MEMORANDUM 1/2023

The Finnish Economic Policy Council

THE FINNISH ECONOMIC POLICY COUNCIL

Revisions of Output Gap and Potential Output in Finland

Adam Rybarczyk

MEMORANDUM 1/2023

Adam Rybarczyk

The Finnish Economic Policy Council
Address:
VATT Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Helsinki, November 2023

Contents

1 Introduction

One of the goals of a government's fiscal policy is to act countercyclically in the presence of economic shocks. This is achieved through both automatic stabilizers as well as one-off policy decisions. However, to perform appropriate fiscal policy, estimates of the cyclical stance of the economy, as well as its output potential, are needed. One of the common indicators of the cyclical stance is the output gap, defined as the relative difference between potential- and realized output.

The output gap and potential output also play an important role in enforcing European Union's Stability and Growth Pact. The medium term objectives therein are defined in terms of structural balances. Estimating these requires knowledge of the output gap, and thus potential output. In the case that a country deviates from their objective, Commission can punish the member state with financial sanctions. (European Commission, 2018).

In practice, the estimation of potential output is challenging. The estimates should be available in real time to allow for a timely response of fiscal policy. However, a large body of literature has shown that the real-time estimates are not too reliable (Chen & Górnicka, 2020; Coibion et al., 2017; Grigoli et al., 2015; Kangur et al., 2019; Orphanides & Norden, 2002). The revisions to the initial estimates are large, and, furthermore, the differences between different methods can be substantial, to the point of changing the sign of the cyclical stance of the economy. This presents a challenge for policy making, as policy might have to be based on flawed real-time estimates (Orphanides & Norden, 2002).

The main problem when trying to compare different methods of estimating the potential output is that neither potential output nor the output gap are observable. Thus different models have to be compared against each other, and model selection has to be made based on other criteria than performance with respect to observables.

In Finland, the Ministry of Finance uses the potential output measuring methodology created by the European Commission (EC) (Valtiovarainministeriö, 2022). Thus, understanding the properties of this methodology is of interest when considering the domestic fiscal policy. In this report, we attempt to gain insight into the properties of the estimates by comparing the EC results against the estimates of IMF and OECD, as well as comparing the estimates against those for Germany and Sweden.

Our approach concentrates on investigating the properties of the estimates across data vintages, that is, observations of the same data as reported in different points in time. This allows us to look into the properties of the revisions in potential output releases. We look at original series on potential output, as well as construct our own estimates using European Commission's methodology, thus removing differences arising from different model specifications throughout time. We compare estimates for Finland across these institutions, as well as estimates by the EC for Finland, Germany, and Sweden

Consistently with the previous literature, we find that the revisions of the output gap estimates are large. Comparing the estimates for Finland across the three considered institutions shows that the EC's estimates are revised the least. Within the EC methodology, we find that revisions as well as forecast errors for Finland are larger than those for Germany and Sweden. We also show that small improvements of the forecast performance of the EC methodology can be achieved by adjustments in the parametrisation of the model.

The next section presents a literature review on articles concerning Finland. Section 3 presents the vintage data releases for potential output and the output gap by EC, IMF, and OECD. Section 4 presents further results we have obtained by applying the EC methodology to vintage releases of the data. Section 5 concludes.

2 Literature Review

As the body of literature investigating the properties of output gap estimates is rather large, this paper focuses on analyses that have specifically investigated output gap estimates for Finland. The references mentioned in the introduction are a good starting point for general literature on output gap estimation.

Billmeier (2004) is one of the earliest authors to investigate output gap estimates specifically for Finland. His work covers both statistical and structural methods for estimating the potential output. He finds large differences in the estimates of the output gap generated by the different methods. For example, in 2002 most of the values range between -1.5% and 1.2%, although there are a few outliers. Haavio (2008) compares the HP-filter, production function approach, and Bank of Finland's AINO dynamic stochastic general equilibrium (DSGE) model for estimating potential output in Finland. The output gaps produced by the DSGE model turn out more moderate than either the HP-filtered or production function approach output gaps.

Melolinna (2010) compared mostly unobserved component models. They find that priors on the parameter values can affect the estimates of the output gap, with the values of output gap estimates for Finland ranging within a couple percentage points of each other across the specifications.

IMF (2014) compared three methods for calculating the potential output in the Finnish context; HP- and multivariate filters, and Cobb-Douglas production function methodology. They find that HP-filter and production function approach estimate the growth of the potential GDP to be positive in the years following the financial crisis, while the multivariate filter estimates it to be negative. Additionally, their calculations also show that the smoothing parameter of the HP-filter is an important determinant of the results, affecting growth rates, potential output, and the speed at which output gaps are closed.

