INTERNATIONAL MONETARY FUND

Staff Country Reports

October 2003
IMF Country Report No. 03/326

Finland: Selected Issues

This Selected Issues paper for **Finland** was prepared by a staff team of the International Monetary Fund as background documentation for the periodic consultation with the member country. It is based on the information available at the time it was completed on **September 24, 2003.** The views expressed in this document are those of the staff team and do not necessarily reflect the views of the government of **Finland** or the Executive Board of the IMF.

The policy of publication of staff reports and other documents by the IMF allows for the deletion of market-sensitive information.

To assist the IMF in evaluating the publication policy, reader comments are invited and may be sent by e-mail to publicationpolicy@imf.org.

Copies of this report are available to the public from

International Monetary Fund • Publication Services
700 19th Street, N.W. • Washington, D.C. 20431
Telephone: (202) 623 7430 • Telefax: (202) 623 7201
E-mail: publications@imf.org • Internet: http://www.imf.org

Price: $15.00 a copy

International Monetary Fund
Washington, D.C.

INTERNATIONAL MONETARY FUND

FINLAND

Selected Issues

Prepared by Helge Berger, Louis Kuijs, and Andreas Billmeier (all EU1)

Approved by European I Department

September 24, 2003

Contents Page

Figures

Tables

Appendices

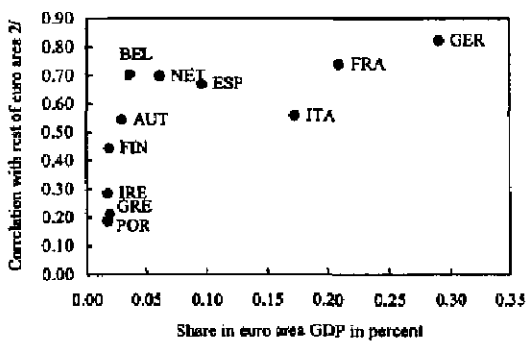
I. THE IMPACT OF FISCAL POLICY IN FINLAND[1]

A. Introduction

1. In light of potential changes in fiscal policy in Finland, the impact of changes in the fiscal stance on short-term economic growth is an issue of interest. The new government's program includes significant expenditure increases for 2003 and 2004, at the same time that tax cuts are being considered. These actions would lower the general government surplus and possibly result in considerable fiscal stimulus. Drawing on analytical work conducted at the Ministry of Finance and the Bank of Finland, IMF staff, in the context of the 2003 Article IV consultation, consider a general government surplus of 4 percent of GDP as a medium-term norm which could be achieved by targeting an increase in the structural primary balance of ½ percent of GDP per year (while allowing the automatic stabilizers to play around the consolidation path). All this underscores why discussions of fiscal policy and its effect on the economy have come to the fore.

2. This paper estimates the short-term impact of fiscal policy on growth in Finland. The estimates are based on a structural vector auto regression (SVAR) model, following the methodology introduced by Blanchard and Perotti (2002) (hereafter called BP) in a study on the United States, and also applied by Céspedes and Hoffmaister (2003) (hereafter called CH) to Spain. Against the background of Finland's participation in the euro area and thus the absence of an independent monetary policy, and because Finland's economy is among the least synchronized with the core euro-area countries (Figure 1), this question takes on an added dimension.

Figure 1. Finland and Selected Countries: Size and Correlation of GDP Growth with the Rest of Euro Area, 1992-2003 1/



Sources: WEO; IFS; and Fund staff calculations.

1/ Based on real quarterly GDP data, seasonally adjusted. Ireland: 1995-2003.
2/ Correlation with euro area *excluding* the country in question.

3. While the application of the structural VAR approach to Finland is somewhat hampered by data limitations, the results suggest that fiscal policy has only a modest impact on activity. Thus, although, for example, expansionary fiscal surprises do indeed increase real GDP in the short run, the impact remains small compared to the results reported in BP and CH for Spain.

[1] Prepared by Helge Berger and Louis Kuijs.

B. Empirical Evidence on the Impact of Fiscal Policy on Growth in OECD Countries

Macroeconomic models

4. Traditionally, estimates of the short-run impact of fiscal policy on growth predominantly came from simulations using structural macro models. Such studies, in which Keynesian effects play a prominent role, suggest fairly large short-term expenditure multipliers for European countries, ranging from 0.6 to 1.5 in "dollar-for-dollar" terms, and somewhat smaller revenue multipliers (see Brunila, Buti, and in 't Veld (2002), Hunt and Laxton (2003) and, for a survey, Hemming, Kell, and Mahfouz (hereafter called HKM) (2002)).[2] The strength of studies based on structural macroeconomic models is also one of their major weaknesses: while structural simulations shed light on the channels via which effects may take place (or be offset), they often model these relationships based on economic theory that may not be firmly rooted in empirical estimates.

5. The study by Brunila, Buti, and in 't Veld (2002) is of interest to the Finnish case. Using the European Commission's Quest model, these authors compare the impact of changes in different kind of expenditure and revenue items. They measure short-term expenditure multipliers from a *temporary* shock in which government expenditures are increased by 1 percent of GDP (government purchases of goods and services, government investment, transfers to households, and government employment). Short-term revenue multipliers are produced by reducing labor taxes, corporate profit taxes, and value-added taxes by one percent of GDP.

6. The impact of higher government expenditures on GDP is modest because of the crowding out of private spending via higher real interest rates and leakage via imports. The extent of crowding out depends on the response of monetary policy, and leakage via imports depends, inter alia, on the openness of the economy. In the Quest model, the majority of households is assumed to be permanent income consumers, whose consumption responds only to a small extent to temporary changes in transfers or taxes. This is why changes in transfers to households and taxes have smaller effects than changes in government consumption and investment. In the case of changes in government purchases of goods and services or investment, the short-term multiplier in Finland is estimated at around 0.65. The multiplier for changes in the government wage bill is close to 1, but the multiplier for changes in transfer payments is only 0.2.

7. As was the case for temporary changes in transfers, the simulations suggest that the impact of temporary changes in labor and corporate income taxes on output is small—with multipliers of around 0.2—because the intertemporal optimizing behavior of economic

---

[2] A multiplier in "dollar-for dollar terms" measures the ratio of the (absolute) amount of additional activity in terms of currency units generated by one unit of additional expenditure.

agents smoothes away most of the impact.[3] In the case of a permanent tax cut, the impact of a tax cut would strengthen over the medium term as the distortionary effects of taxation are reduced. For instance, a reduction in labor taxes has a direct demand effect through its impact on disposable income and a positive supply effect. The latter also allows for lower wage costs and improved competitiveness, further boosting (foreign) demand. The impact of temporary changes in indirect taxes is significantly higher than in the case of direct taxes—estimated at one-half for Finland—as private agents are assumed to bring expenditures forward in anticipation of a return to higher tax rates in subsequent years.

8. In all, in the case of temporary changes to fiscal policy, the impact of expenditure changes is found to be larger than that of revenue changes. In contrast, in the case of more permanent fiscal policy changes, the impact of expenditure changes would fade out, while the supply side effects of tax changes become more important.

Case studies of "expansionary" fiscal contractions

9. In circumstances of high government debt, the credibility effects of a fiscal contraction can offset (in part or fully) the traditionally assumed Keynesian effects. Indeed, the experience of certain European countries that undertook fiscal consolidation in the 1980s and 1990s in circumstances of high government debts and deficits generated a literature on "expansionary fiscal contractions." Discussing some of the available evidence, HKM conclude that there indeed appear to have been episodes of expansionary fiscal contraction, and that some episodes share certain characteristics.

10. HKM stress, however, that caution is needed in drawing general conclusions from these experiences and point to methodological flaws such as selection bias problems (country experiences are "handpicked"), simultaneity bias (strong growth led to lower fiscal deficits), and omitted variables (not taking into account sharp devaluations that accompanied fiscal contractions and influenced growth). Indeed, although credibility effects seem plausible, it is unlikely that they have been large enough to offset the Keynesian effects of fiscal consolidation. Credibility effects can largely be captured by the reduction in long-term interest rates, the effect of which on activity can be measured separately and which is typically found to be relatively modest, compared to the Keynesian effects of fiscal policy.

Studies based on structural VAR models

11. Empirical estimates of fiscal multipliers using structural VAR models—which have a strong empirical element—typically find significant, positive multipliers. Data constraints have tended to limit their application to large industrial countries. While studies focusing on the U.S. typically find significant multipliers reinforcing Keynesian priors, recent papers on large European countries have mixed conclusions. Aarle, Garretsen, and Gobbin (2001)

---

[3] In the absence of liquidity constraints, temporary tax cuts that are later reversed would not have any effect on spending.

found considerable variation in the size and signs of multipliers for EU countries. In a study that includes Germany (as well as four Anglo-Saxon countries), Perotti (2002) finds that the effects of fiscal policy on GDP have become weaker over time; they are substantially smaller in the post-1980 sample than the pre-1980 sample. He finds that in the post-1980 sample only in Germany is the effect of government spending on GDP significantly positive on impact. But even there, the effect turns negative by the fourth quarter. His findings on the impact of taxation on GDP are mixed, but Keynesian effects seem to be present for Germany, with multipliers in the range of 0.2 to 1 in the first three years. In the above-mentioned study for Spain, CH find significant multipliers for both revenues and expenditures. Notwithstanding the diversity of results, in most cases the short-term fiscal multipliers estimated by structural VAR models tend to be lower than those found by simulations of macro models, but significantly higher than suggested by the literature on "expansionary fiscal contractions."

Other methods

12. In a preliminary empirical study of the impact of fiscal policy and other factors on growth in euro area countries, IMF staff (forthcoming), using a panel of annual time series data (1980–2001), find that GDP growth has been determined mainly by changes in partner countries' import growth and the fiscal stance. The fiscal multipliers, estimated at around 0.4 to 0.5, are not as large as typically assumed in structural macro models, but larger than found in some other recent studies, including those on "expansionary fiscal contractions." Moreover, the study suggests that fiscal multipliers have not become smaller in the 1990s. However, tentative estimations on a country-by-country basis suggest that, overall, fiscal multipliers are weaker in smaller and more open countries, as would be expected on theoretical grounds. While for Finland no statistically significant results were found in the individual country estimation, the cross-country results suggested a multiplier of around 0.4 to 0.5 for an economy with a share of total trade to GDP equal to Finland's.

C. Evidence from a Simple VAR Exercise for Finland

13. The empirical evidence on the impact of fiscal policy on output in Finland, using the structural VAR model proposed by BP and applied by CH, is based on quarterly data from 1991 onwards. The data limitations are across two dimensions. First, quarterly information on general government expenditures and revenue excludes several items of the overall fiscal flows.[4] Thus, the available data were combined with quarterly national accounts data on government consumption and investment, and mapped into total revenues and expenditures data using the composition of annual total data. Second, the resulting time series are short, covering only the period 1991 to 2001. Standard unit root tests showed that transforming the fiscal data and GDP (all in real terms) into logs renders the data trend-stationary, with no clear indication of a cointegration relationship.

---

[4] The available quarterly fiscal data accounts for 34 percent of expenditures, and data on revenues that accounts for 84 percent of the total.

The model

14. Consider a trivariate system in which r_t is real government revenues, g_t is real government expenditures, and y_t is real GDP at time t (all in logs). With the vector X_t defined as $(r_t, g_t, y_t)'$, the VAR can be written as

$$X_t = A(L)\, e_t, \tag{1}$$

where A(L) is a lag polynominal (3 x 3) matrix containing the *reduced-form* dynamic effects of the system and e_t is a vector of serially independent reduced-form shocks to revenues, expenditures, and GDP, with $E[e_t] = 0$ and $E[e_t\ e_t'] = \Omega$, where Ω is the variance/covariance matrix of the reduced-form VAR surprises.[5] In a structural VAR, e_t is interpreted as a linear combination of independently distributed *structural* shocks to revenues, expenditures, and GDP, that is

$$e_t = B\, u_t, \tag{2}$$

where B is a (3 x 3) matrix and u_t is the vector of structural shocks in revenues, expenditures, and GDP.

Blanchard-Perotti identification steps

15. The identification process consists of obtaining the matrix B (see BP and CH for details). Once B is obtained, the information contained in the estimated reduced-form unexpected movements e_t can be used to uncover the structural shocks u_t.

Imposing the assumptions of the BP approach, equation (2) can be reduced to

$$\begin{aligned} e^r &= a_1\ e^y + a_2\ u^g + u^r \\ e^g &= b_1\ e^y + b_2\ u^r + u^g \\ e^y &= c_1\ e^r + c_2\ e^g + u^y \end{aligned} \tag{3}$$

where u^r, u^g, and u^y are the mutually uncorrelated structural shocks to revenue, expenditure, and GDP that are to be recovered.

16. The first line of the equations in (3) states that unexpected movements in taxes can be due to the response to unexpected movements in GDP, captured by e^y, and the response to structural (or, discretionary) shocks to expenditure and revenues (respectively, u^g and u^r). A similar interpretation applies to unexpected movements in spending in the second line. The

[5] Sometimes e_t is referred to as the vector of unexpected movements or surprises based on the reduced-form VAR model.

third line states that unexpected movements in output can be due to unexpected movements in taxes, unexpected movements in spending, or to other, structural, shocks in GDP (u^y). Thus, (3) leaves us with six unknown coefficients and three equations. To be able to estimate the impact of structural, discretionary fiscal shocks on GDP, at least three of the unknown coefficients have to be pre-identified.

