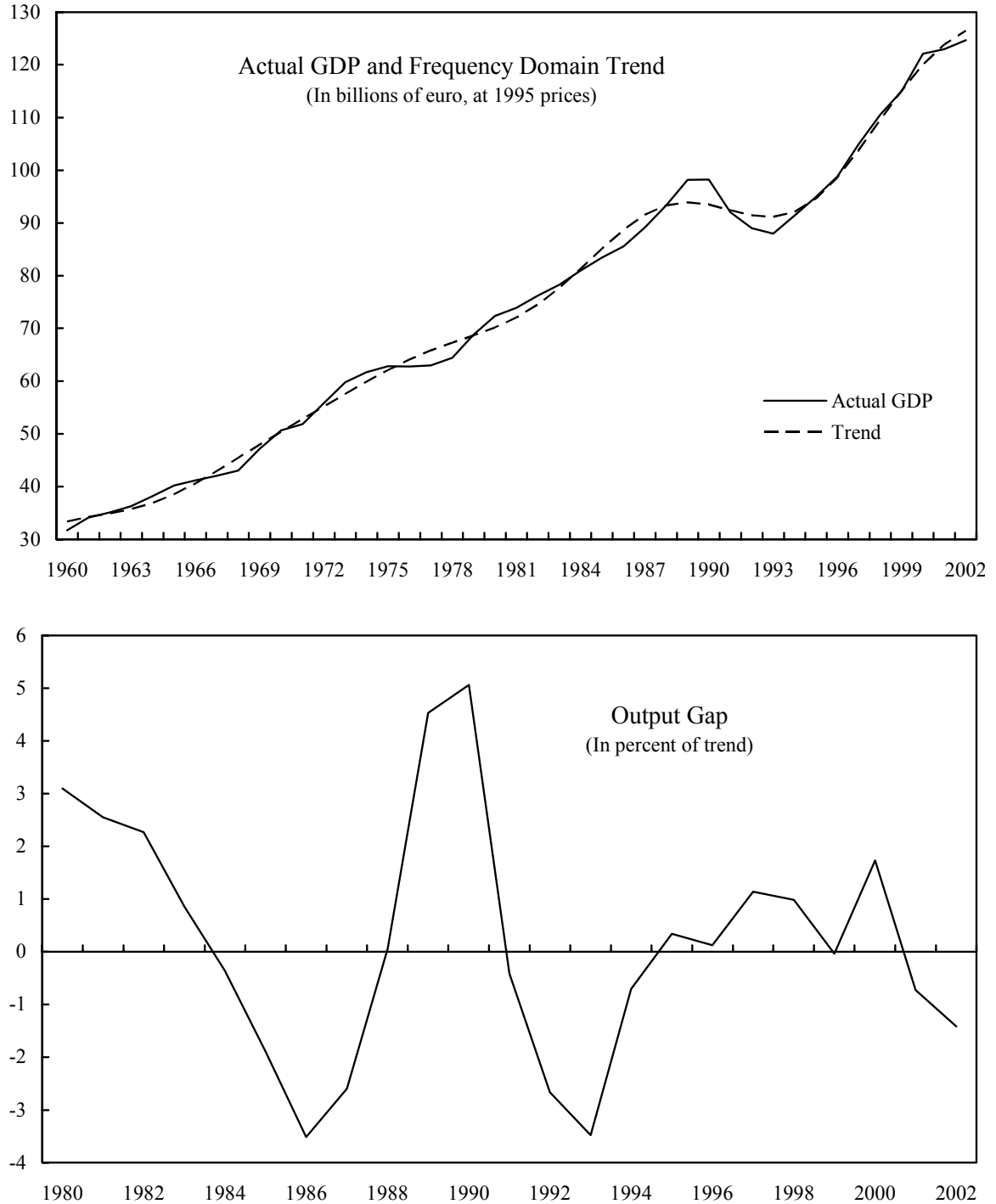
Figure 5 shows the results of the filtering process, applied to Finnish data.¹⁴ In terms of the intuitive criteria, the average output gap over the period 1980–2002 is close to zero, and the detrending procedure reproduces the Finnish boom and bust period around 1990. In particular, the overheating in the late 1980s is associated with a substantial positive output gap, which reaches 5 percent of (potential) GDP using frequency domain filtering. After the trough in 1993, the gap closed quite rapidly according to this method, remaining (marginally) positive over the period 1995–2000. The growth slowdown after the burst of the ICT bubble 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.

The results from the FD approach yield a few interesting details, especially when compared with other approaches. With regard to the extremes of the output gap, the frequency domain 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 in 1990 compared with the trough in 1993 of -3.5 percent of potential output. This is in stark contrast with the result from the Beveridge-Nelson decomposition. Based on the latter, the peak output gap (1.3 percent of trend output) is substantially smaller and reached one year earlier; the trough (-6.4 percent of trend GDP) is almost double the size of the one estimated with the frequency domain filter and occurs two years earlier. Regarding the duration of the economic downturn in the early 1990s, the frequency domain filter results in a negative gap limited to the period 1991–94, contrary to, for example, the results stemming from the HP filter (where

¹³ See Corbae, Ouliaris, and Phillips (2002) for the analysis of the asymptotic case.

¹⁴ Given that the data is annual, 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 (2004).

Figure 5. Finland: Frequency Domain (FD) Detrending, 1960-2002



Source: IMF, *IFS*; and IMF staff calculations.

the duration is much longer). Moreover, the output gap stayed relatively small during the second half of the 1990s, according to the frequency domain filter. This approach, hence, attributes much of the high actual growth rates during the late 1990s to the underlying trend, and little to cyclical factors—consistent with the view that the ICT boom had an enhancing impact on potential growth.

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 and the output gap.