Huovari et al. (2017) have investigated some properties of the European Commission's Commonly Agreed methodology (EUCAM) for output gap estimation with Finnish data. They use the latest revision of the data, and compared the output gap estimates obtained with different model assumptions to the historic estimates from the European Commission. They find that results are quite sensitive to the assumptions about parameters, especially in the estimation of the trends of hours worked and participation rates. On the other hand the production function parameter α used in the Cobb-Douglas production function did not significantly affect the estimates.

The work of Jysmä et al. (2019) investigates the effects of substituting the constant elasticity of substitution (CES) production function instead of the Cobb-Douglas one used in EUCAM. An advantage of this approach is the ability to distinguish between capital- and labour-augmenting technology. The authors find that the efficiency of capital use decreases in a downturn, while that of labour stays somewhat constant, perhaps even increases. Their results suggest that cyclical movements could have affected potential output more than European Commission's estimates suggest.

Orjasniemi and Kuusela (2021) have investigated the properties of multivariate filter for output gap estimation using quarterly data from Finland. The revisions of the output gap estimates obtained using their methodology are less than those obtained using the HP-filter, especially in the build-up towards the financial crisis. The variables used in the estimation were GDP, unemployment rate, inflation, and inflation expectations.

3 Estimates and Methodologies of Major Institutions

EC, IMF, and OECD all provide access to the vintage time series of their potential output and output gap estimates, as well as other variables. These are usually released twice a year. EC's AMECO database contains the spring and autumn releases of Commission's economic forecasts, with 2013–2017 additionally containing a winter forecast. IMF publishes its World Economic Outlook twice a year in spring and autumn. OECD releases its Economic Outlook in summer and winter. One should note that the winter releases of Commission were published at the beginning of the year, while those of OECD at the end of the year.

3.1 Vintage Releases

Figures 3.1 and 3.2 present the level of potential output and its growth rate for Finland respectively, as estimated by EC, IMF, and OECD. Each line in a graph represents one vintage of the data. The earliest vintage releases are autumn 2002 for EC, spring 2000 for IMF, and winter 1996 for OECD. The levels are normalized so that different base years and data definitions can be compared. Latest available vintage of the data is displayed in red. From figure 3.1 it can be seen that the real-time estimates of the level of potential output were overly optimistic for multiple years following the financial crisis, when compared to the latest revision of the data. Figure 3.2 shows that Commission's methodology produces lower growth rates during both the financial crisis and the downturn of the 1990s. This is reflected in higher growth rates outside of these periods.

The estimates of potential output growth rates by IMF are considerably less smooth than the estimates of other two institutions. There are large jumps between vintages in the 90s. This is probably related to IMF's country desk officers having the freedom to exercise substantial judgment in forming the estimates. This will be elaborated further in section 3.2.

An interesting feature of the OECD data is that there was a sharp revision in the level