17. The response of taxes and expenditures to changes in GDP stem from the so-called *automatic stabilizers* and a possible *discretionary adjustment of fiscal policy to cyclical conditions* as measured by the reduced-form unexpected movements in GDP. Coefficients a_1 and b_1 capture both effects for revenues and expenditures, respectively. As BP argue convincingly, the use of quarterly data "virtually eliminates the second channel." Thus, they suggest setting these coefficients equal to estimates of the revenue and expenditure elasticities. Based on recent OECD (2003) estimates for Finland, elasticities of 1 for revenues and around -0.4 for expenditure would appear to be reasonable starting points.[6]

18. The coefficients a_2 and b_2 in (3) describe possible contemporaneous *impacts of structural shocks* in expenditure and revenues on unexpected movements (i.e., the VAR residuals) in revenues and expenditure. How large will these effects be? The answer depends on the institutions of policy making. If tax decisions are made before expenditure decisions, a_2 is zero. As suggested by BP, b_2 could then be estimated freely. Alternatively, b_2 could be assumed to be zero, and a_2 estimated freely. As a rule, with the budget process based on the entire fiscal year (rather than quarters), both effects will be rather small, however. Thus, to economize on degrees of freedom, a_2 and b_2 are both set at 0.[7]

19. Under these assumptions, estimating the impact of discretionary changes in (or structural shocks to) fiscal policy on GDP becomes a straightforward exercise. With a_1 and b_1 known, the" cyclically adjusted" unexpected movements in revenues and expenditures can be computed as:

$$e^{r*} = e^r - a_1 e^y$$

and

$$e^{g*} = e^g - b_1 e^y,$$

which are then used to obtain estimates of c_1 and c_2 in (3):

$$e^y = c_1 e^{r*} + c_2 e^{g*} + u^y. \tag{4}$$

---

[6] See CH for comparable results for Finland. The empirical section below comments on the robustness of these assumptions.

[7] The assumption is in line with preliminary estimates of (3) which found no robust evidence that either a_2 or b_2 is statistically and economically significant.

Given our assumptions, c_1 and c_2 capture the impact of structural or discretionary fiscal policy changes on GDP.

Results

20. To empirically identify the contemporaneous and dynamic implications of discretionary fiscal policy changes on GDP, a standard VAR is estimated with revenues, expenditure, and GDP as endogenous variables (all in real terms and logs) using four lags. In addition, the model includes seasonal dummies and a linear and non-linear trend as exogenous variables.[8] The upper two panels in Figure 2 summarize the results of that exercise.

21. Using the residuals from the VAR and the OECD's (2003) revenue and expenditure elasticities (i.e., $a_1 = 1$ and $b_1 = -0.4$), c_1 and c_2 can be estimated, and the structural shocks uncovered (see lower panel in Figure 2). Table 1 summarizes the parameter values. On a dollar-for-dollar basis, the point estimate of the *contemporaneous impact* of a discretionary change in revenues on GDP (c_1) is about -0.1; the impact of a discretionary change in government spending (c2) is about 0.2.[9] The direction of these effects is in line with our priors: an increase in taxes lowers GDP while higher expenditure has a positive impact on real activity, with the latter dominating the former in quantitative terms. However, the quantitative effect is notably smaller than what CH report for Spanish and BP for the U.S. data. CH report a contemporaneous effect of revenues of about -2.8 and of expenditures of about 1.4, and BP find effects of about -0.9 and 1.0, respectively.[10] Moreover, in the Finnish case neither coefficient is statistically significant at conventional levels.

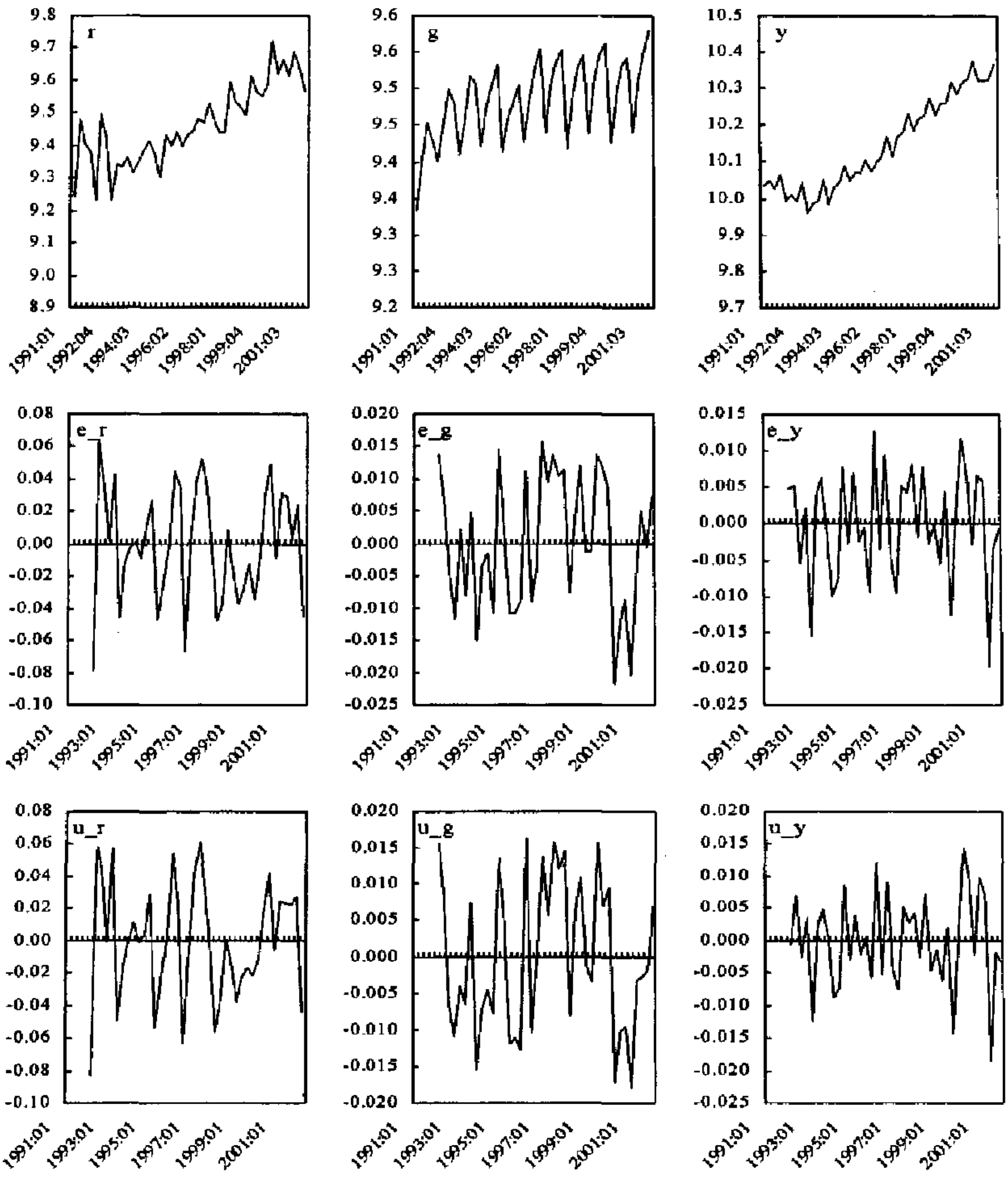
22. The results for c_1 and c_2 are fairly robust with regard to different assumptions for the (exogenous) elasticities of revenues and expenditures to unexpected movements in GDP.

[8] To help the model capture some of the exogenous structural changes influencing GDP growth during the 1990s—that is, the breakdown of Finland's trade relations with the former Soviet Union, the banking crisis in the early 1990s, and the rise of the export-driving ICT sector—both a linear time and a non-linear time trend are included. The latter is a Hodrick-Prescott approximation of actual GDP growth with the smoothing factor set to 1,000.

[9] The (log-) estimates of the coefficients are transformed into dollar-for-dollar based on sample means of the revenue-to-GDP and expenditure-to-GDP ratios, respectively. The dynamic results discussed below are treated equivalently.

[10] Comparisons with BP are based on their model assuming a deterministic trend in the VAR (as in our model). The data definition used by BP differs from ours, but comparable results following their specifications can be obtained. All CH results have been transformed into dollar-for-dollar terms assuming a revenue-to-GDP and expenditure-to-GDP ratio of 0.45.

Figure 2. Finland: VAR Model, 1991-2002



Source: Fund staff calculations.

Table 1. Finland: Summary Coefficients

$$e^r = a_1 \ e^y + a_2 \ u^g + u^r$$
$$e^g = b_1 \ e^y + b_2 \ u^r + u^g$$
$$e^y = c_1 \ e^r + c_2 \ e^g + u^y$$

Parameter	Value (dollar-for-dollar terms)	
a_1	0.5	1/
b_1	-0.2	1/
a2	0	
b2	0	
c1	-0.1 (-1.4)	2/
c2	0.2 (0.95)	2/

Source: Fund staff estimates.

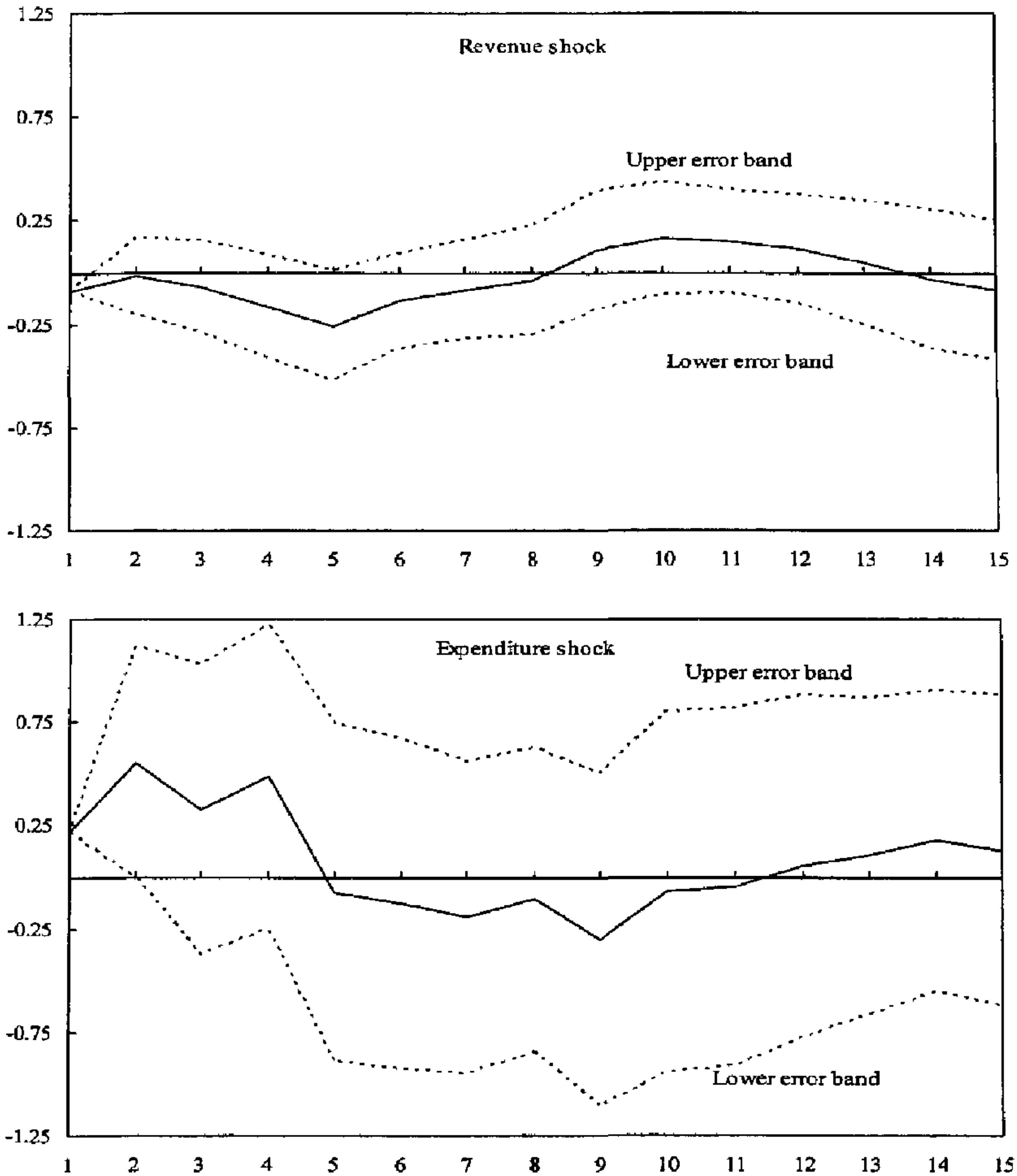
1/ The elasticity is around two times as high.
2/ The elasticity is around two times as low.

Assuming higher elasticities of 1.1 for revenues and -0.5 for expenditures leaves the contemporaneous impact of revenues on GDP unchanged, while the impact of expenditures on GDP increases somewhat to 0.3 on a dollar-for-dollar basis. However, neither estimate is significant even at the 10 percent level. Choosing lower elasticities, 0.9 for revenues and -0.3 for expenditures, yields an expenditure impact on GDP of just 0.1 and unchanged estimates for the impact of revenues—both insignificant.

23. Having identified the contemporaneous effects, the VAR model can be employed to take a closer look at the *dynamic effects* of discretionary fiscal policy on the economy. Figure 3 shows the intertemporal responses of GDP to structural shocks to revenues and expenditures. The impulse responses are expressed in dollar-for-dollar terms. The results incorporate our baseline assumptions on revenue and expenditure elasticities, but comparable figures are obtained using the alternative assumptions on elasticities discussed above.

24. The impact of a shock to revenues on output is very modest. It peaks at about -0.25, after five quarters. This is notably lower than the results found by BP for the United States (with the strongest effect at about -0.75 after around six quarters) and it also appears to be

Figure 3. Finland: Impulse Responses of GDP to Discretionary Fiscal Policy Shocks 1/



Source: Fund staff calculations.
1/ Figures show the reaction of quarterly real GDP with a 2-standard-error band. All effects are expressed as dollar-for-dollar.

lower than the results found by CH for Spain.[11] Moreover, the impulse response function is not statistically different from zero even at its maximum.