B. Theory-Based Measures

Measures of potential output that rely to a larger extent on economic theory comprise, among others, the permanent-transitory decomposition by Blanchard and Quah (1989) and the production function approach.¹⁵ In the case of the Blanchard-Quah approach, the economic reasoning is tied to the conventional distinction of “demand” versus “supply” shocks, whereas the production function approach is based on a model of the aggregate production structure of the economy. The latter approach, hence, offers a variety of ways to accommodate economy-specific knowledge in the estimation, with a potentially favorable impact on the reliability of the estimates and an enhanced understanding of the economic rationale underlying the results.

The Blanchard-Quah 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, 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 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. Given the chosen structure, it seems natural to label the shocks as supply and demand shocks.¹⁶

¹⁵ See Benes and N'Diaye (2004) for a recent application of a multivariate filter.

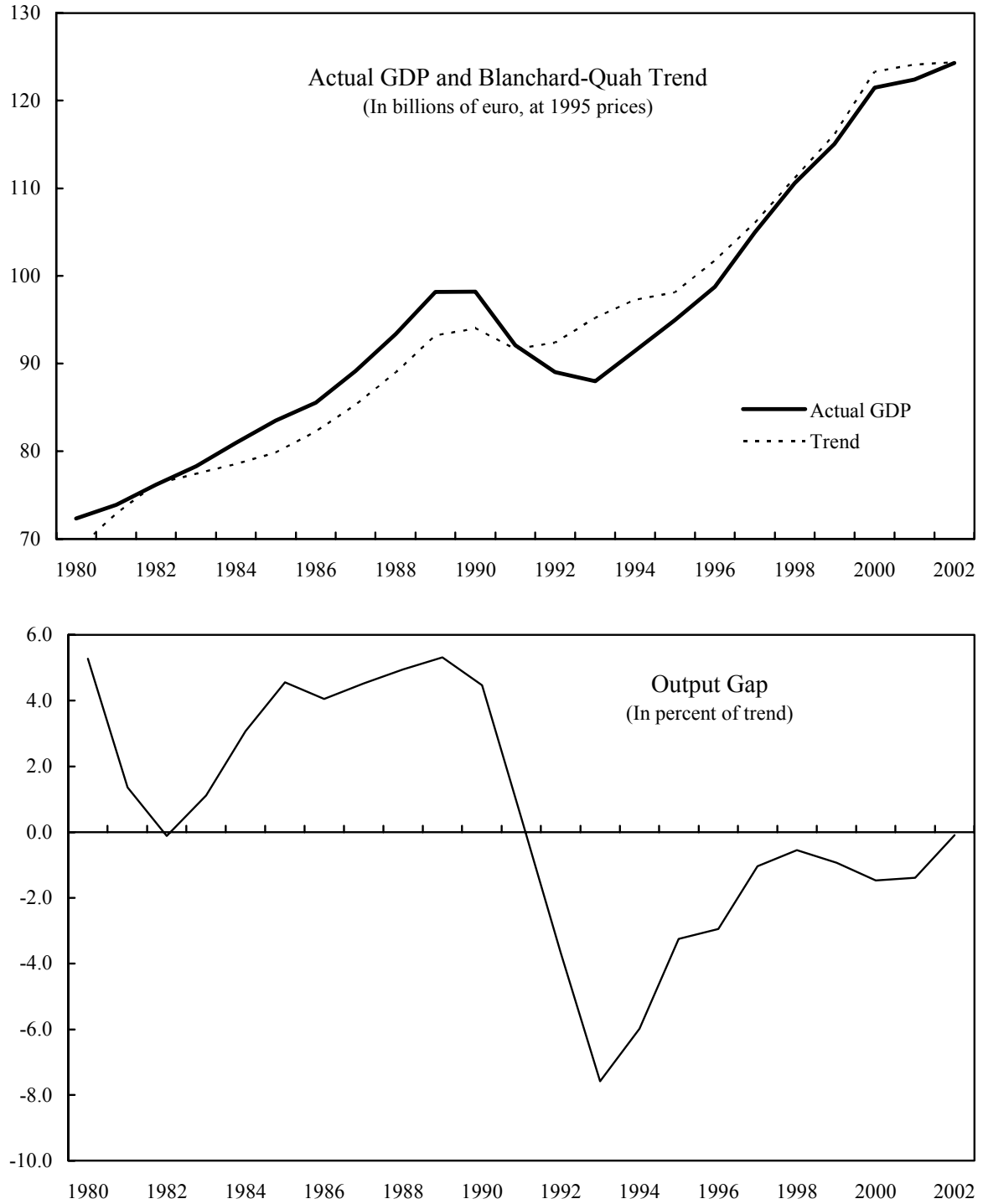
¹⁶ Blanchard and Quah (1989) also show that small violations of the identification scheme (e.g. lasting effects on output stemming from nominal shocks through a wealth effect) are of minor consequence.

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, benefits from explicit economic foundations. Furthermore, the gap—identified as the demand component of output—is not subject to any end sample bias. On the other hand, 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. However, 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 Blanchard-Quah (BQ) decomposition is—on some but not all accounts—similar to those stemming from other approaches (Figure 6).¹⁷ 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—underscoring the conventional wisdom that the Finnish economy was able to achieve strong growth almost without inflationary pressure during the second half of the 1990s precisely because the output gap had not yet been closed. In other words, the BQ decomposition attributes much of the strong real growth in the second half of the 1990s to supply side factors and high growth of potential output.

¹⁷ 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.

Figure 6. Finland: Blanchard-Quah (BQ) Detrending, 1980-2002



Source: IMF, *IFS*; and IMF staff calculations.

The Production Function Approach

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 technological progress/total factor productivity is at its trend level.¹⁸ 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 :

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

Assuming constant returns to scale implies that $\alpha + \beta = 1$; 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 from the above equation:

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

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

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

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 (in logs) as:¹⁹

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

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, a feature

¹⁸ At the IMF, early work on the production function approach includes Artus (1977). Subsequent research has reviewed the production function approach, for example, in the context of developing and transition countries; see De Masi (1997).

¹⁹ 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).