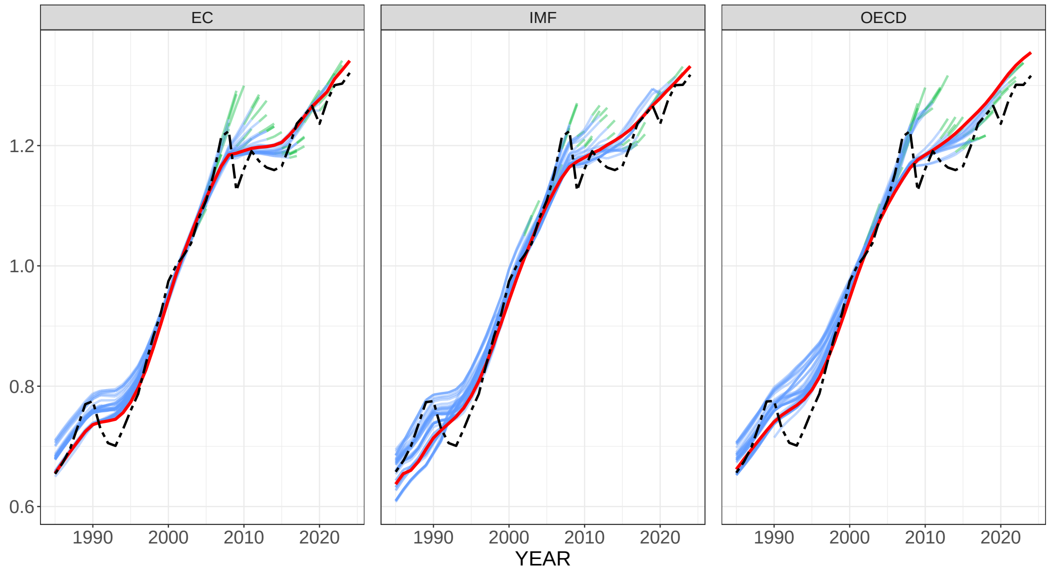

Figure 3.1: *EC, IMF, and OECD estimates of potential GDP (solid blue lines), latest vintage of potential GDP (thick red line), latest vintage of real GDP (dashed line), and forecasts of potential GDP (green) for Finland. All values are normalized such that value in 2001 is equal to 1. Potential GDP series is shifted by the value of output gap in 2001.*

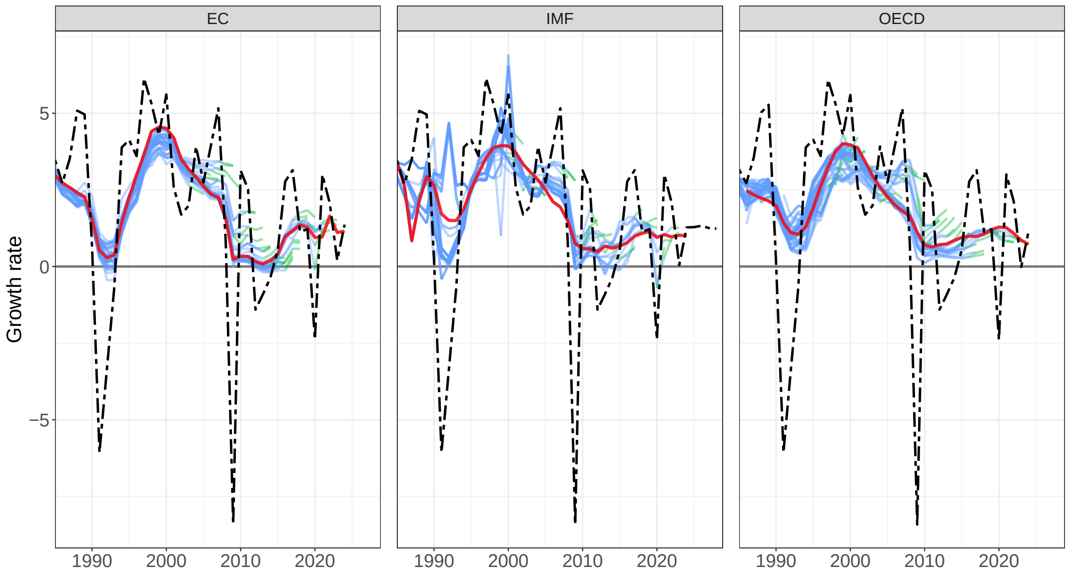

Figure 3.2: *EC, IMF, and OECD estimates of potential GDP growth rate (solid blue lines), latest vintage of potential GDP growth (thick red line), latest vintage of real GDP growth (dashed line), and forecasts of potential GDP (green) for Finland.*

of potential output starting from the 2012 vintage, which resulted in a higher output gap for years 2007–2011. For comparison, the revisions of EC and IMF estimates were more gradual.

Figure 3.3 presents the vintage output gap estimates for Finland. EC has a slightly more pessimistic view of the level of potential output in Finland during the financial crisis. This implies an output gap closer to zero in the following years. The difference is quite clear when compared to OECD estimates, in which the real output has not reached the potential since 2011, except for a couple initial estimates that were revised later. EC and IMF have reported a roughly zero output gap during 2017–2019.