25. The dynamic impact of an expenditure shock on GDP, while statistically significant only in the second period, is larger than that of a discretionary change in revenues. It peaks at about 0.5 in the second quarter and reaches a level close to this in the fourth. With a significant share of government expenditure directly impacting GDP through public consumption and investment, this is perhaps not surprising. But the effect remains small compared to the results reported by BP for the United States, where the impact measures close to 1 in the first quarter and rises to even higher levels around quarter fifteen. A comparison with Spain based on CH would seem to support a similar conclusion.

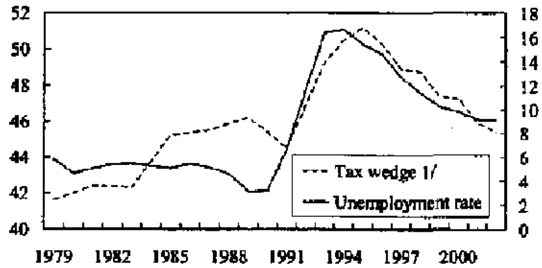
26. While these VAR estimates for Finland are in some ways different from VAR estimates for the United States and Spain, they are roughly consistent with simulations with macro models for Finland and with estimates from a panel of EU countries conducted at the IMF (see paragraph 12).

D. Some Considerations on the Longer-run Impact of Tax Changes

27. The limited effectiveness of revenue-centered fiscal policy is a particularly striking feature of the findings described in Section C. But due to the short-term nature of the VAR approach, VAR models are ill-equipped to capture the longer-term effects of tax changes on employment and growth, especially in less-than-fully flexible labor market environments such as in Finland, where labor demand and supply reactions are likely to take time.

28. Identifying the longer-term impact of changes in labor taxes on employment and growth is not an easy task, however. A first complication arises because tax rates are endogenous to macroeconomic and fiscal developments (Koskela and Uusitalo 2003). While there is indeed a striking correlation between the tax wedge on labor and unemployment in Finland (Figure 4), the increase in labor taxes during the early 1990s was required to balance the fiscal accounts after the (preceding) drastic increase in unemployment in the wake of the breakdown of Finland's trade with Russia and the consequences of the banking crisis. Moreover, a variety of taxes is

Figure 4. Finland: Labor Taxes and Unemployment, 1979-2002



Sources: OECD; and Fund staff calculations.
1/ Income tax plus employee and employer contributions in percent of labor costs for single persons without children.

[11] The comparison for Spain is less straightforward, since CH present nonstandardized impulse responses. However, based on the time profile of the GDP reaction to the tax shock and the high contemporaneous impact described above, the dynamic effect would clearly appear to exceed the one reported in Figure 3.

likely to influence employment, including payroll, income, and consumption taxes. Their respective impact on the labor market will depend, inter alia, on the elasticities of labor demand and supply and the degree to which economic agents perceive the underlying fiscal decisions as permanent. This further complicates identifying the relation between taxes and employment.

29. But even in the face of these difficulties, there seems to be a significant and positive longer-term correlation between labor taxes and unemployment (and a negative correlation with employment) in Finland. Roughly half of a change in income tax rates is translated to a change in labor costs, which, in turn, influence employment. By comparison, the labor-cost and employment effect of changes in payroll taxes (e.g., social security contributions for employers) seems to be somewhat higher, as more than half of these taxes are shouldered by employers (Koskela and Uusitalo 2003). Honkapohja and others (1999), using industrial-level time series data, find that the long-run elasticity of employment to changes in labor cost is about -0.7 in Finland. This implies that a 1 percentage point decrease in labor taxes would increase employment by about 0.4 percent. While there are some indications that the longer-term impact of tax changes in Finland (as well as in other Nordic economies) could be somewhat weaker than in other countries (Daveri and Tabellini, 2000), the order of magnitude of these effects is clearly nonnegligible.[12]

30. The discussion holds a crucial message for policy makers, putting the VAR-based findings into perspective: even though tax cuts might have only limited impact as a short-term macroeconomic policy tool, cutting taxes on labor can very well have a significant positive impact on employment and output in the longer run—but this requires keeping the public finances healthy so that tax cuts are not seen as being only of a temporary nature.

E. Concluding Remarks

31. The main message of this study is that, at least over the short run, discretionary fiscal measures have only a modest impact on the Finnish economy. While fiscal policy influences GDP in the short-run in line with our priors, the impact remains small compared to results reported from related studies on the U.S. and Spain. This difference in policy impact, often explained by the relative smallness and openness of the Finnish economy, is in line with findings in other recent studies conducted by the European Commission and the IMF.

---

[12] There are some indications that the longer-term impact of tax changes in Finland (as well as in other Nordic economies) could be somewhat weaker than in other countries (Daveri and Tabellini, 2000). Koskela and Uusitalo (2003) argue this could be due to the more centralized bargaining systems in the Nordic region which might make wage formation less sensitive to changes in taxation compared to less centralized systems. The explanation is not fully compelling, however. In the Finnish case, for example, the government has used the centralized wage negotiations of recent years to condition tax policy on wage behavior, fostering wage moderation through promises of tax cuts on labor.

32. A caveat to these findings is that the applicability of the structural VAR approach to the Finnish case is somewhat limited by the lack of data. Comprehensive high frequency fiscal data are not readily available and, to the extent it can be constructed, covers only a relatively short period marked by notable real shocks that are hard to model endogenously. This calls for a degree of caution in interpreting the results.

33. A second qualification to the VAR-based approach—whose importance to policy makers is hard to overplay—concerns its short-term perspective. Labor market decisions can play an important part in the transmission of fiscal policy decisions to the real economy, but they also might require time, especially in less flexible institutional environments. And indeed the discussion of existing Finnish evidence reveals that the impact of tax changes, which the VAR-model characterizes as rather small, is more prominent in the longer run. Thus, while the short-term impact of tax cuts seems small, they would seem to be a more effective tool to fostering growth over longer time periods, especially when undertaken in the context of expenditure consolidation to ensure that the tax cuts are not seen as reversible and that a sustainable fiscal position is obtained.

REFERENCES

Aarle, B., H. Garretsen, and N. Gobbin, 2001, "Monetary and Fiscal Policy Transmission in the Euro area: Evidence from a Structural VAR Analysis," paper presented at a Vienna Institute for International Economic Studies seminar, Vienna, Austria, January.

Blanchard, O., and R. Perotti, 2002, "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *The Quarterly Journal of Economics* (November).

Brunila, A., M. Buti, and J. in 't Veld, 2002, "Fiscal Policy in Europe: How Effective are Automatic Stabilizers?," *European Economy* (September), (Brussels: European Commission).

Céspedes, L., and A. Hoffmeister, 2003, "Fiscal Policy and Macroeconomic Volatility in Spain: An Empirical Assessment", *Spain: Selected Issues*, IMF Country Report No. 03/41 (Washington: International Monetary Fund).

Daveri F., and G. Tabellini, 2000, "Unemployment, Growth and Taxation in Industrial Countries," *Economic Policy*, (April), Vol. 15, No. 30, pp. 47–104, (Oxford, U.K. and Boston, USA: Blackwell Publishers Ltd.).

Hemming, K., and Mahfouz, 2002, "The Effectiveness of Fiscal Policy in Stimulating Economic Activity—A Review of the Literature," IMF Working Paper 02/208 (Washington: International Monetary Fund).

Honkapohja, S., E. Koskela, and R. Uusitalo, 1999, "Employment, Labor Taxation, and the Balance of the Public Sector," *Finnish Economic Journal*, Vol. 95, pp. 74–95.

Hunt, B., and D. Laxton, 2003, "Some Simulation Properties of the Major Euro Area Economies in MULTIMOD," IMF Working Paper 03/31 (Washington: International Monetary Fund).

International Monetary Fund, "Adopting the Euro in Central Europe: an Examination of Policy Issues, " (forthcoming: Washington).

Koskela, E. and R. Uusitalo, 2003, "The Un-intended Convergence: How the Finnish Unemployment Reached the European Level," Mimeo, January (Helsinki: University of Helsinki).

Perotti, R., 2002, "Estimating the Effects of Fiscal Policy in OECD Countries", paper presented at an ISOM conference, Frankfurt, Germany, June.

Organization for Economic Cooperation and Development, 2003, Economic Outlook, Sources and Methods, No. 73, June (Paris).

II. Estimating the Output Gap in Finland[1]

A. Introduction

1. Assessing the degree of slack in the economy is important for a number of analytical reasons and the output gap—which measures the deviation of GDP from its potential—is a frequently used indicator for this purpose. Among the reasons, 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. A third reason is that the magnitude of the output gap is relevant for assessing economic growth—that is, can variations in actual growth be attributed to cyclical factors (such as slow growth in trading partner economies) or to a longer-term change in potential growth?

Box 1. Volatility in the Real Economy

During the last 15 years, the Finnish economy has undergone (i) one of the sharpest recessions among euro-area countries; (ii) a strong boom period in the second half of the 1990s—with real GDP growth averaging about 5 percent between *1994–2000*, led by the Information and Communication Technology (ICT) sector, especially in the second half of the period; and, lately, (iii) a significant slowdown in economic growth. During the crisis in the early 1990s, 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 volatility of economic conditions is captured in the table to the right by showing that the standard deviation of real GDP growth rates was comparatively high.

Growth: Finland vs. Other Euro Area Countries

	Real GDP	
	growth 1/	volatility 2/
Finland	1.8	3.9
Euro area 3/	2.3	1.8
Portugal	2.9	2.4
Greece	2.4	2.5
France	1.9	1.7

Sources: IFS; and WEO databases.
1/ Growth refers to the average of quarterly year-on-year growth rates.
2/ Volatility is measured by the standard deviation of the quarterly growth rates described in footnote 1.
3/ Unweighted average excluding Finland; Ireland is also excluded due to a lack of data over the entire period, while data for Germany starts in 1992.

2. Two broad 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

[1] Prepared by Helge Berger and Andreas Billmeier. The authors would like to thank the Finnish authorities for their interest in the subject—which served as a catalyst for this paper.

can be viewed as akin to different economic concepts of potential output. Under the first approach, potential output is driven by productivity shocks, and temporary deviations of actual output 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. 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.

3. As this paper will show, the pronounced volatility of output in Finland makes it particularly difficult to estimate potential output and there is, therefore, considerable uncertainty about the size of the output gap. The observed volatility is due, at least in part, to the development and swings in the performance of the Information and Communication Technology (ICT) sector. At the same time, distinct problems are associated with each of the two 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 must rely on estimates of both labor input at full employment and the stock of effective capital input in production as components of a stable production function.[2] Moreover, 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.

4. The focus of this paper is to estimate the output gap using different methodologies and, while highlighting the uncertainties, also to provide some guidance and make some

[2] 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).

judgments about its size.[3] This is done by establishing a number of "intuitive criteria," against which the various gap measures could be assessed. First, the mean of the gap measure should be close to zero over longer time horizons. Second, the gap measure should produce "reasonable" maxima and minima, in terms of magnitude. Third, 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.

5. 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. However, expectations of slowing inflation and some rise in unemployment in 2003 (see, for example, the Finland–Staff Report for the 2003 Article IV Consultation, SM/03/313, 9/9/03) cast a degree of doubt on results indicating a still positive gap. Overall, the production function approach appears to have a number of comparative strengths, in part because it incorporates a substantial amount of additional information on the economy (such as developments related to the labor market). This method suggests that output in Finland remained below its potential in 2002 by about 1 percent; and, with actual growth likely to stay below potential in 2003, the output gap is expected to widen somewhat. Nevertheless, the uncertainty surrounding the various point estimates of the output gap is pronounced.

B. Measures of Potential Output and the Output Gap

6. Defined as the difference between actual (y_t) and unobservable potential output ($ypot_t$), the output gap (gap_t) is itself an unobserved variable:

$$gap_t = \frac{y_t - ypot_t}{ypot_t} \tag{1}$$

[3] Relatively little empirical work has been done comparing different output gap estimation methods. However, 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 towards 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). De Masi (1997) reviews related research done at the IMF.

In what follows, purely statistical approaches are investigated first, followed by techniques that draw more on economic models. In both, the estimated underlying trend is assumed to coincide with potential output, and the two expressions are used interchangeably.