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.²⁰ Moreover, the strong movements of the unemployment rate since the crisis in the early 1990s convey valuable information on labor market conditions.²¹ 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, of course, adds an element of discretion.²² In the simplest case, potential output is a linear combination of HP-filtered series. However, the approach can also be implemented more flexibly by using more sophisticated filtering procedures, including those that incorporate structural assumptions based on economic theory.

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. From a conceptual point of view, however, this approach rests on the assumption that a natural rate of unemployment exists—in other words, that the Phillips curve is partly vertical (see the Appendix).

Evaluation of the NAWRU

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

²⁰ See Jalava and Pohjola (2001), and Wagner (2001).

²¹ Looking ahead, demographic developments will play a significant role in Finland, with the baby boomer generations expected to retire soon. In a longer-term analysis, this kind of information could be taken into account using the production function approach.

²² 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.

(autoregressive moving average) 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; no exogenous regressors are added.²³

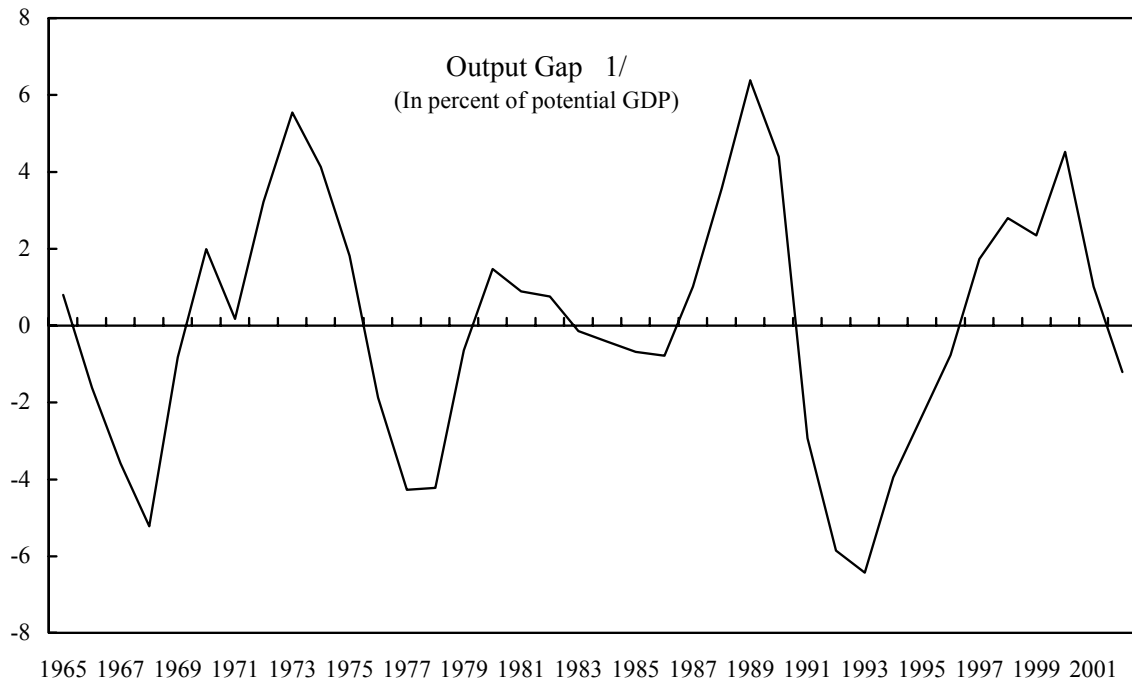
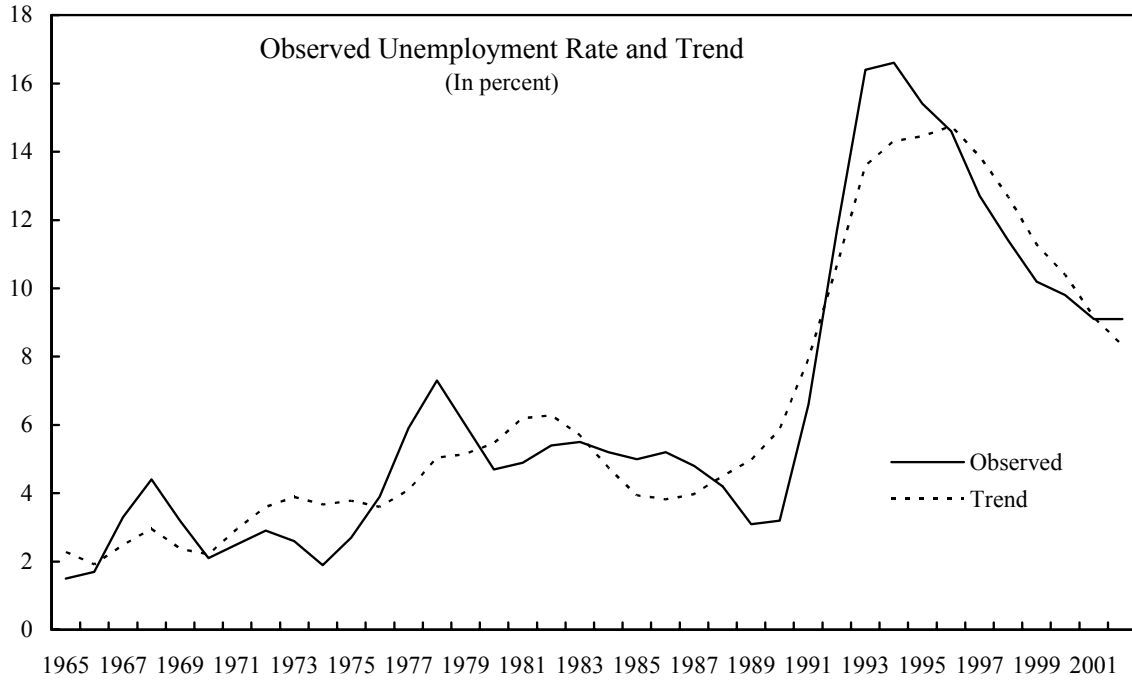
Figure 7 (upper panel) presents the estimated trend unemployment rate tracking the observed unemployment rate with a small lag. In particular, the highest rate of trend unemployment is achieved in 1996, at approximately 14.7 percent. Since then, and until 2001, the trend unemployment rate remained above 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. With respect to the modeling framework, 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 understanding of the interaction between both variables, and 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 deterioration of the system's performance from a statistical point of view.