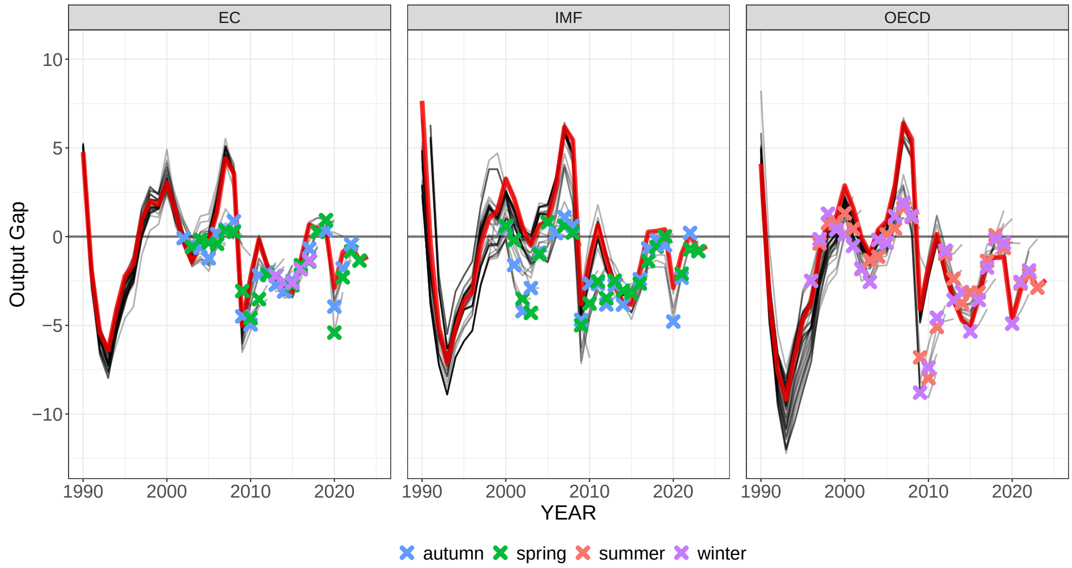

Figure 3.3: *EC, IMF, and OECD estimates of the output gap (solid lines), latest vintage of output gap (thick red line), for Finland. Crosses represent real-time estimates. Colour of the crosses represents the vintage season of the release.*

It can be seen that the errors in the real-time estimates of the output gap were especially severe during the financial crisis, when all institutions underestimated the size of the upswing of the business cycle. OECD also made a large negative error when estimating the size of the recession in 2009-2011, while IMF made large real-time errors in the early 2000s. There is also a difference in the growth rates of the real-time estimates leading up to the financial crisis. OECD estimates appear to have caught some of the upward movement of the output gap, whereas the estimates of the Commission are all hovering around zero, with a very modest upward trend. The IMF estimates also appear to have an upward trend, however the starting point has a larger error than the estimates of EC and OECD.

The large errors of the real-time estimates in the build-up towards the financial crisis are not a universal rule in the data. Figure 3.4 presents Commissions output gap estimates for Finland and Germany. It can be seen that the real-time estimates in 2008 worked significantly better in the case of Germany, with the autumn forecast being essentially spot-on with the latest available data vintage.

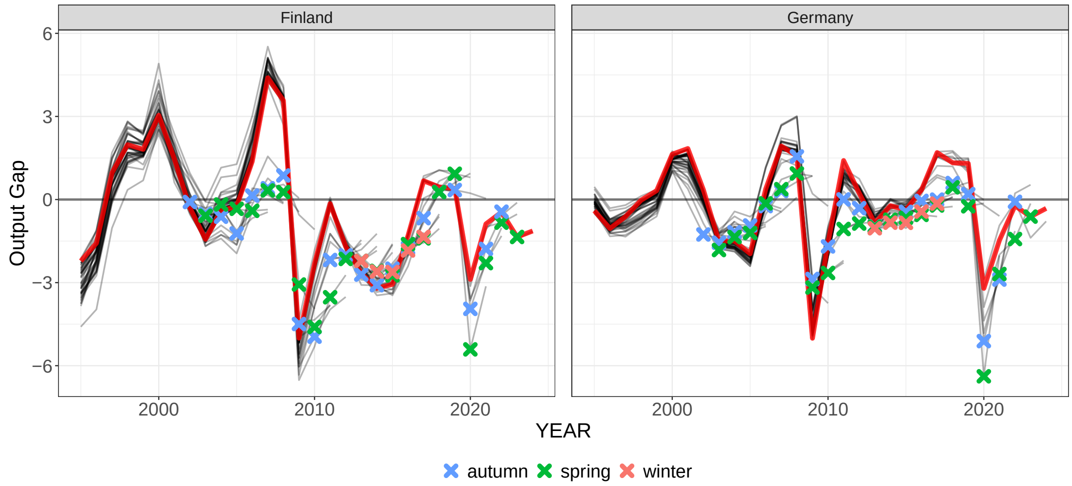

Figure 3.4: *EC estimates of the output gap (solid lines), latest vintage of output gap (thick red line), for Finland and Germany. Crosses represent real-time estimates. Colour of the crosses represents the vintage season of the release.*