An overview of statistical measures of potential output

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

Arithmetic trends

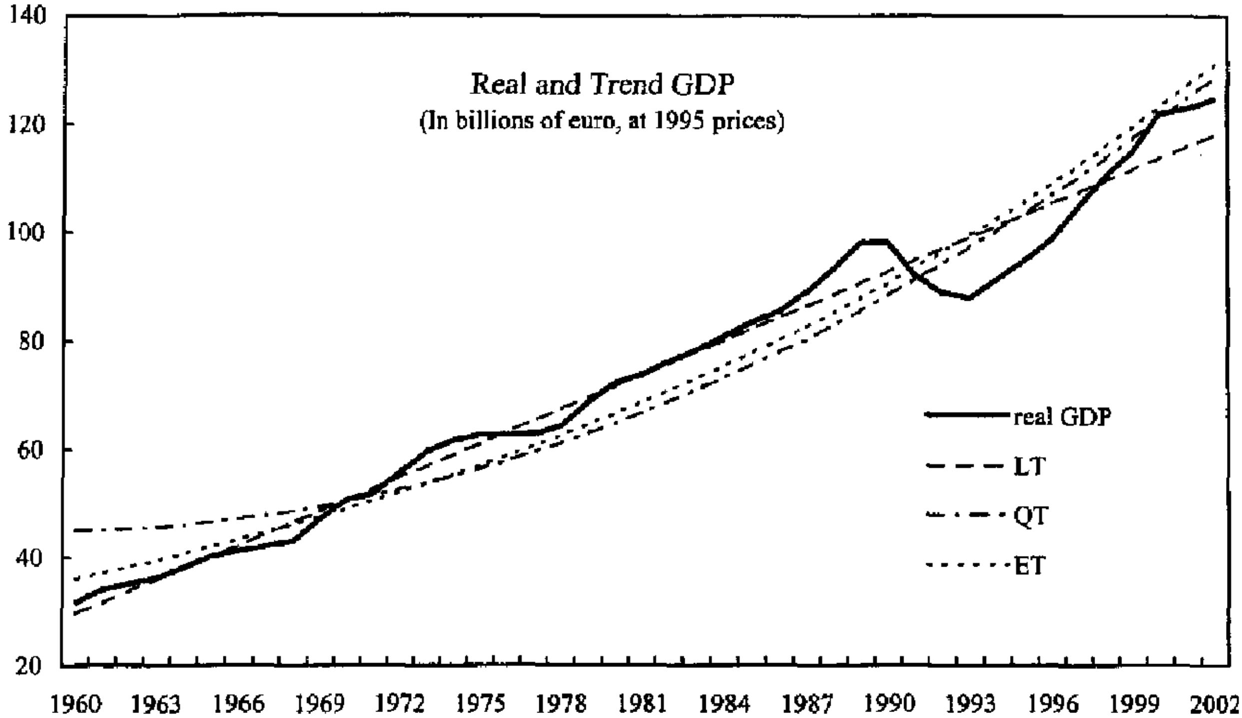
8. With linear, quadratic, or exponential detrending, potential output is assumed to follow a deterministic process, which can be approximated by a polynomial or exponential function of time. Furthermore, the trend and cycle are assumed to be uncorrelated. While these measures lack theoretical underpinning,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

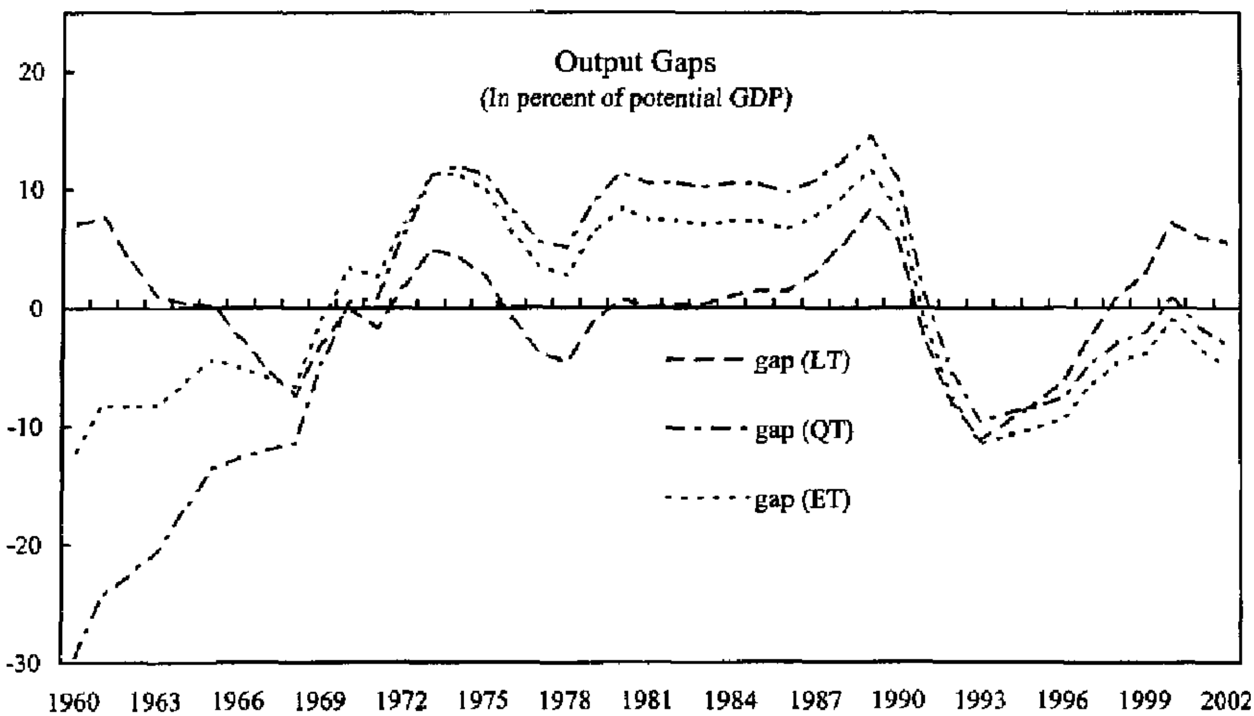
9. Results from linear (LT), quadratic (QT) and exponential (ET) detrending are somewhat similar in shape, but indicate rather unrealistic swings in the output gap (Figure 1). The upper panel contains the raw and the smoothed series, whereas the lower panel shows the resulting output gap. All measures 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 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 in paragraph 4), the quadratic trend fails to fulfill the criterion of an average output gap of 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 huge 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 in principle not implausible, raise concerns about the underlying methodology. These, in the next section, methods are used which attempt to glean more information from the original series by applying statistical filtering techniques.

---

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

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





Sources: IFS, WEO; and Fund staff calculations.
Note: Trend abbreviations as described in the main text.

Univariate statistical filters

10. Statistical filters can extract information either in the common 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 filter

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

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

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. 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.

12. 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 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

[5] See, for example., 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 (2), 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.

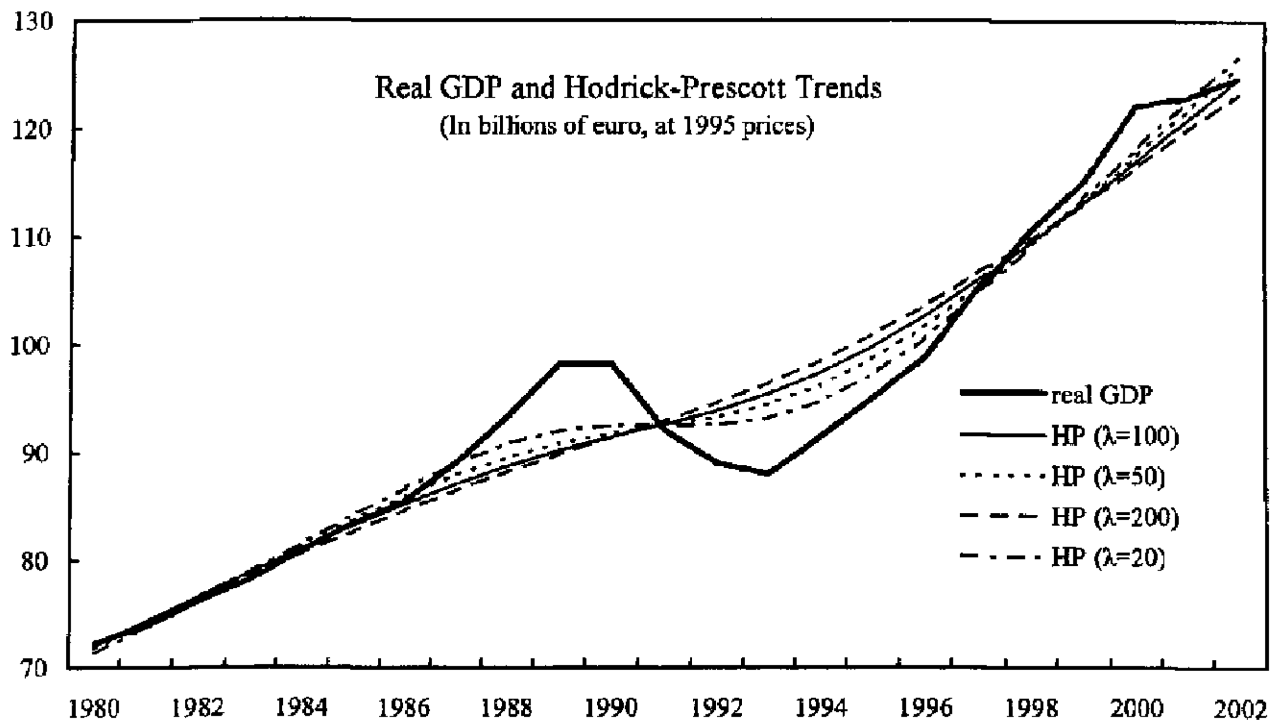
practice has been to extend the observation period by forecasting. The discussion that follows focuses on the consequences of differing assumptions on the smoothness of the trend, and on the end sample bias in the case of Finland.

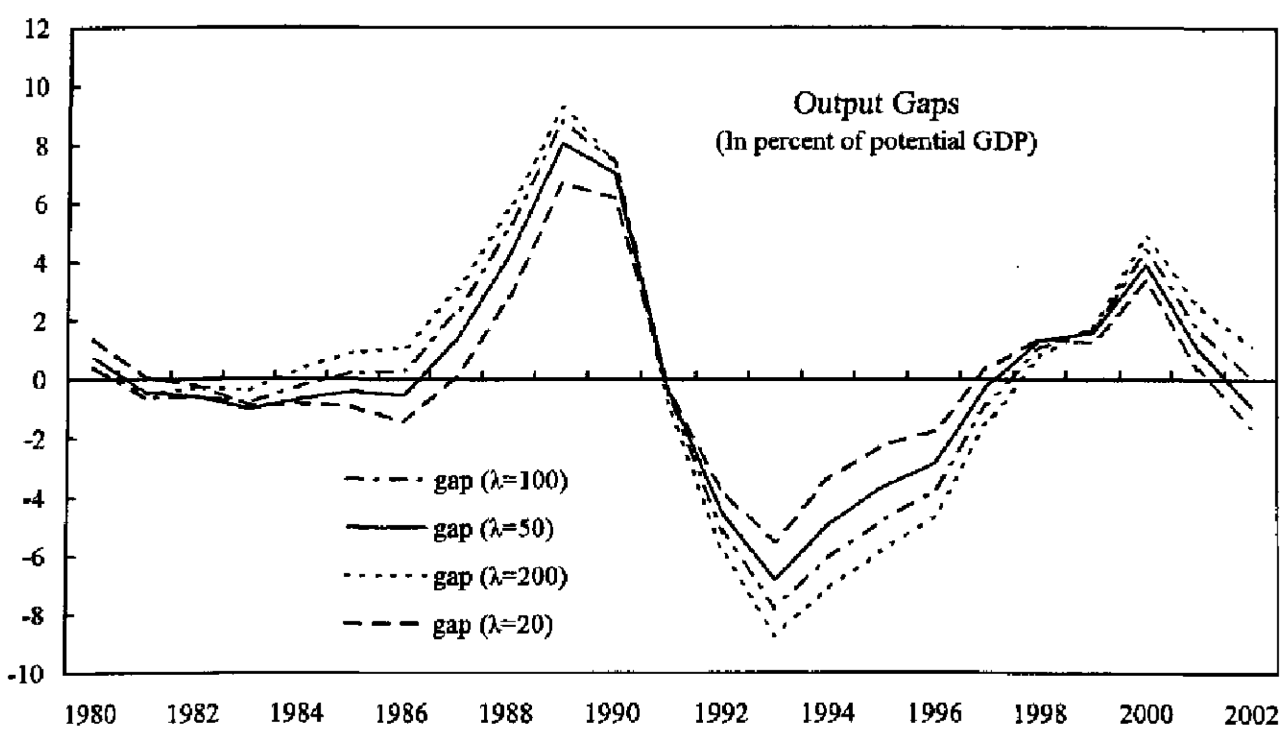
13. While various HP filters find a similar pattern of the output gap, the assumption on the smoothness of the trend has strong implications for the magnitude of the gap at the end of the observation period. In Figure 2, the results of HP filtering of Finnish real GDP—for various parameter values for λ—over the sample period 1980–2002 are shown.[7] Trend estimates are provided for $\lambda = 20, 50, 100, 200$ with a lower parameter value indicating a smoother trend. As far as the intuitive criteria are concerned, the estimated gap using all filters are close to zero on average over the sample period 1980–2002, 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. In addition, 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 actual GDP was above trend by 1.1 percent of potential GDP, whereas the least smooth trend indicates actual output was 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 end-point problem.

14. The severity of the endpoint bias is depicted in Figure 3 by adding a few forecasted values of the variable to be filtered. Accordingly, values for real GDP were added using the WEO forecast, in the first step until 2005, and until 2008 in the second step. The assumed medium-term recovery of the Finnish economy has an effect on the estimates of the output gap in 2002. For both assumed parameter values—$\lambda = 20, 100$—gap estimates based on the extended series are higher than for the original series. 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 extension of the estimation period is even more pronounced for the less smooth trend, $\lambda = 20$. In this latter case, both estimates indicate a positive output gap on the order of 0.3 percent of GDP, whereas the gap was negative using the original series and equal to -1.6 percent of trend GDP.

[7] Estimation spans the period 1960–2002.