Implications for the Output Gap and the ICT Sector

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 7, lower panel). Its mean is, in terms of the intuitive criteria, very close to zero (0.2 percent on average per year between 1980 and 2002). 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 those provided by the Beveridge-Nelson and Blanchard-Quah decompositions, which attribute a very 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. While the positive gap over the period 1997–2001 seems somewhat high, given other statistics pointing to an essentially closed gap,

²³ See Denis, Mc Morrow, and Roeger (2002), and Planas and Rossi (2003). This methodology substitutes for more “traditional” approaches—such as the HP filter—and, at the same time, unifies the European Commission's efforts toward a consistent representation of business cycles in the member countries. In the terminology of the Commission, the NAWRU model is known as the “GAP model.” A more detailed description of the model set-up for the Finnish case can be found in an earlier version of this paper; see the appendix in Berger and Billmeier (2003). There, it is shown that both the ARMA structure chosen to model wage inflation and the inclusion of additional regressors can have substantial impact on trend unemployment.

Figure 7. Finland: Production Function Approach, 1963-2002



Sources: IMF, *IFS*, European Commission; and IMF staff calculations.

1/ Based on the ARMA(2,3) model.

the negative gap in 2002 is in line with general economic indicators.²⁴ Moreover, compared with the purely statistical methods, the production function approach shifts attention away from aggregate growth to the various factors of production, thereby underscoring more forcefully the importance of the ICT sector.

The boom in the 1990s and the recent slowdown highlight the strong link between Finnish overall growth and the ICT sector in terms of growth rates—but also with regard to growth volatility (Figure 8).²⁵ With total factor productivity (TFP) being the production function residual, not related to labor and capital input, this unobservable “factor of production” is often associated with the level of “technology.” The decomposition of total growth by input factor indicates that labor contributed positively to total growth since 1995—but that the share of its contribution was small in 2002, reflecting the stalled recovery of the labor market. Capital, on the other hand, a steady and nonnegligible contributing force to overall growth before the crisis years, has not yet regained the role it played in the past.²⁶ With respect to TFP, its increasing contribution coincides with the surge of the ICT sector (and thus TFP and ICT are used somewhat interchangeably below).²⁷ Even taking into account two spikes in the late 1960s and early 1970s, the role of TFP has never been as powerful over an extended period as between approximately 1994 and 2000. In fact, TFP contributed on average 3.5 percentage points to total actual growth during the boom—almost double the combined contribution of capital and labor (see Figure 8, upper panel). At the same time, the contribution of the ICT sector was highly volatile: it varied strongly between less than 2 percentage points (1999) and almost 5 percentage points (1997). With the onset of the slowdown in growth, the ICT share declined substantially and dipped briefly into negative territory in 2001. In 2002, while growth contributions from capital and labor added approximately half of a percentage point, the growth contribution stemming from TFP almost reached 1 percentage point.

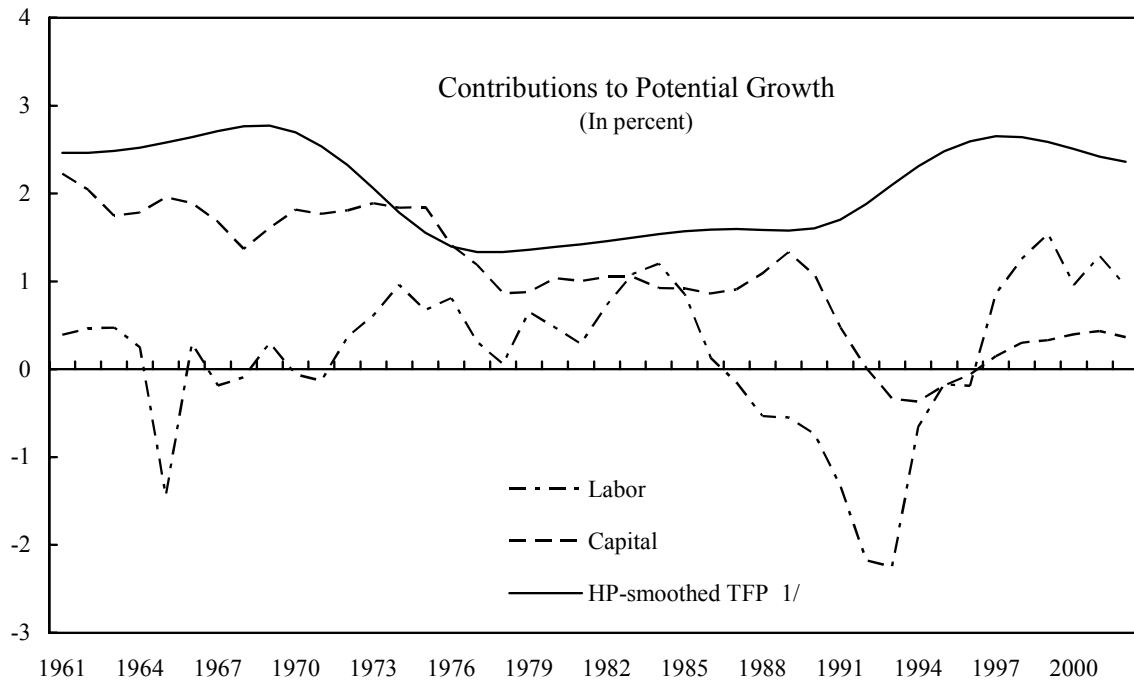
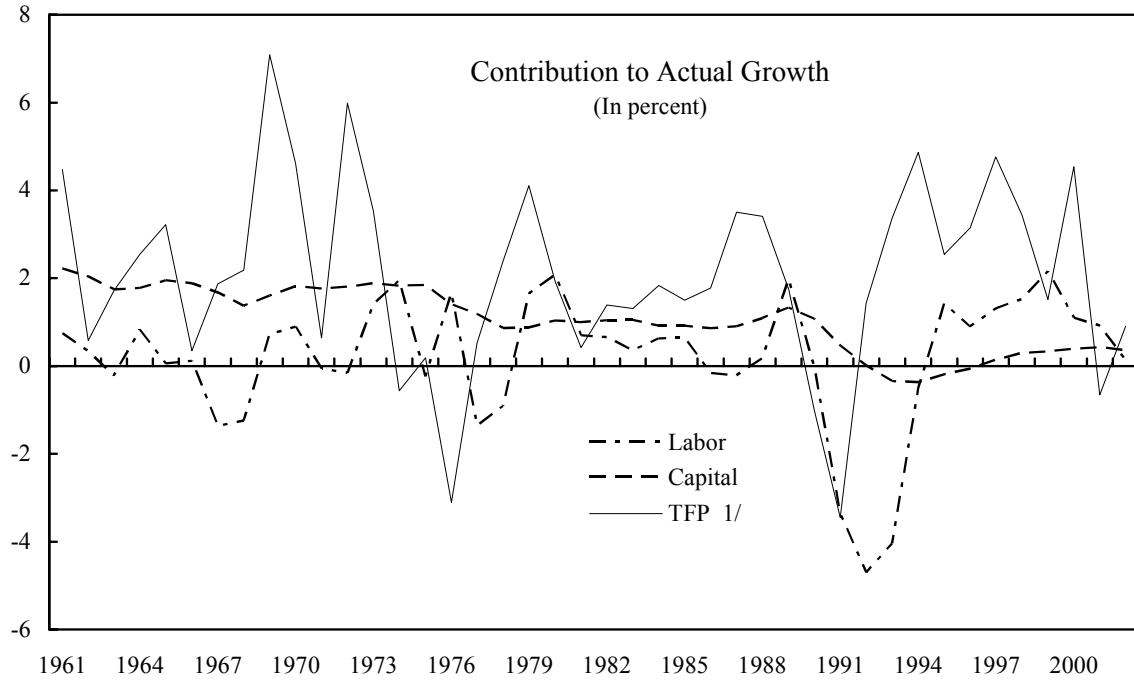
²⁴ See IMF (2003).