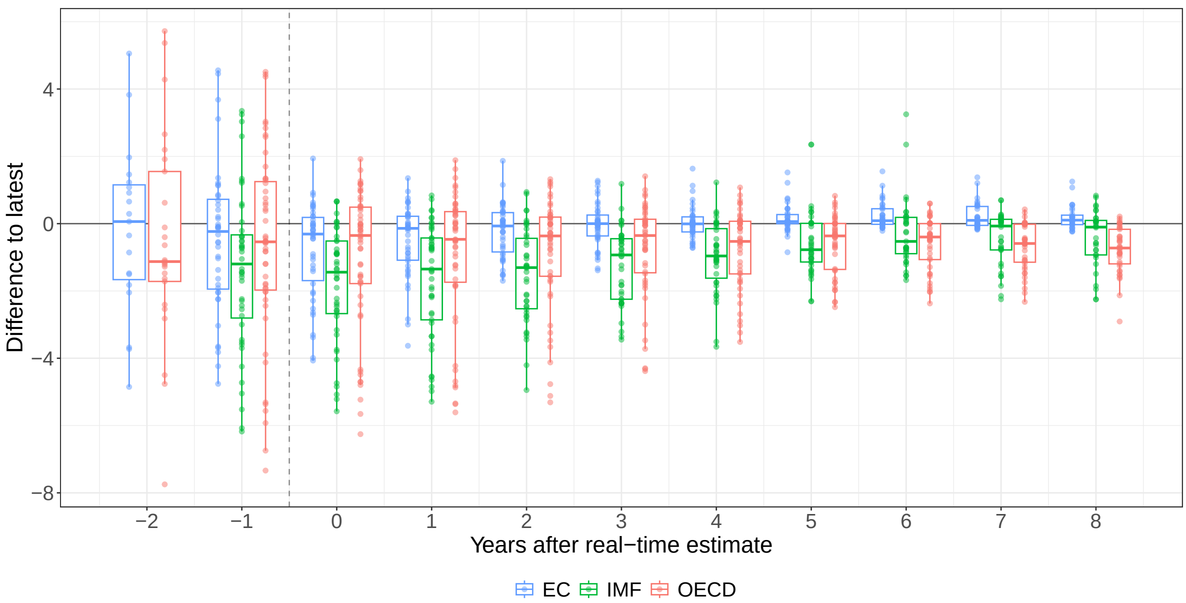

Figure 3.5: *Gaps to the latest vintage of output gap estimate by years after the real-time estimate for Finland. Dots are the observed values. Boxes present the 25th percentile, median, and 75th percentile. Whiskers extend to the furthest value no more than 1.5 times the interquartile range from the boxes. Positive values indicate that the estimate is above the latest estimate.*

	Finland			Germany			Sweden		
	EC	IMF	OECD	EC	IMF	OECD	EC	IMF	OECD
Mean	1.04	2.06	1.51	0.31	1.00	0.98	0.57	1.27	1.57
Mean Absolute	1.44	2.30	2.10	0.80	1.13	1.07	1.07	1.70	1.75
Median	0.56	2.09	1.17	0.42	1.03	1.12	0.37	1.31	1.49
SD	1.68	1.91	2.30	0.94	0.92	0.91	1.23	1.79	1.56
Sample Size	28	26	28	28	26	28	28	26	28

Table 3.1: *Summary statistics for output gap revisions seven years after the initial estimate is made. Only vintages starting in 2002 are included.*

Figure 3.5 furthermore investigates the convergence of output gap estimates for Finland. It can be noted that EC estimates converge towards the final estimate faster than the estimates of either IMF or OECD. This is seen as the blue boxes shrink towards zero faster over time, as compared with red and green ones. Furthermore the median estimate of EC is closer to the final estimate than both IMF and OECD estimates starting from the earliest forecasts all the way to six years after the initial estimate.

Table 3.1 presents summary statistics for the revisions of output gap estimates by institution for Finland, Germany, and Sweden. The revision is defined as the difference between the value of the output gap estimated seven years later after the initial estimate, and the initial estimate. Thus a positive value indicates an upwards revision of the output gap. Seven years was chosen as the horizon of interest as by that time the output gap estimates have somewhat converged. The same horizon was used by Grigoli et al. (2015).

The values show that the Commission's estimates are revised the least for all three countries in consideration, as measured by either mean revisions, mean absolute revisions, or median revisions. Furthermore the standard deviations of EC's revisions are lowest for Finland and Sweden, but OECD's and IMF's estimates have a lower value for Germany. There is no clear pattern between the results of IMF and OECD.

One of the goals of the Commission's methodology is to "reduce the degree of cyclicality of the potential growth estimates to an absolute minimum in order to avoid the mistakes of the past." This has been mentioned in all the major releases concerning the methodology (D'Auria et al., 2010; Denis et al., 2002, 2006; Havik et al., 2014). This goal was set out due to overly optimistic forecasts of budgetary developments made in the past.