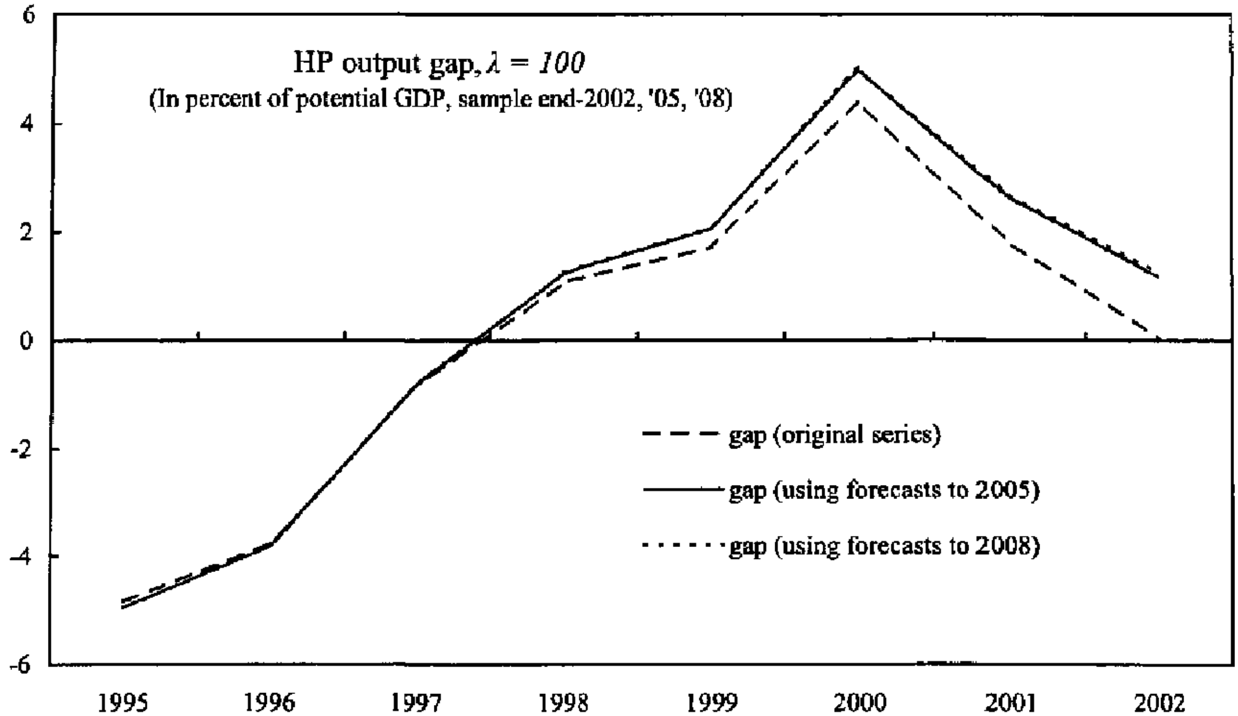
Figure 2. Finland: Hodrick-Prescott Detrending, 1980-2002 1/

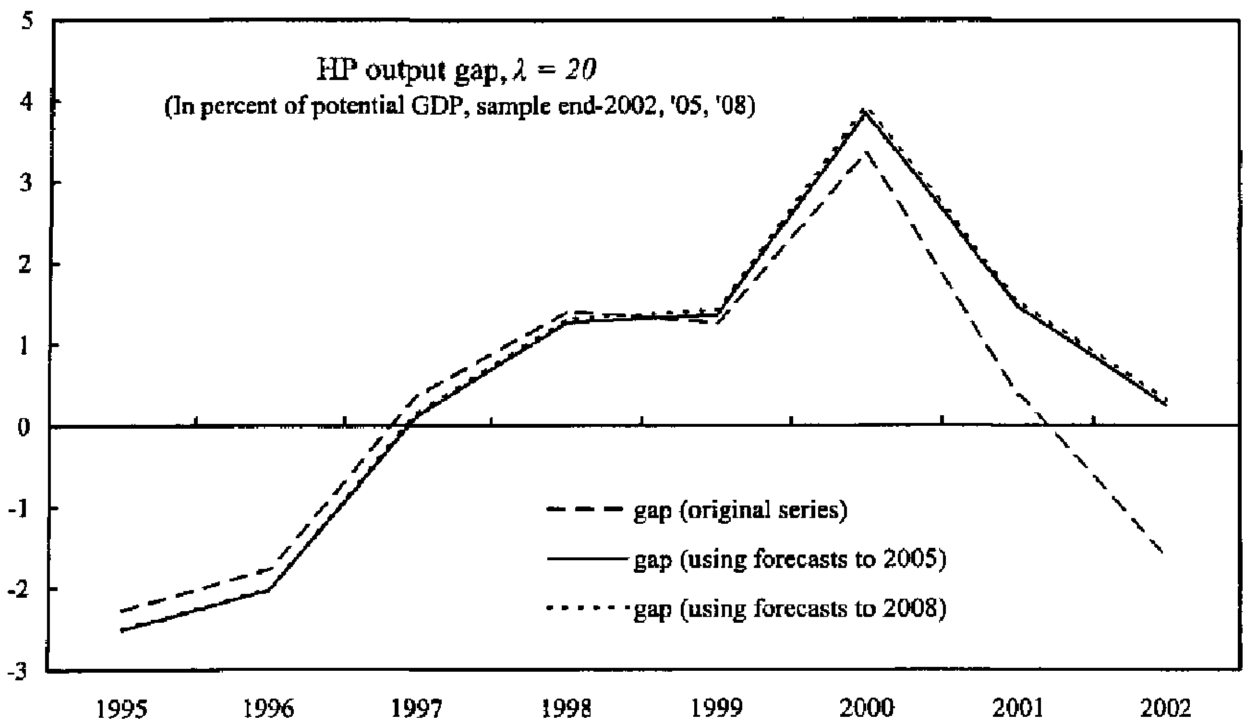




Source: IFS; and Fund staff calculations.
1/ The last observation of the original time series is 2002.

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





Sources: IFS, WEO; and Fund staff calculations.

15. Results as varied as these underscore the need for a multi-dimensional assessment of the output gap. Regarding the HP estimation approach, there is 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] 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 decomposition

16. The approach pioneered by Beveridge and Nelson (1981) constitutes another decomposition of 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.[9] The decomposition rests on the crucial assumption that innovations in the permanent (trend) and the transitory (cyclical) component are perfectly negatively correlated.[10] Based on inspection of the resulting trend component, an ARIMA(1,1,1) model was chosen to represent the structure of the underlying series.[11]

17. 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

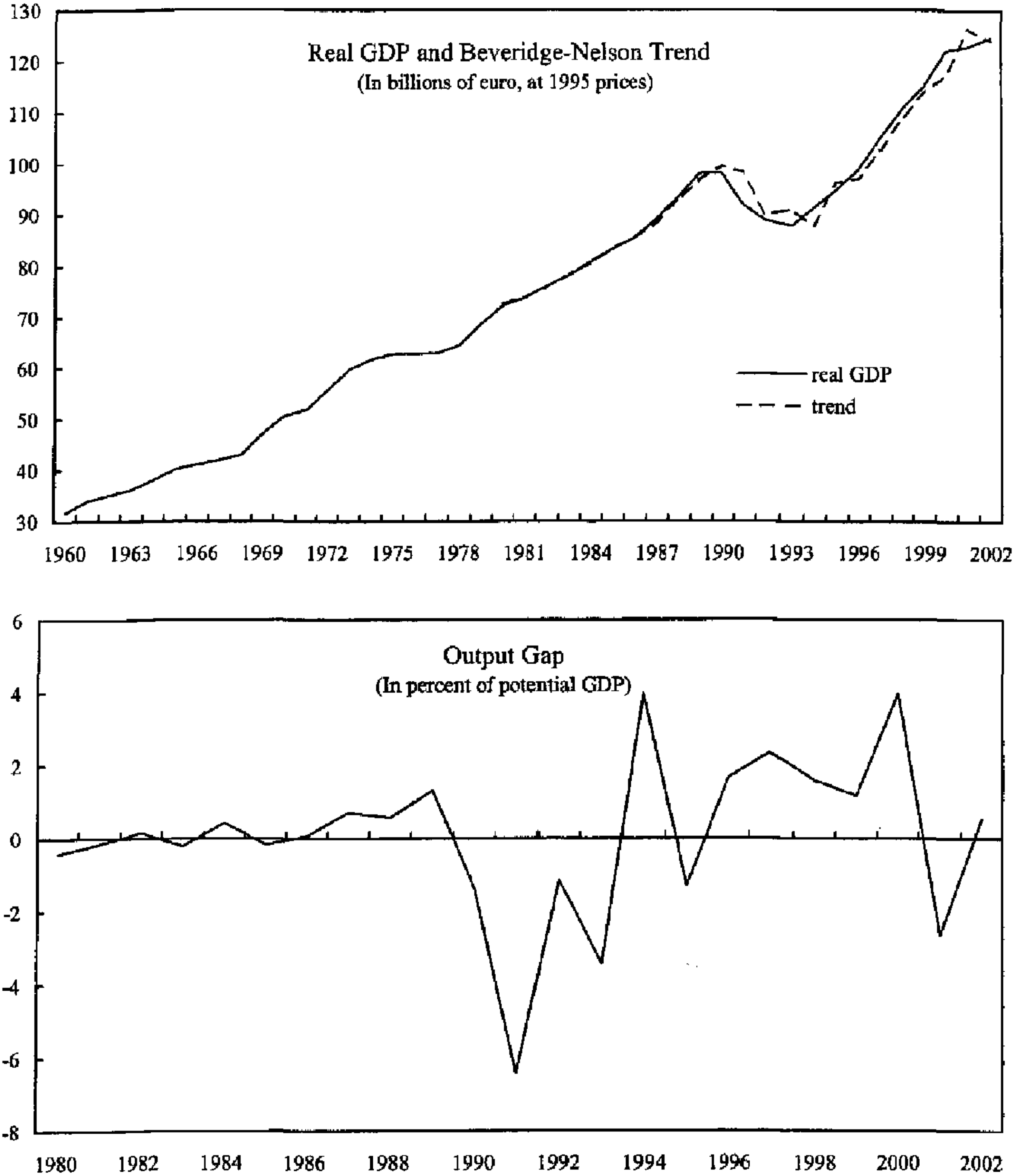
[8] In what follows, given the very similar results for the time series ending in 2005 and 2008, the 2005 end point results are reported when comparing the HP filter to other trend estimation techniques.

[9] In this 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, $d = 1$.

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

[11] Note that 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, resulted in unrealistic volatility of the trend component around the peak-trough period in the late 1980s/early 1990s.

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



Source: IFS; and Fund staff calculations.

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

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.

18. 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.

Frequency domain filters

19. 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.[12] In this context, Corbae and Ouliaris (2002) provide a consistent band pass filter for nonstationary data.[13]

[12] 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, see Baxter and King (1999). They argue that upfront detrending of the series in order to apply discrete Fourier transforms involve 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.

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

20. 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 was 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.

21. The results from the frequency domain approach yield a few interesting particulars, 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 to the trough in 1993 of -3.5 percent of potential output. This is in stark contrast with the result from the Beveridge-Nelson decomposition. With 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 the duration is much larger). 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.

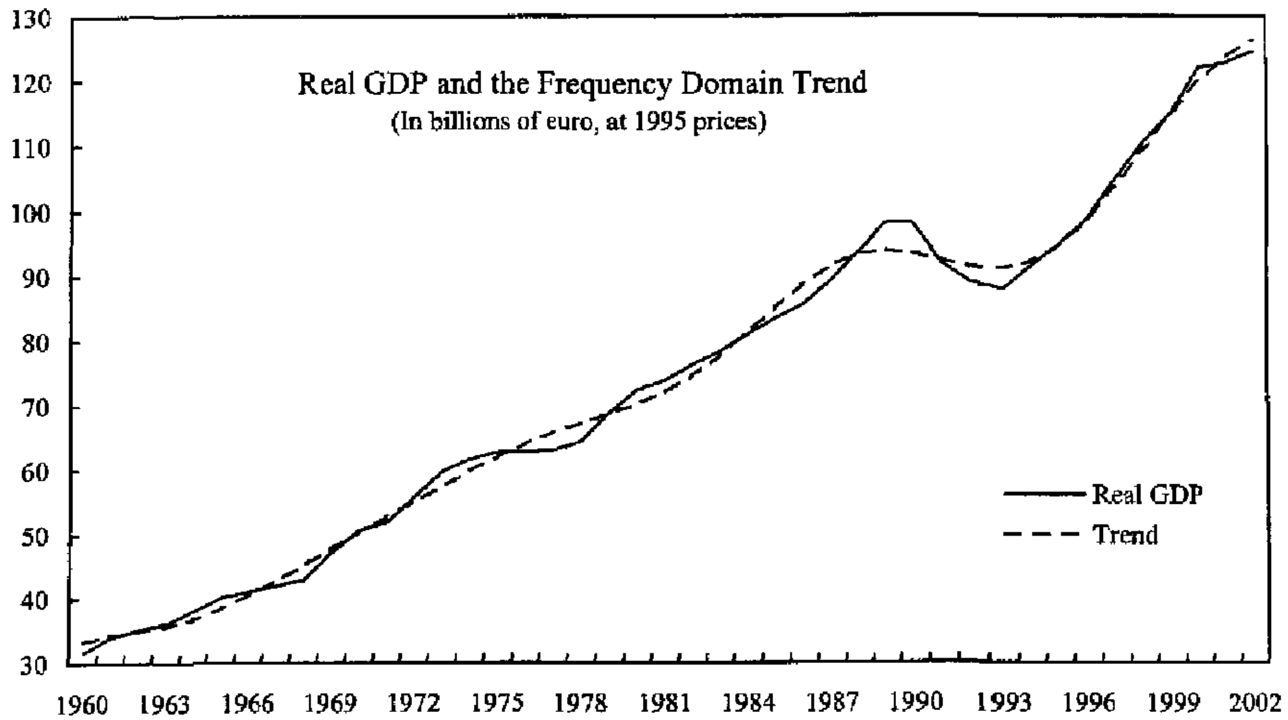
22. While the major advantage of the frequency domain approach and, indeed, other statistical methods is their simplicity, they are subject to the criticism of lacking foundation in economic theory. Thus, the next section turns to theory-based models of trend GDP and the output gap.