²⁵ This parallels earlier results: in Jalava and Pohjola (2001), it is found that the importance of “multi-factor productivity”—a concept similar to total factor productivity in the present paper—almost doubled in the second half of the 1990s; see also Wagner (2001).

²⁶ See Table 4 in Wagner (2001) for a qualitatively similar result.

²⁷ Assuming that TFP is entirely associated with “technological progress” may bias the interpretation of the results to the extent that there are, for example, errors in measuring the capital stock—though the results of Jalava and Pohjola (2001) suggest good reasons to believe in the association.

Figure 8. Finland: Growth Accounting, 1961-2002



Sources: IMF, *IFS*, European Commission; and IMF staff calculations.

1/ Total factor productivity.

IV. ECONOMETRIC EVALUATION

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. Omitting the simple arithmetic detrending procedures, inflation forecasts based on the remaining output gap estimates are compared using a simulated out-of-sample methodology.²⁸ The forecasting model is a variant of the Phillips curve:

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

where $\pi_{t+1|t}$ denotes the one-year ahead inflation in the price 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.²⁹ 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—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 “real-time” series, HP.rt, is constructed for the forecasting evaluation assuming two parameter values— $\lambda = 20, 100$. This new series consists of “last observations”—that is, real-

²⁸ The Beveridge-Nelson output gap is also omitted from the econometric evaluation due to the loss of observations needed for the initialization of the filter.

²⁹ 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 also the discussion in the Appendix. 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.

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 1 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, consists of 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.

Table 1. Output Gap Measures: Forecast Performance

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

1/ Root mean square error (multiplied by 100).

2/ Theil's (1971) U-statistic.

Note: PF, production function approach; FD, frequency domain filter; HP100(rt), (real-time) Hodrick-Prescott filter with $\lambda=100$; HP20(rt), (real-time) Hodrick Prescott with $\lambda=20$; BQ, Blanchard-Quah decomposition.

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 output gap”), however, performs even better than both the FD and the BQ filter. The relatively low RMSE of the inflation forecast 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 (see Figure A1 in the Appendix). 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.³⁰ Both versions, however, cannot provide useful information for the forecast, nor can the PF approach.

The disappointing performance of the output gap measures in improving the inflation forecast is most likely due to the high volatility of output itself, hampering the determination of a statistically satisfying measure of potential output. An additional effect may stem from the fact that the analysis is carried out with annual data. While quarterly data would have resulted in a substantially higher number of observations, and hence better model fit, other

³⁰ Clark and McCracken (2003) 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.

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.³¹

V. SUMMARY OF DESCRIPTIVE RESULTS AND CONCLUDING REMARKS

Overall, there are significant differences across the various methods—visible, for instance, in the 2002 gap. In other words, there is considerable uncertainty about the size of the output gap in Finland. Table 2 summarizes descriptive statistics for most of the output gap estimates considered above (and reproduced in Figure 9).