In figure 3.6 we have plotted the revisions of the output gap estimates against the real-time estimates of real output growth, and least squares estimates of the relationships. In this graph the goal of reducing cyclicality would correspond to having the slope coefficient of the least squares estimate close to zero. The motivation of this exercise is that if the revisions of the output gap estimates display cyclicality, then so will the structural budget balances. OLS estimates are reported in table 3.2. From the results we can see that in this interpretation, all three institutions have some degree of cyclicality in their estimates. Furthermore OECD's estimates are the least dependant on real GDP

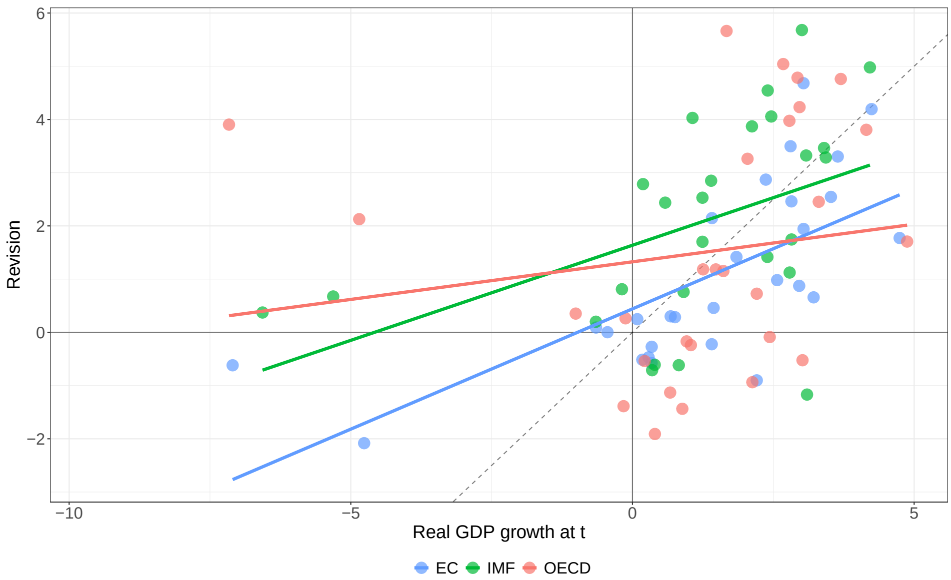

Figure 3.6: *Revision of the output gap estimate seven years after the initial estimate is made and real GDP growth as estimated in real time for Finland. Positive values of revision indicate that output gap was revised upwards. Lines are the least squares estimates. Dashed line is the 45 degree angle. Vintages are limited to start from 2002.*

Table 3.2

	Sample:					
	Full			Outliers removed		
	EC	IMF	OECD	EC	IMF	OECD
Real GDP Growth	0.452***	0.357***	0.141	0.751***	0.746**	0.969***
	(0.120)	(0.138)	(0.259)	(0.157)	(0.321)	(0.236)
Constant	0.441	1.637***	1.325	−0.188	0.863	−0.402
	(0.318)	(0.521)	(0.869)	(0.231)	(0.700)	(0.455)
Observations	28	26	28	26	24	26
R^2	0.464	0.212	0.024	0.508	0.262	0.354
Adjusted R^2	0.444	0.179	−0.014	0.487	0.228	0.327

Table 3.3: *OLS regression results of output gap revision seven years after the initial estimate is made against real-time estimate of real GDP growth. Heteroscedasticity and autocorrelation robust standard errors in parenthesis. Only vintages starting in 2002 are included. Outliers refer to observations with GDP growth of less than -3%.* $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

growth, while in fact EC estimates are the most dependant. However, if we remove the outlier values of low real GDP growth, the cyclicality between institutions becomes similar, with EC and IMF estimates having roughly equal cyclicality. It should also be pointed out that IMF results are based on two fewer observations, due to missing data.

3.2 Overview of EC, IMF, and OECD methodologies

The methodologies that the institutions use are somewhat similar between EC and OECD, while IMF is more of an outlier. As noted in De Masi (1997) for IMF, "In all cases, estimates of potential output incorporate a substantial amount of judgement and country-specific expertise of desk officers. Staff estimates of potential have usually been presented and discussed with the national authorities." Furthermore, survey results of De Resende (2014) show that out of IMF desk economists in the European Department, 56.3% use structural methods, 50.0% statistical methods, 78.1% use judgement, and 9.4% use other methods.[1] A country report on Finland (IMF, 2014) compared the HP-filter, production function approach, and the multivariate filter. The use of judgement and expertise appears to be reflected in figure 3.5; the IMF estimates never really converge as much as EC and OECD values.