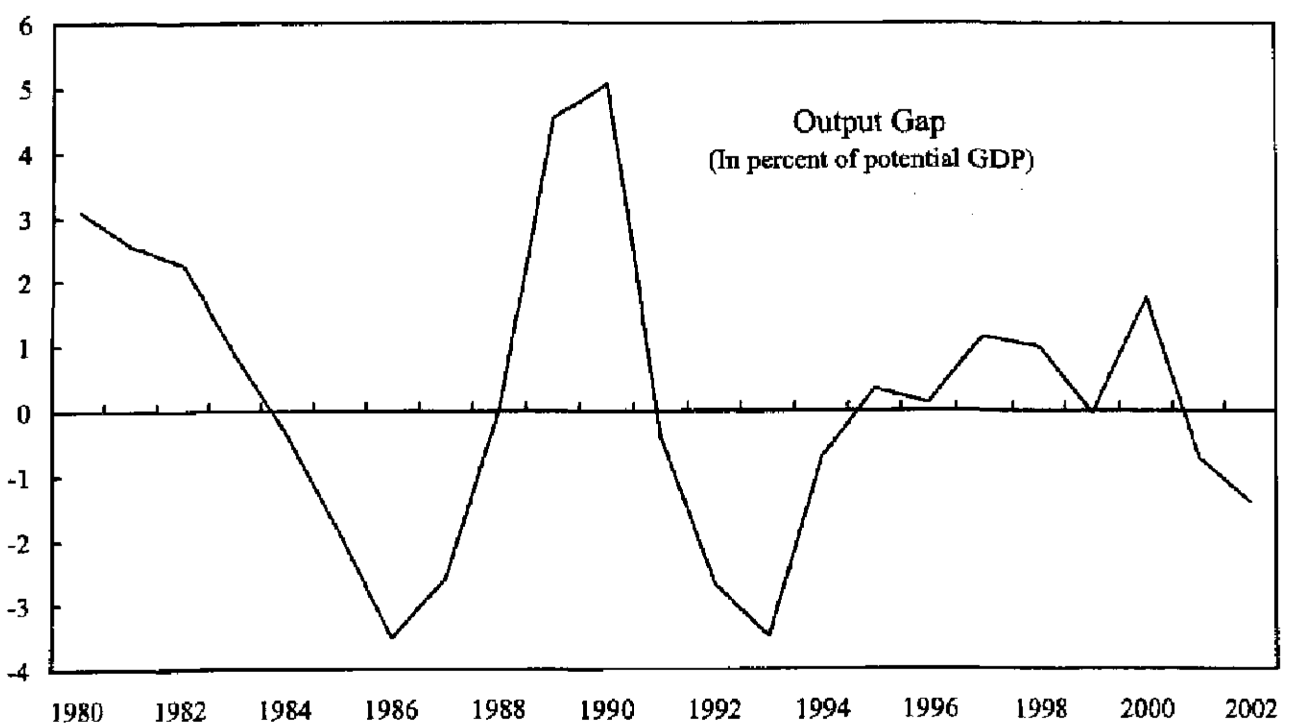
Theory-based measures

23. 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. 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 accomodate

[14] Given that the data are annual, a periodicity for the business cycle between 2 and 8 years has been assumed. Experiments with somewhat longer and shorter cycles yielded broadly similar results.

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





Sources: IFS; and Fund staff calculations.

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.

Structural VAR: the Blanchard-Quah approach

24. The appeal of the approach by Blanchard and Quah (1989) to the identification of structural shocks in a VAR stems from its compatibility with a wide array of theoretical models. 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 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.[15]

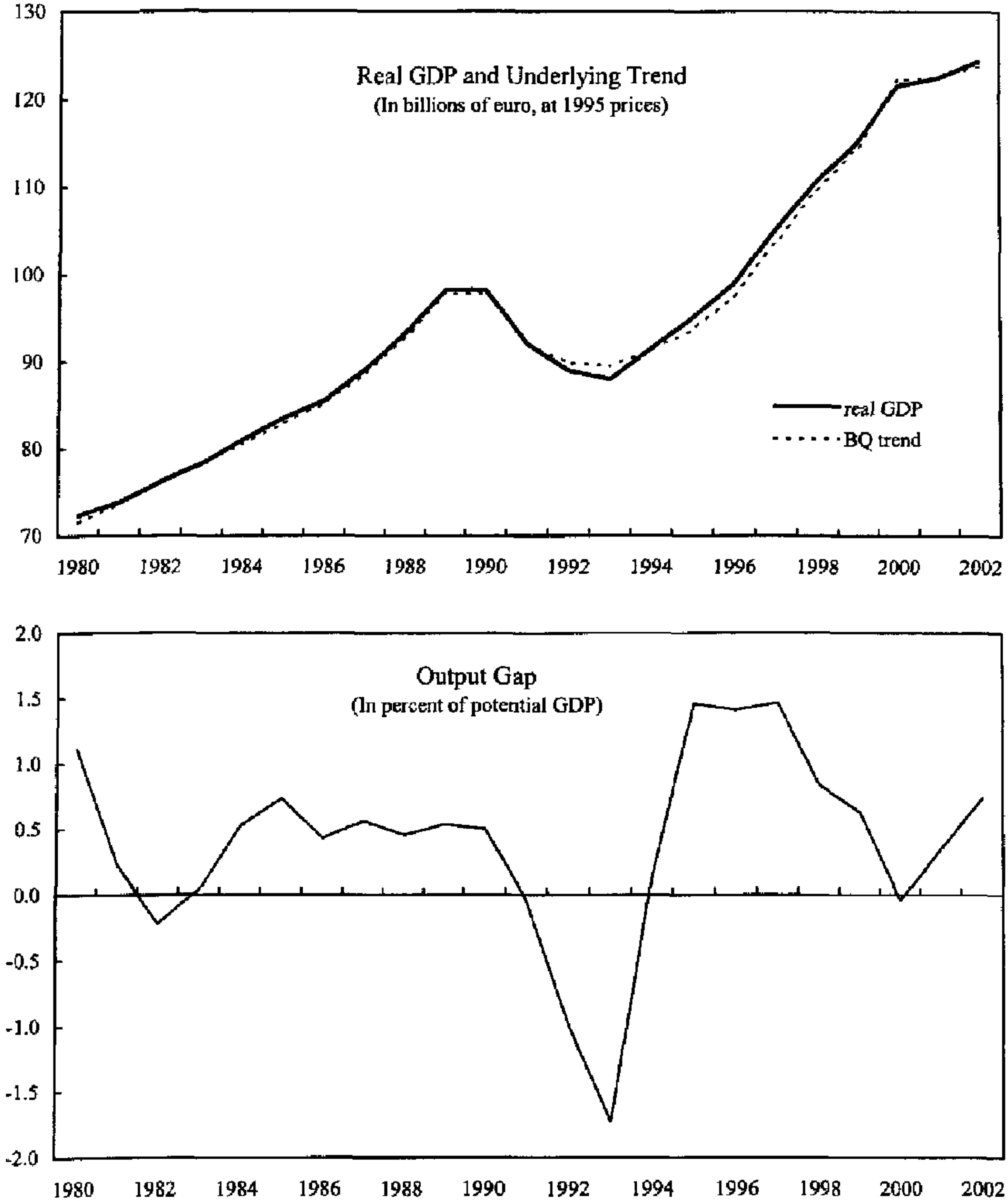
25. In the present context, potential output is associated with cumulated supply shocks, whereas the output gap reflects cyclical (temporary) swings in aggregate demand. This approach, hence, benefits from explicit economic foundations. Furthermore, the gap—identified as the demand component of output—is 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. Hence, while appealing on theoretical grounds, the applicability of the Blanchard-Quah approach might be limited.

26. In terms of the intuitive criteria, the estimated output gap—as characterized by the Blanchard-Quah decomposition—shares only a limited amount of characteristics with other approaches (Figure 6).[16] For example, a negative output gap characterizes the crisis period in

[15] 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.

[16] The VAR model underlying the estimation includes, in addition to a constant, four 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, mostly on statistical grounds.

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



Sources: IFS; and Fund staff calculations.

the early 1990s with the trough in 1993. However, the magnitude of the gap seems to be limited, barely reaching 1.5 percent of trend GDP. A number of other findings are somewhat surprising as well. First, the output gap during the crisis in the early 1990s is closed as early as 1994, barely one year after the trough. Instead, this decomposition yields a positive output gap for the period 1994 onwards, with one little dip into barely negative territory in 2000. Second, the (positive) gap seems to be biggest in the immediate aftermath of the crisis, 1995 –97, contradicting 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 been closed yet. Third, the overheating of the Finnish economy in the late 1980s is only insufficiently captured by the measure, with actual output being only 0.5 percent above potential during 1984–90.

27. Overall, while the measure is compatible with some of the intuitive criteria discussed earlier, the small magnitude of the gap and in particular the long positive gap since 1994 cast doubt on the reliability of the Blanchard-Quah approach to the identification of the output gap. With regard to the years 1995–97, the approach seems to underestimate growth of potential output, attributing strong real growth mostly to demand side effects, and, hence, evoking inflationary pressures, as captured by the positive output gap. One way to curb the problem of assigning shocks to demand or supply origins is to start from a growth-accounting perspective.

Production function approach

28. The production function 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.[17] A convenient functional form is the Cobb-Douglas type, where output Y_t depends on labor L_t and capital K_t, as well as the level of total factor productivity TFP_t:

$$Y_t = TFP_t K_t^{\alpha} L_t^{\beta} \tag{3}$$

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:

$$tfp_t = y_t - \alpha k_t - (1 - \alpha) l_t \tag{4}$$

[17] Early work on the production function approach includes Artus (1977). Subsequent research has refined the approach in various directions, see, for example, De Masi (1997).

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

$$L_t^* = LF_t\left(1 - UR_t^*\right) \tag{5}$$

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.[18] Potential output can be written (in logs) as:

$$y_t^* = \alpha k_t + \left(1 - \alpha\right)l_t^* + tfp_t^* \tag{6}$$

Given the assumption that capital is always employed at full potential, no capacity adjustment is made to the capital stock.[19]

29. 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.[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. In the simplest case, potential output is a linear combination of HP-filtered series.[22] However, the approach can

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

[19] See Everaert and Nadal de Simone (2003).

[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] 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 (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.

also be implemented more flexibly by using more sophisticated filtering procedures, including those that incorporate structural assumptions based on economic theory.

C. The Production Function Approach: An Application to Finland

30. 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).[23] 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 exist, in other words that the Phillips curve is partly vertical.

Preliminaries: is there a long-run Phillips curve?

31. 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, that is, 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 non-vertical 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 United States 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.

32. In the case of Finland, simple tests indicate that, while both wage inflation and unemployment are non-stationary, there is no sign of cointegration, a result conducive to the NAWRU approach.[24] 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.[25] Due to the lack of strong priors in favor of a trend

[23] See Denis, Mc Morrow, and Roeger (2002). This methodology substitutes for more "traditional" approaches—such as the Hodrick-Prescott filter—and, at the same time, unifies the Commission's efforts toward a consistent representation of inflationary pressures in the member countries.

[24] Note that these conclusions also hold for CPI inflation instead of wage inflation.

[25] With three lags, no significant residual autocorrelation emerged, whereas more parsimonious models reveal problems of autocorrelation at the first lag (statistics not reported).

restricted to the cointegrating space, a system with an unrestricted constant was estimated.[26] Likelihood ratio tests (distributed as χ^2(dgf)) of the time series properties reveal that both series appear to be nonstationary;[27] in addition, the unemployment rate can be considered weakly exogenous from a statistical point of view. 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: lag length selection

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

Cointegration: time series properties

DGF	dwage	ur
	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.

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.

33. 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.

[26] In a model with a restricted trend, all variables (including the trend) appeared to be excludable from the system.

[27] 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).

Evaluation of the NAWRU

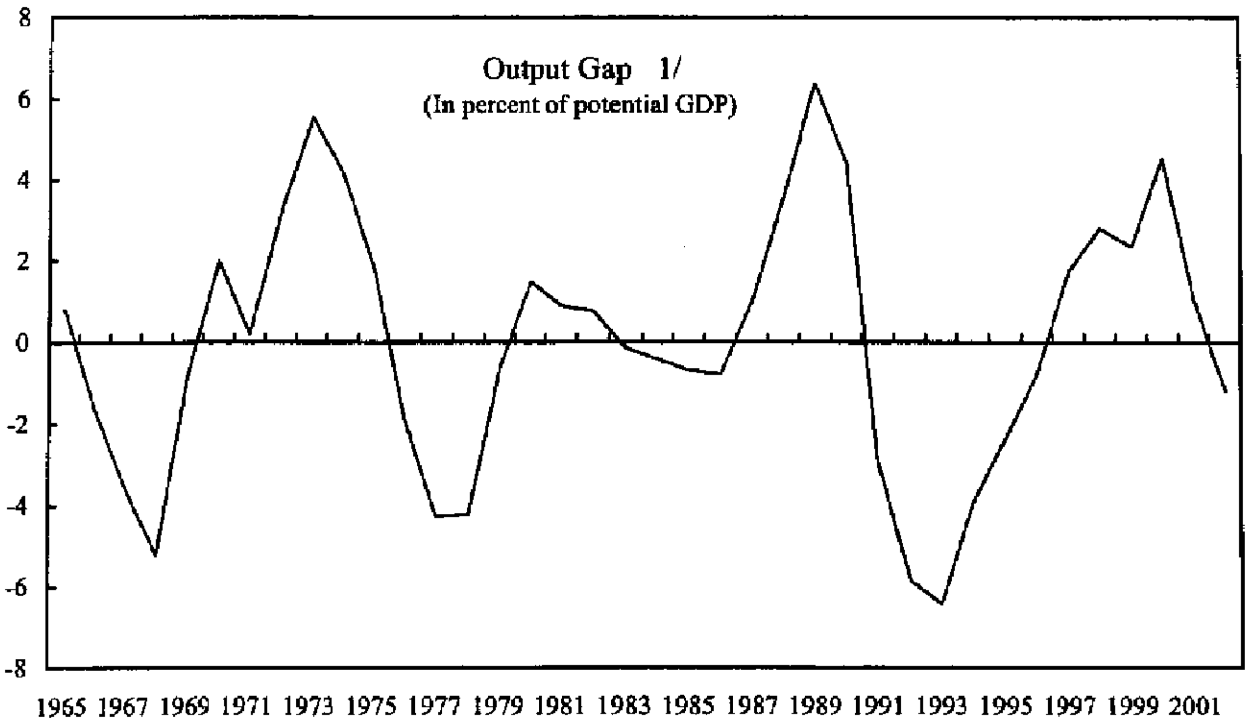
34. Under the latent variable approach, the Finnish natural rate of unemployment—defined here as the NAWRU—is computed using a Kalman filtering process on the observable unemployment rate, to extract the cyclical component. The procedure employs a bivariate model, where the observables "unemployment rate" and "change in wage inflation" (i.e., second differences of wages) play the role of endogenous variables. While the first equation contains a simple decomposition of the observed unemployment rate in trend and cyclical component, the second equation—in principle a Phillips curve—relates the wage inflation to a number of regressors, including lags of wage inflation and the cyclical component of unemployment. Given the error term, wage inflation is assumed to follow an ARMA process. The trend unemployment rate, in turn, serves to determine the (full-employment) stock of labor entering the production function. Estimation takes place in the state-space form, no exogenous regressors are added.[28]

35. 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 (i) 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;[29] and (ii) 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 labor productivity, and measures of the terms of trade resulted in deterioration of the system's performance from a statistical point of view.