Table 2. Output Gap Measures: Descriptive Statistics (1980-2002)

	Gap measure									
	LT	QT	ET	HP($\lambda=100$)	HP($\lambda=200$)	HP($\lambda=20$)	FD	BN	BQ	PF
Mean	0.0	3.0	0.5	0.1	0.1	0.1	0.2	0.1	0.4	0.2
Min	-11.2	-9.6	-11.5	-7.9	-8.7	-5.7	-3.5	-6.4	-7.6	-6.5
Max	8.3	14.5	11.6	8.8	9.2	6.7	5.1	4.0	5.3	6.4
Standard deviation	5.4	8.2	7.7	4.0	4.4	2.8	2.3	2.2	3.6	3.1
Output gap in 2002	5.3	-3.4	-5.3	0.9	1.2	0.0	-1.4	0.5	-0.1	-1.5

Table 3. Output Gap Measures: Correlations 1/

	Gap measure									
	LT	QT	ET	HP($\lambda=100$)	HP($\lambda=200$)	HP($\lambda=20$)	FD	BN	BQ	PF
LT	1	0.743	0.766	0.950	0.962	0.893	0.560	0.138	0.777	0.812
QT	0.657	1	0.999	0.886	0.882	0.892	0.750	-0.182	0.895	0.670
ET	0.677	1.000	1	0.902	0.899	0.907	0.756	-0.161	0.902	0.693
HP($\lambda=100$)	0.943	0.656	0.677	1	0.999	0.985	0.756	0.105	0.883	0.888
HP($\lambda=200$)	0.959	0.705	0.724	0.996	1	0.975	0.725	0.088	0.877	0.870
HP($\lambda=20$)	0.860	0.571	0.593	0.963	0.938	1	0.858	0.157	0.891	0.922
FD	0.423	0.386	0.400	0.562	0.516	0.745	1	0.258	0.764	0.840
BN	0.185	-0.014	-0.005	0.168	0.155	0.198	0.192	1	-0.110	0.395
BQ	0.737	0.907	0.912	0.751	0.788	0.662	0.348	0.004	1	0.648
PF	0.831	0.572	0.593	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: LT, linear trend; QT, quadratic trend; ET, exponential trend; HP, Hodrick-Prescott filter; FD, frequency domain filter; BN, Beveridge-Nelson decomposition; BQ, Blanchard-Quah decomposition; PF, production function approach.

With regard to the intuitive criteria, most measures' averages are close to zero, with the notable exception of the quadratic trend, and, to a lesser extent, the exponential trend and the Blanchard-Quah approaches. Exceptionally large extremes cast doubt on the arithmetic filtering techniques (linear, quadratic, and exponential approaches), and high standard

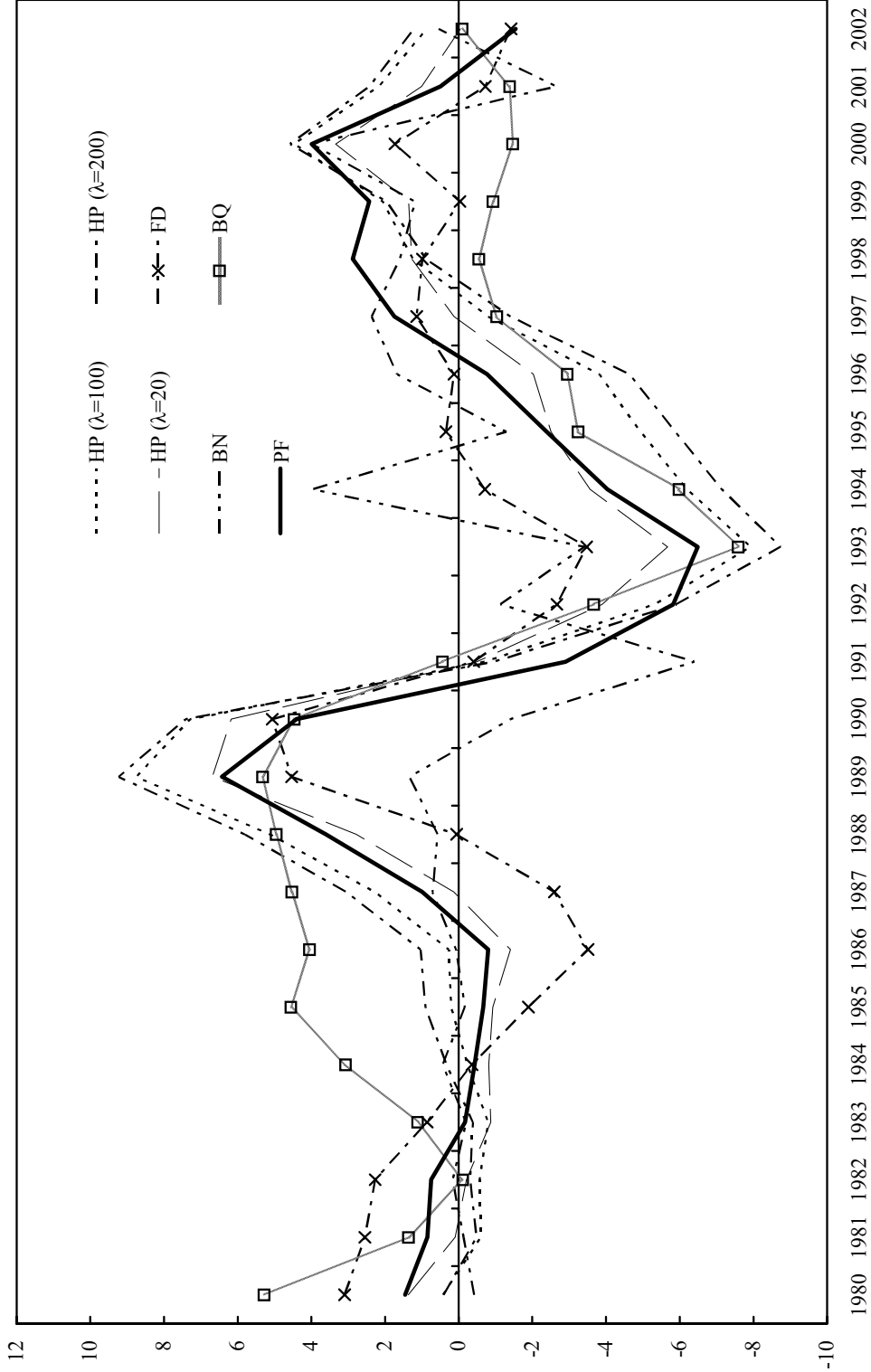
³¹ In a related paper exploring a larger sample of countries, Billmeier (2004) shows that these results for Finland are robust to a change in the maximum number of lags. The analysis of French data yields a similar outcome, whereas for other European countries (Greece, Italy, United Kingdom), selected output gap measures produce inflation forecasts that consistently outperform the univariate model.