Both EC's (Havik et al., 2014) and OECD's (Chalaux & Guillemette, 2019) methods for estimating potential output are Cobb-Douglas production function based approaches, thus both essentially concern estimating the potential level of TFP and labour input. The values of parameter α, corresponding to the output elasticity of labour, are similar, with $\alpha = 0.65$ in EUCAM and $\alpha = 0.67$ in OECD's methodology. These are constant and equal across all countries.

The methodologies differ slightly in the definition of potential labour input. EC's methodology uses potential total hours worked in the economy, while OECD's uses potential employment. Both methodologies use the number of 15–74-years-olds as the working age population. They also assume that the inputs of potential labour include working age population, non-accelerating wage rate of unemployment (NAWRU), and participation rate. EC's methodology additionally uses hours worked per emplyee as a factor, while OECD's methodology does not. However, the latter does include an adjustment factor for the total employment, which aims to correct for inconsistencies between national accounts and labour force surveys.

OECD applies an HP-filter to the working age population to extract the trend, whereas EC's methodology does not do this. NAWRU is estimated with a similar method in both methodologies, as both use a Kalman filter augmented with either forward- or backward looking Phillips curve in the estimation. Trend participation rate is obtained via an HP-filter in EUCAM, while OECD uses an approach in which the raw participation rate is adjusted using estimated impulse responses of the participation rate to the gap in unemployment. The trend of hours worked per person employed in EUCAM is estimated using an HP-filter.

For trend productivity, both institutions augment the raw productivity series with

[1] More than one answer was allowed.

additional variables that are assumed to be correlated with the productivity cycle. Both methodologies include survey-based measures of capacity utilisation, with EUCAM augmenting the series with business survey data for construction and services. OECD's methodology additionally uses investment-to-GDP ratio, current account, and commodity price indicators. These are selected on a country basis. In the case of Finland, the used indicators are capacity utilisation and the investment ratio.

In EUCAM capital stock is used as-is. This is justified by the fact that "the maximum potential output contribution of capital is given by the full utilisation of the existing capital stock in an economy" (Havik et al., 2014). This is why the series is also not detrended; furthermore, it is already rather stable as investment contributes a small fraction of capital stock each year. OECD uses a measure of productive capital stock, which excludes housing. This series is also used without further transformations.

3.3 Details of EUCAM

A detailed description of the current EUCAM methodology is provided in Havik et al. (2014), with a briefer explanation together with a guide on the corresponding software provided by Blondeau et al. (2021). Earlier papers describing the methodology in detail are D'Auria et al. (2010) and Denis et al. (2002, 2006).

The EUCAM methodology is a Cobb-Douglas (CD) production function (PF) approach, that is, output Y at time t is assumed to be produced according to

$$\begin{aligned} Y_t &= (U_K \cdot E_K \cdot K_t)^{1-\alpha} \cdot (U_L \cdot E_L \cdot L_t)^{\alpha} \\ &= (U_K \cdot E_K)^{1-\alpha}(U_L \cdot E_L)^{\alpha} K_t^{1-\alpha} L_t^{\alpha} \\ &\equiv A_t K_t^{1-\alpha} L_t^{\alpha}. \end{aligned} \tag{3.1}$$

K_t is the capital stock, L_t is the labour input, and the parameter α is given the value 0.65. Furthermore, U refers to the degree of capacity utilisation and E refers to the efficiency of capital and labour augmenting technology. These are collected into A_t, the total factor productivity (TFP). The link between capacity utilisation and TFP will later be exploited when recovering the potential value of TFP.

The production function approach to estimating potential output has been in use at the Commission since 2002. Prior to that the Hodrick-Prescott (HP) filter was used to detrend GDP. However, as mentioned in section 3.1, this method was not satisfactory. The goal of moving towards the PF approach was to reduce the cyclicality and biases of the estimates, as well as keeping the methodology transparent and ensuring an equal treatment of all member states (Havik et al., 2014).

Estimation of potential in EUCAM boils down to estimating potential TFP and labour inputs. Denoting potential variables with a bar over the corresponding letter, the potential output is then

$$\bar{Y}_t = \bar{A}_t \cdot K_t^{1-\alpha} \cdot \bar{L}_t^{\alpha}. \tag{3.2}$$

Note that the capital stock enters the equation without further transformations, as the size of the capital stock directly determines the possible contribution of capital to output. Additionally the time series of capital stock is quite smooth without any further transformations.