[28] A more detailed description of the model set-up can be found in Appendix I. In the terminology of the European Commission, the NAWRU model is known as the "GAP model." As the analysis in the appendix shows, both the assumed representation of wage inflation and the inclusion of additional regressors can have substantial impact on trend unemployment.

[29] Here, the model abstracts from arguments related to the role of trade unions, and centralized wage bargaining.

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



Sources: IFS, European Commission; and Fund staff calculations.
1/ Based on the ARMA(2,3) model.

Implications for the ICT sector and the output gap

36. 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).[30] With 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 slowing 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.[31] 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).[32] Even taking into account two spikes in the late 1960s/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—on top of a combined contribution of capital and labor of less than two percentage points (see Figure 8, upper panel). At the same time, the contribution of the ICT sector was highly volatile: contributions from the ICT sector 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 ½ percentage point, the growth contribution stemming from TFP almost reached 1 percentage point.

37. The measure of the output gap as characterized by the production function approach combines many of the positive 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,

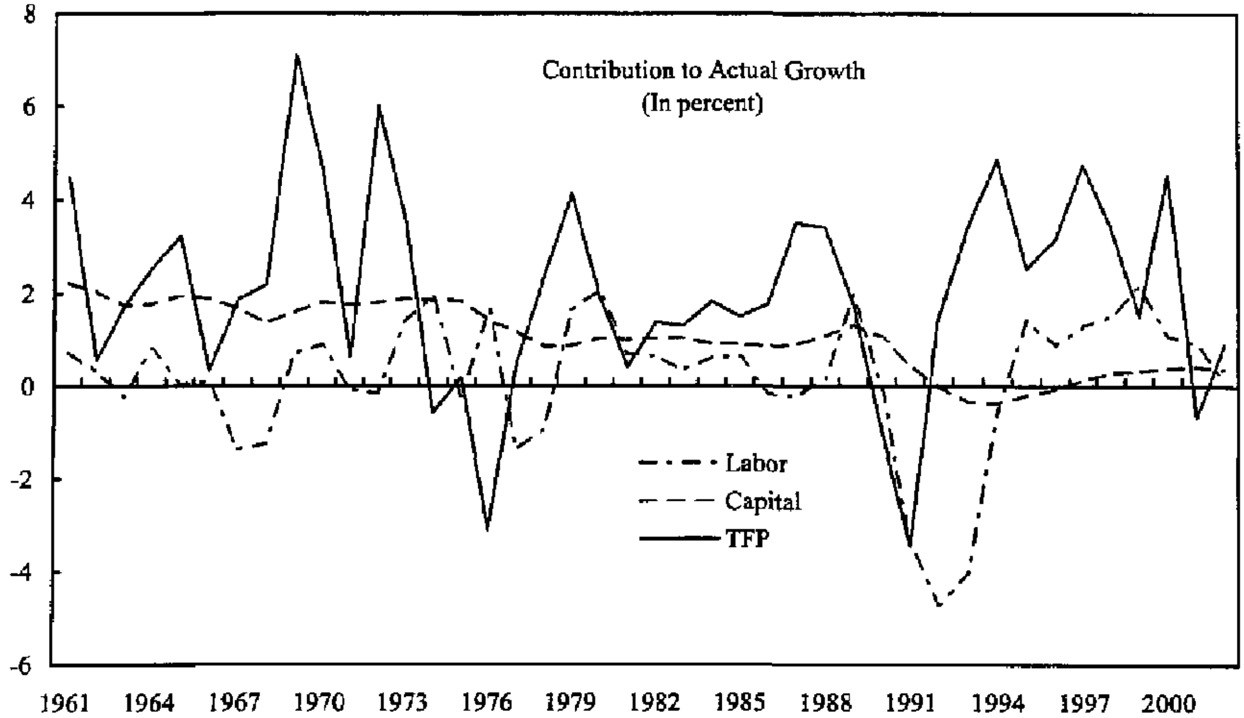
---

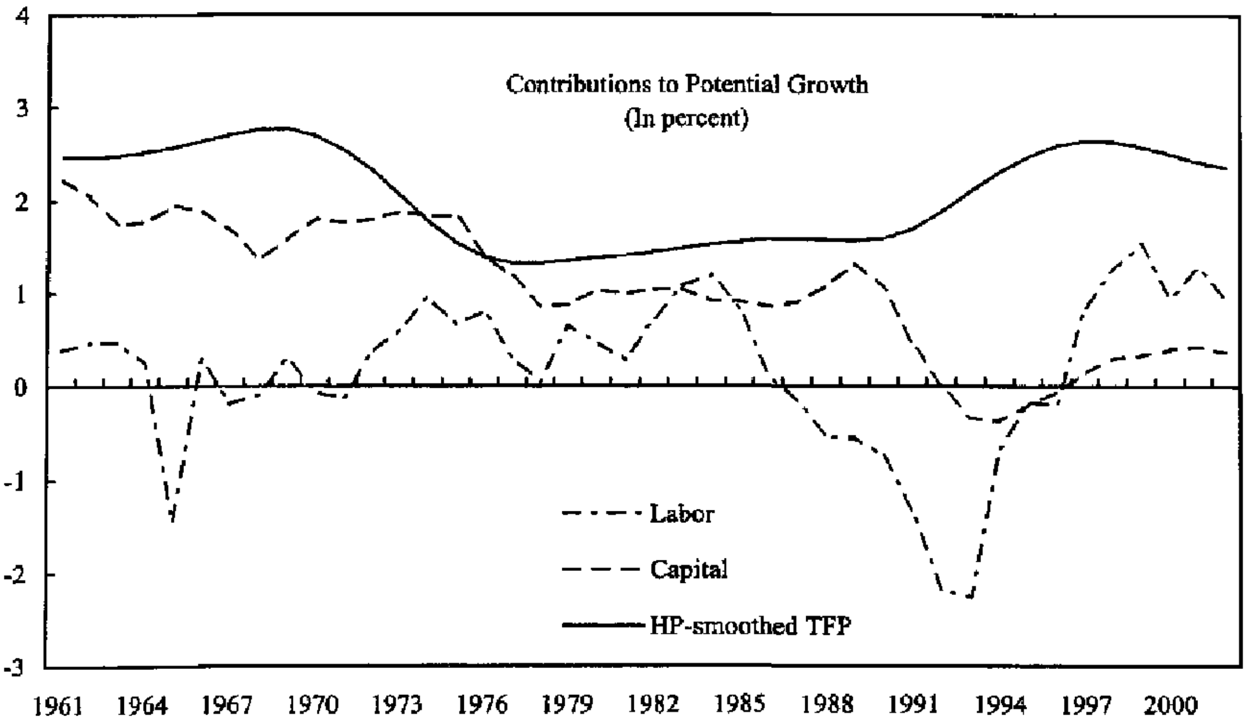
[30] This parallels earlier results: in Jalava and Pohjola (2001), it is found that the importance of "multi-factor productivity," a concept similar to total factor productivity in the present paper, almost doubled in the second half of the 1990s, see also Wagner (2001).

[31] See Table 4 in Wagner (2001) for a qualitatively similar result.

[32] Assuming that TFP is entirely associated with "technological progress" may bias the interpretation of the results to the extent, for example, that there are 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: IFS, European Commission; and Fund staff calculations.

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, the negative gap in 2002 is in line with general economic indicators.[33] Moreover, compared to the purely statistical methods, the production function approach tends to underscore with more force the importance of the ICT sector—while, more generally, shifting attention away from aggregate growth to the various factors of production.

D. Summary of Results and Concluding Remarks

38. Table 1A presents descriptive statistics for the various estimates of the output gap considered above (and reproduced in Figure 9).[34] Overall, there are significant differences across the various methods—visible, for instance, in the 2002 gap—and thus a main conclusion is that there is considerable uncertainty about the size of the output gap in Finland. With regard to the intuitive criteria, most measures are centered around zero, with the notable exception of the quadratic trend (QT), while the exponential trend (ET) and the Blanchard-Quah (BQ) approaches are also further away from zero than the others. Exceptionally large extremes would seem to cast doubt on the arithmetic filtering techniques (linear (LT), QT and ET approaches), while high standard deviations further lessen their attractiveness. The remaining measures—Hodrick-Prescott (HP), the frequency domain filter (FD), Beveridge-Nelson (BN), and production function (PF)—display standard deviations in the more appealing range of 2–4 percent. However, the associated estimates of the output gap in 2002 vary between -1.4 percent and 1.5 percent of potential. Given prospects of low and falling inflation and high unemployment, estimates suggesting actual output above potential seem counterintuitive.

[33] See Wagner (2001).

[34] The simple linear, quadratic, and exponential trend measures have been omitted.

Table 1A: Descriptive Statistics for the Output Gap Measures Considered (1980-2002)

	Gap measure									
	LT	QT	ET	HP(λ=100)	HP(λ=200)	HP(λ=20)	FD	BN	BQ	PF
Mean	0.1	3.0	0.6	0.2	0.2	0.1	0.2	0.1	0.4	[illegible]
Min	-11.2	-9.6	-11.5	-7.9	-8.7	-5.7	-3.5	-6.4	-1.7	[illegible]
Max	8.3	14.5	11.6	8.8	9.2	6.7	5.1	4.0	1.5	[illegible]
Standard deviation	5.5	8.2	7.7	4.0	4.4	2.9	2.3	2.2	0.7	[illegible]
Output gap in 2002	5.6	-3.2	-5.0	1.2	1.5	0.2	-1.4	0.5	0.7	[illegible]

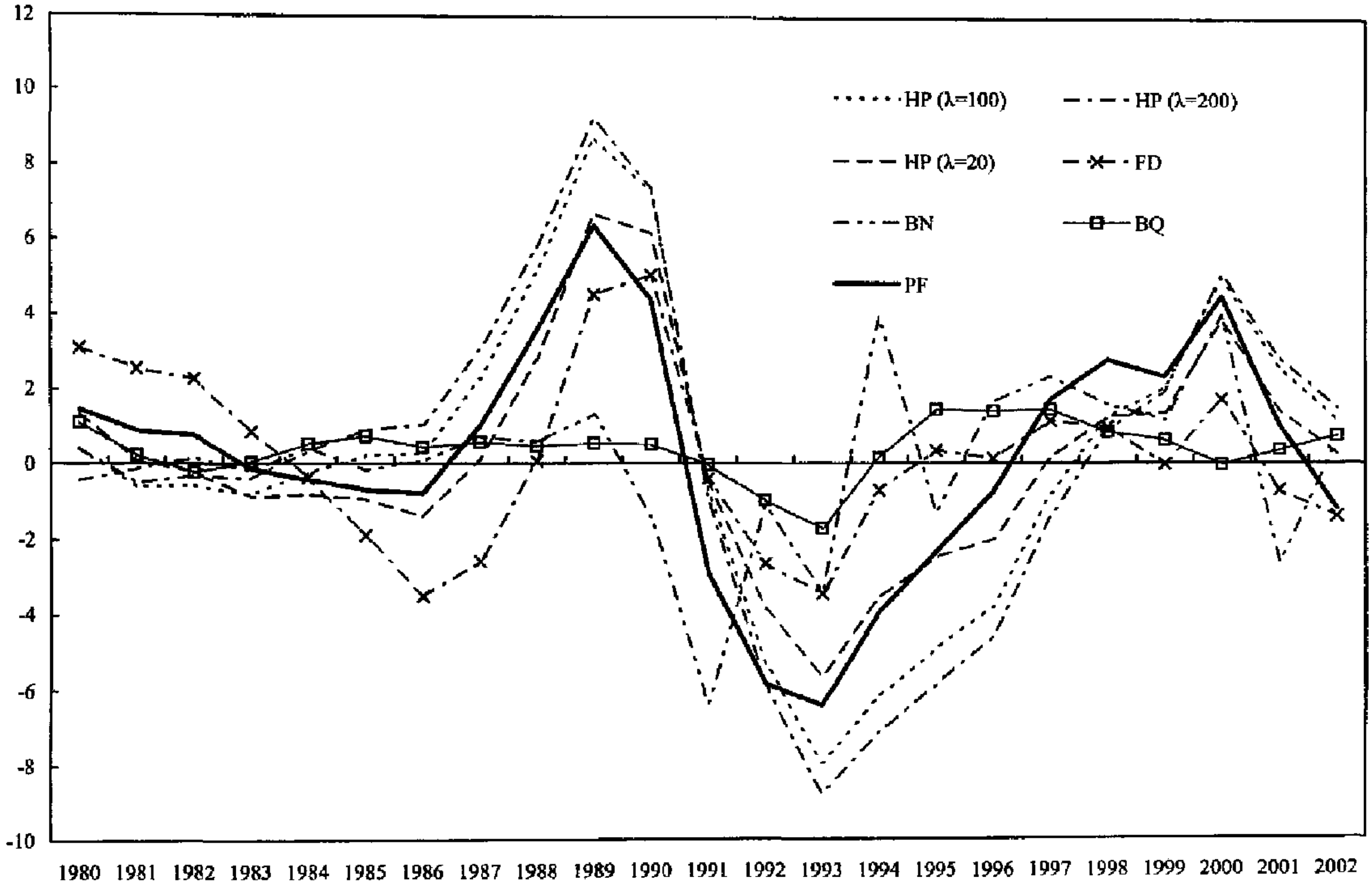
Table 1B: Correlation of Output Gap Measures 1/