deviations further lessen their attractiveness. The remaining measures—Hodrick-Prescott, frequency domain, Beveridge-Nelson, Blanchard-Quah, and production function—display standard deviations in the more appealing range of 2 to 4 percent. However, the associated estimates of the output gap in 2002 vary between -1.5 percent and 1.2 percent of potential. Given prospects of low inflation and high unemployment, estimates suggesting actual output above potential seem counterintuitive.³²

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 the specific attention given to the derivation of full-capacity labor input via the NAWRU estimation, and the potential association between the ICT sector and total factor productivity. Nevertheless, future work could better take into account long-run demographics and the ICT sector even more explicitly along the lines of Jalava and Pohjola (2001). This should also significantly enhance its value in forecasting inflation. The frequency domain approach has broadly similar properties in terms of the intuitive criteria (with the (negative) output gap in 2002 estimated to be 0.1 percent of potential larger than under the production function approach), and the correlation between the two approaches is rather high (0.84) over the sample 1990–2002 (see Table 3). For this measure, the definition of the business cycle duration is crucial, and further research could be dedicated to the “optimal” frequency. A similar argument applies to the Hodrick-Prescott filter. In addition, whereas both variants (two-sided and one-sided/“real-time”) are not able to improve the univariate inflation forecast, the respective correlation with the production function and frequency domain approaches is high. 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.

³² See IMF (2003).

Figure 9. Finland: Comparing Measures of the Output Gap, 1980-2002
(In percent of potential GDP)



Source: IFS, WEO; and staff calculations.

1/ HP, Hodrick-Prescott; BN, Beveridge-Nelson; 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.

APPENDIX. IS THERE A LONG-RUN PHILLIPS CURVE IN FINLAND?

A modeling framework based on a (time-varying) NAWRU—understood as the natural rate of unemployment underlying the economy—implicitly assumes that the Phillips curve is vertical at said natural rate; in other words, that the unemployment rate is independent of (wage) inflation. In other words, empirical inference along these lines rules out the existence of a long-run nonvertical Phillips curve, and, hence, an underlying relationship between inflation and the unemployment rate. This prior has been questioned recently by a number of authors (see, for instance, Beyer and Farmer (2002), and Schreiber and Wolters (2003)), who found empirical evidence against the vertical Phillips curve assumption in U.S. and German data, respectively. As the latter argue, the existence of the NAWRU can be rejected if both the unemployment rate and the rate of (wage) inflation are nonstationary and cointegrated, indicating a long-run relationship, similar to a Phillips curve.

In the case of Finland, simple tests indicate that, while both wage inflation and unemployment are nonstationary, there is no sign of cointegration, a result conducive to the NAWRU approach (see Figure A1).³³

	<i>d_gf</i>	<i>d_wage</i>	<i>ur</i>
	LR test for exclusion		
1		5.02	7.03
	LR test for stationarity		
1	7.03		5.02
	LR test for weak exogeneity		
1	4.97		0.72

Note: Bold test statistics indicate significance at the 5 percent level, the critical value with 1 degree of freedom being 3.84.

In fitting a bivariate VAR to the basic data, the lag length was chosen according to the Schwarz and the Hannan-Quinn information criteria, which both propose three lags (in levels); see text table.³⁴ Due to the lack of strong priors in favor of a trend restricted to the cointegrating space, a system with an unrestricted constant was estimated. Likelihood ratio tests (distributed as $\chi^2(d_{g}f)$) of the time-series properties reveal that both series appear to be nonstationary;³⁵ in addition, the unemployment rate can be considered weakly exogenous from a statistical point of view.

Lags	Schwarz	Hannan-Quinn
4	-6.0	-6.5
3	-6.4	-6.8
2	-6.2	-6.5
1	-5.9	-6.0

³³ Note that these conclusions also hold for CPI inflation instead of wage inflation (results not reported here).

³⁴ With three lags, no significant residual autocorrelation emerged, whereas more parsimonious models reveal problems of autocorrelation at the first lag.

³⁵ On theoretical grounds, the unemployment rate is bounded by the interval (0;1) and hence not truly I(1). The fact that it cannot grow out of bounds in the long run, however, does not preclude it from behaving like an integrated process in the shorter run, as evidenced by the test statistics. The stationarity tests presented above do not allow for a structural break in the series analyzed. The strong rise of the unemployment rate in the early 1990s—as described above—could be viewed as such a break. This proposition is not investigated further since a stationarity result for the unemployment rate when allowing for a break in the series even underscores the case for the NAWRU approach; see Schreiber and Wolters (2003).