	Gap measure									
	LT	QT	ET	HP(λ=100)	HP(λ=200)	HP(λ=20)	FD	BN	BQ	PF
LT	1	0.751	0.774	0.955	0.965	0.905	0.550	0.140	0.277	[illegible]
QT	0.653	1	0.999	0.886	0.883	0.890	0.746	-0.176	0.076	[illegible]
ET	0.674	1.000	1	0.903	0.900	0.906	0.751	-0.155	0.094	[illegible]
HP(λ=100)	0.946	0.650	0.671	1	0.999	0.987	0.743	0.109	0.302	[illegible]
HP(λ=200)	0.961	0.699	0.719	0.996	1	0.979	0.713	0.092	0.280	[illegible]
HP(λ=20)	0.868	0.560	0.583	0.965	0.940	1	0.840	0.163	0.372	[illegible]
FD	0.417	0.387	0.401	0.556	0.511	0.736	1	0.258	0.543	[illegible]
BN	0.187	-0.012	-0.003	0.170	0.158	0.202	0.192	1	0.380	[illegible]
BQ	0.273	0.089	0.100	0.287	0.269	0.338	0.332	0.356	1	[illegible]
PF	0.840	0.564	0.586	0.908	0.889	0.935	0.709	0.403	0.466	[illegible]

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

39. On balance, notwithstanding caveats, the following approaches seem the most promising. 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, further refinements left to 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).The frequency domain approach has broadly similar properties in terms of the intuitive criteria (with the (negative) output gap in 2002 estimated to be 0.2 percent of potential larger than under the PF approach) and the correlation between the two approaches is rather high (0.83) over the sample 1990–2002 (see Table 1B). Finally, the HP(λ = 20) filter using an extended forecast period—while indicating an essentially closed gap and, hence, contradicting to some extent the results from the two aforementioned filters—may have some merit as a rather pragmatic approach due to its ease of applicability. Moreover, its correlation with the PF and FD 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.

Figure 9. Finland: Comparing Measures of the Output Gap, 1980-2002 1/
(In percent of potential GDP)



Sources: IFS, WEO; and Fund staff calculations.

1/ The results from the Hodrick-Prescott filter are based on the extended data series until 2005.

The Baseline NAWRU Model

The filtering process yielding the NAWRU is based on the unobserved components approach, see Kuttner (1994), Denis, Mc Morrow, and Roeger (2002), and Planas and Rossi (2003). The starting point for the first component of the bivariate model is the definition:

$$U_t = C_t + \tilde{T}_t, \tag{A1}$$

which decomposes the observable unemployment rate U_t in a cyclical component, C_t, and a non-cyclical component, $\tilde{T}_t$. Additional exogenous regressors—$M \leq 3$—are assigned to the latter component, such that

$$\tilde{T}_t = T_t + \sum_{m=1}^{M} \alpha_m Z_{m,t}, \tag{A1'}$$

where T_t represents the underlying long-term trend, or NAWRU. Without additional exogenous regressors, i.e., $M = 0$, the non-cyclical component and the NAWRU-trend coincide.

The trend component is modeled according to its statistical properties, i.e. no economic information (e.g., on structural breaks) is included. The most general specification (see further below) is given by a random walk with drift, where the drift term μ_t is itself a random walk (and the trend T_t, hence, a second-order random walk):

$$\begin{aligned} T_t &= \mu_t + T_{t-1} + z_t, \qquad \textit{with} \\ \mu_t &= \mu_{t-1} + a_t \end{aligned} \tag{A2}$$

Both errors, z_t and a_t, are n.i.i.d; if $\mathrm{Var}(a_t) = 0$, the model collapses to a 1[st] order random walk with drift. On the other hand, the cyclical component in (1) is specified as an AR(N) process:

$$C_t = \sum_{n=1}^{N} \phi_n C_{t-n} + v_t \tag{A3}$$

where $N \leq 2$. To guarantee stationarity of the cyclical component, it must hold that $\sum \phi_n < 1$.

The second component of the generic model is given by

$$\Delta\pi_t^w = \mu \underbrace{\left[+\sum_{l=1}^{L} \rho_l X_{lt}\right]}_{a} \underbrace{\left[+\sum_{s=1}^{S} \theta_s \Delta\pi_{t-s}^w\right]}_{b} \underbrace{\left[+\gamma(1-L)^d U_{t-1}\right]}_{c} \underbrace{\left[+\sum_{r=0}^{R} \beta_r C_{t-r}\right]}_{d} + \underbrace{b_t}_{e}, \quad \textit{where} \quad b_t = \sum_{i=0}^{I} \varepsilon_{t-i} \tag{A4}$$

and L is the lag operator.

This Phillips curve relationship links the change in wage inflation to (i) exogenous determinants of wage inflation, X_t, such as (changes in) labor productivity or (changes in) the terms of trade, with $0 \leq L \leq 10$ in the empirical application; (ii) autoregressive terms of the wage inflation (with $0 \leq S \leq 2$); (iii) the d-th difference of the lagged observed unemployment rate U_t; (iv) the cyclical unemployment component C_t, (with $0 \leq R \leq 4$), and (v) an error term, which can have a MA(I) structure, $I \leq 3$.

In the estimation, the most generic model is used: with respect to (A1), the specification chosen is a bivariate autoregressive model, with the trend expressed as a second order random walk, hence, $\mathrm{Var}(a_t) \neq 0$ in (A2). Furthermore, an AR(2) specification is selected for the cyclical component in (A3), as indicated by preliminary tests (not reported here). Experiments with additional exogenous regressors in (A1), that is, $M > 0$, have resulted in a deterioration of the statistical fit.

Regarding (A4), the second difference of the lagged first series ($\Delta^2 U_{t-1}$), as well as the contemporary cyclical component of unemployment were included, that is, $d = 2$ and $R = 0$. The choice of a 2nd order RW specification in (A2) implies for (A4) that $d = 2$, that is, the lagged unemployment series regressor enters in second difference in order to obtain a stationary regressor. In (A4), no exogenous regressors were employed in the baseline case.[35]

Table A1 illustrates the impact of different assumptions on the ARMA structure of the Phillips curve equation (A4) on key results of the model, including the estimated coefficient β_0, the implied 2002 NAWRU, as well as selected test statistics. The significance of the estimator β_0—which multiplies the contemporaneous cyclical unemployment component—indicates whether changes in wage inflation respond to the general economic environment as represented by the cyclical unemployment component. It is expected to be negative and significant, reflecting the dampening effect of an adverse economic environment on the size of wage increases. The estimate of the NAWRU in 2002 (as opposed to the official unemployment rate of 9.1 percent) is reported to allow a plausibility check of the results. Whether the model gives a statistically acceptable description of the endogenous series' first two moments is checked by means of a Ljung-Box residual test statistic, with the null hypothesis being that the residuals are white noise.

[35] Maximization of the likelihood function was carried out by two algorithms. While the simulated annealing algorithm is slower than a Newton type algorithm, it is more likely to identify a global maximum. Since our experiments showed that local maxima posed a problem using the Newton-type algorithm, the simulated annealing algorithm was applied in the estimation.

Table A1: Descriptive Results for the NAWRU Model (Simulated Annealing Algorithm)

	Model Specification: ARMA(s,i)					
	1,1	1,2	1,3	2,1	2,2	2,3
beta0 1/	0.0215	0.0196	0.0194	0.0130	-0.0093	-0.0101
t-value	2.32	1.97	1.93	1.64	-2.13	-2.33
NAWRU 2002	9.07	9.00	9.00	8.94	8.31	8.33
Ljung-Box p-value residuals from (A1)	0.03	0.03	0.03	0.02	0.15	0.15
Ljung-Box p-value residuals from (A4)	0.19	0.80	0.92	0.91	0.21	0.27
-2*loglikelihood	111.78	107.25	107.20	106.62	101.78	101.33
R^2 (uncentered)	0.21	0.31	0.31	0.35	0.43	0.44

1/ beta0 represents the coefficient of the cyclical component of unemployment in the Phillips curve.

The table points to an unsatisfactory description of the data in the first four ARMA specifications (in particular with regard to the unemployment rate, see the first Ljung-Box test statistic). In addition, the positive value for β_0 is clearly counterintuitive. The ARMA (2,2) and (2,3) specifications instead yield a reasonable statistical description: normality assumptions on both equations are not significant at a 10 percent level. The estimated β_0 is significant for both models. Increasing the number of moving average terms in (A4) raises the t-value (in absolute terms). At 8.3 percent, the NAWRU derived for 2002 (in both specifications) is about 0.8 percentage points lower than observed unemployment. Using the ARMA(2,3) model, the final specification of (A4) is hence:

$$\Delta\pi_t^w = \mu + \rho_1 ddtot_t + \rho_2 dwprod_t + \sum_{s=1}^{2}\theta_s \Delta\pi_{t-s}^w + \gamma\Delta^2 U_{t-1} + \beta_0 C_t + u_t, \quad where \quad u_t = \sum_{i=0}^{3}\varepsilon_{t-i} \quad \text{(A5)}$$

Figure 6 in the main text illustrates the NAWRU implied by (A4), and (A1), (A2), and (A3), specified as:

$$\begin{aligned} U_t &= T_t + C_t \\ T_t &= \mu_t + T_{t-1} + z_t, \qquad with \quad \mu_t = \mu_{t-1} + a_t, \quad Var(a_t) \neq 0 \qquad \text{(A6)} \\ C_t &= \sum_{n=1}^{2}\phi_n C_{t-n} + v_t \end{aligned}$$

Additional Technical Restrictions of the Baseline Model

Additional restrictions imposed during the estimation process relate to parameters of the maximization technique (simulated annealing), and other parameters. In addition, if exogenous regressors were to be employed, boundaries on these variables in both equations (A1') and (A4) could be imposed as well.

Table B1. Additional Technical Assumptions and Restrictions

A. Assumptions on the "Simulated Annealing" Parameters

Initial temperature	1
Temperature redution factor	0.8
Tolerance	0.005
Number of cycles	15
Number of iterations	50
Maximum number of evaluations	1000000

B. Restrictions/Boundaries on "Other" Parameters

	lower bound	upper bound	estimated model
Equation (A1)			
Drift	-0.5	0.5	n.a. 1/
AR1	-1.96	1.96	1.32
AR2	-0.97	0.97	-0.82
Cycle MA1	-1.96	1.96	n.a. 1/
Cycle MA2	-0.97	0.97	n.a. 1/
Trend innovation variance	0.01	0.3	0.01001
Trend slope variance	0.75	0.95	0.75
Cycle innovation variance	0	0.406	0.31
Equation (A2)			
Intercept	-1	1	-0.004
Gamma - lag 1	-2	2	-0.003
Beta - lag 0	-2	2	-0.01
AR1	-1.96	1.96	-0.14
AR2	-0.97	0.97	-0.85
MA1	-0.97	0.97	0.38
MA2	-0.97	0.97	0.9
MA3	-0.97	0.97	-0.18
Innovation variance	0	0.000474	0.0003

1/ Restrictions not applicable due to the choice of trend specification (second order random walk).

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*, Vol. 24, pp.1–35. (Washington: International Monetary Fund).

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

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

Beyer, A., and R. E. A. Farmer, 2002, "Natural Rate Doubts," ECB Working Paper 121 (Frankfurt: European Central Bank).

Blanchard, O., and D. Quah, 1989, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, Vol. 79, pp. 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*, Vol. 109, pp. 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*, Vol. 55, pp. 313–28.

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).

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

De Masi, P. R., 1997, "IMF Estimates of Potential Output: Theory and Practice," IMF Working Paper No. 97/177 (Washington: International Monetary Fund).

Denis, C., K. Mc Morrow, and W. Roeger, 2002, "Production Function Approach to Calculating Potential Growth and Output Gaps—Estimates for the EU Member States and the U.S.," 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 No. 03/128 (Washington: International Monetary Fund).

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

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

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

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 and Economic Statistics*, Vol.12, pp. 361–8.

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).

Ross, K., and A. Ubide, 2001, "Mind the Gap: What is the Best Measure of Slack in the Euro Area?" IMF Working Paper No. 01/203 (Washington: International Monetary Fund).

Scacciavillani, F., and P. Swagel, 1999, "Measures of Potential Output: An Application to Israel," IMF Working Paper No. 99/96 (Washington: 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).

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