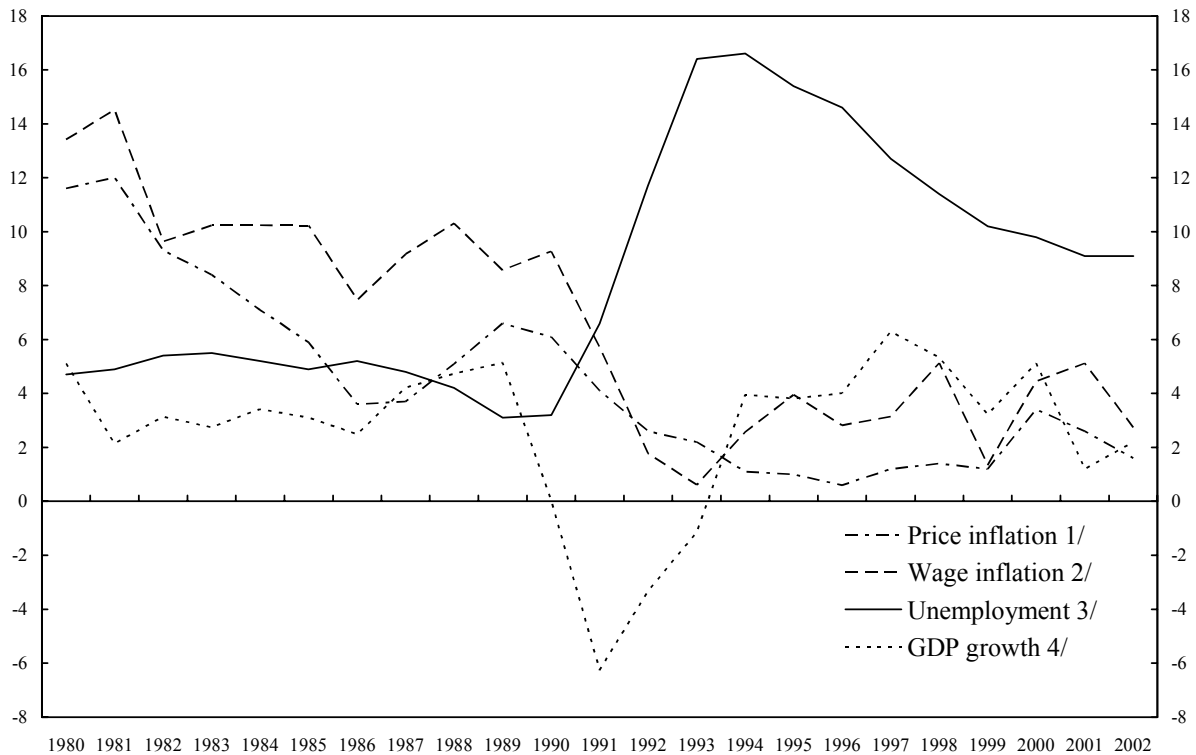
Based on this, the analysis indicates no cointegration between the unemployment rate and wage inflation: the null hypothesis of $r = 0$ —that is, no cointegrating relationship—cannot be rejected at conventional levels. Hence, the data cannot provide evidence of a long-run relationship between wage inflation and the unemployment rate.

Cointegration: Test Statistics				
Null:	L-max	Trace	L-max90	Trace90
$r = 0$	8.67	10.23	10.6	13.31
$r = 1$	1.56	1.56	2.71	2.71

Note: L-max and Trace describe the maximum eigenvalue and trace test statistics, and the appropriate 90 percent critical values for r cointegrating vectors, see Johansen (1995), p.215.

To confirm further the applicability of the NAWRU approach to the Finnish data, a number of additional considerations are of interest. The lack of cointegration between the two series could be due to a structural break in the cointegrating relationship during the observation period, in particular given the sharp rise in unemployment during the early 1990s. In results not reported here, experimenting with various dummies did not soften the evidence against cointegration. Moreover, the limited number of observations used in the empirical assessment may introduce a small sample bias. Correcting for the bias—for instance, along the lines of Cheung and Lai (1993)—introduces even higher critical values, however, such that the hypothesis $r = 0$ would be accepted even more easily.

Figure A1. Finland: Macroeconomic Variables, 1980-2002
(Percent change over previous year unless otherwise noted)



Sources: Statistics Finland, Finnish Ministry of Finance; and IMF (*IFS*).

- 1/ Consumer price index (CPI)
- 2/ Economy-wide average wage
- 3/ Unemployment rate, in percent of total labor force
- 4/ Real GDP

REFERENCES

- Apel, M. and P. Jansson, 1997, "System Estimates of Potential Output and the NAIRU," Sveriges Riksbank Working Paper No. 41 (Stockholm: Sveriges Riksbank).
- 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.
- Benes, J. and P. N'Diaye, 2004, "A Multivariate Filter for Measuring Potential Output and the NAIRU: Application to the Czech Republic," IMF Working Paper 04/45 (Washington, DC: International Monetary Fund).
- 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.
- Beyer, A. and R.E.A. Farmer, 2002, "Natural Rate Doubts," ECB Working Paper 121 (Frankfurt: European Central Bank).
- Billmeier, A., 2004, "Ghostbusting: Which Output Gap Measure Really Matters?" *forthcoming* as IMF Working Paper (Washington, DC: International Monetary Fund).
- Blanchard, O. and D. Quah, 1989, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–73.
- Box, G.E.P. and G.M. Jenkins, 1976, *Time Series Analysis, Forecasting and Control* (San Francisco: Holden Day).
- 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: Bank of Finland).
- 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.
- Cheung, Y.-W., and K.S. Lai, 1993, "Finite Sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration," *Oxford Bulletin of Economics and Statistics*, 55, 313–28.

- 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: Federal Reserve Bank of Kansas City).
- Corbae, D. and S. Ouliaris, 2002, "Extracting Cycles from Non-Stationary Data," (unpublished; Austin, Texas and Singapore: University of Texas and National University of Singapore).
- , 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: 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).
- Everaert, L. and F. Nadal De Simone, 2003, "Capital Operating Time and Total Factor Productivity," IMF Working Paper 03/128 (Washington, DC: International Monetary Fund).
- Gylfason, T., 1998, "Unemployment, Efficiency, and Economic Growth: The Case of Finland," *Ekonomiska Samfundets Tidskrift*, 3, 193–203.
- 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: United Nations University).
- Johansen, S., 1995, *Likelihood based Inference on Cointegration in the Vector Autoregressive Model* (Oxford: Oxford University Press).
- 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: Centre Interuniversitaire de Recherche en Analyse des Organizations).
- 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: Bank of Finland).
- Ravn, M.O., and H. Uhlig, 2002, "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics*, 84(2), 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, DC: 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).
- Schreiber, S. and J. Wolters, 2003, "What's Wrong with the (German) NAIRU?" (unpublished; Berlin and Frankfurt: Free University Berlin and University of Frankfurt).
- 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).
- 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).