It should be mentioned that the fit of the CD-PF to the Finnish economy has been questioned by Jysmä et al. (2019) and Ripatti and Vilmunen (2001), who both find the constant elasticity of substitution (CES) production function superior to CD-PF. The usage of CES-PF would additionally allow to disentangle TFP into the labour and capital augmenting parts. Jysmä et al. showed that the capital-augmenting technology is more pro-cyclical than labour-augmenting one.

3.3.1 Potential Labour Input

The potential labour input is obtained as a product of the working age population, the trend of the participation rate, the non-accelerating wage rate of unemployment (NAWRU), and the trend of hours worked per person employed. Working age population is defined as the number of 15–74 -year-olds. The trends of participation rate and hours worked per person employed are obtained by first producing an extension from $T+3$ to $T+8$ using a univariate autoregressive model. After this, the time series are smoothed using an HP-filter. The goal of the autoregressive extension is to reduce the end-point bias present in HP-filter.[2]

The method of estimating NAWRU is presented in Hristov et al. (2017). Specifically it is assumed that the unemployment rate U_t is decomposed into NAWRU n_t and a cycle c_t as $U_t = n_t + c_t$. Furthermore the movement of NAWRU and the cycle are assumed to follow a stochastic process as follows,

$$\begin{aligned} n_t - n_{t-1} &= a_{nt} + \eta_{t-1} \\ \eta_t - \eta_{t-1} &= a_{\eta t} \\ c_t &= \phi_{c,1} c_{t-1} + \phi_{c,2} c_{t-2} + a_{ct}, \end{aligned} \tag{3.3}$$

where $a_{nt}, a_{\eta t}$, and a_{ct} are normally distributed white noise. The cyclical component is then related to either a backward- or forward looking Phillips curve. The backward-looking Phillips curve is currently in use for Austria, Belgium, Germany, Italy, Luxembourg, Malta, and the Netherlands, and is specified as

$$\Delta \pi_t = \mu_\pi + \beta_0 c_t + \beta_1 c_{t-1} + \gamma' z_t + a_{wt}. \tag{3.4}$$

π_t is the wage inflation, and z_t is an exogenous vector of terms-of-trade, labour productivity, and changes in wage share. γ is a country-specific vector of coefficients, and a_{wt} is normally distributed white noise.

The rest of the EU countries use a forward-looking Phillips curve,

$$\Delta rulc_t = \phi_r \Delta rulc_{t-1} + \beta_0 c_t + \beta_1 c_{t-1} + a_{wt}, \tag{3.5}$$

where $rulc_t$ is an indicator of real unit labour cost.

[2] At the end of the sample the HP-filter becomes one-sided due to mossing future observations. This leads to large revisions in the estimated trend once the new data becomes available.

For the countries using the forward-looking Phillips curve, including Finland, an adjustment of NAWRU estimates is allowed after the initial estimation. This is done by first calculating the mean difference between NAWRU obtained via backward-looking and forward-looking specifications. If the difference is positive, then this difference is subtracted from NAWRU estimates in each year. In the case of Finland, this procedure yields a -0.72 percentage points revision in the estimates of NAWRU.

The mean-adjustment has been criticized by Huovari et al. (2017). They point out that there is no theoretical argument for the practice, and, furthermore, that it is not symmetric. There is no reason for the differences to remain constant. For Finland, this adjustment results in roughly 0.5% higher potential GDP, and 0.3 percentage point higher cyclically-adjusted budget balance.

3.3.2 Potential TFP

TFP is calculated as a Solow residual,

$$sr_t = \ln Y_t - \alpha \ln L_t - (1-\alpha) \ln K_t \tag{3.6}$$

This residual is decomposed into a potential and cyclical part, $sr_t = p_t + c_t$. The dynamics of the model are such that

$$\begin{aligned} \Delta p_t &= \mu_p + \eta_{t-1} + a_{pt} \\ \eta_t &= \rho \eta_{t-1} + a_{\eta t} \\ c_t &= 2A \cos(2\pi/\tau) c_{t-1} - A^2 c_{t-2} + a_{ct} \end{aligned} \tag{3.7}$$

As noted in the paragraph following equation 3.1, there is a link between TFP and capacity utilisation. Thus in the model the cycle of productivity is connected to an indicator of capacity utilisation, CUBS, via the equation

$$\begin{aligned} cubs_t &= \mu_{cu} + \beta_{cu} c_t + e_t \\ e_t &= \phi_{cu} e_{t-1} + a_{cu,t}. \end{aligned} \tag{3.8}$$

Shocks $a_{pt}, a_{\eta t}, a_{ct}$, and $a_{cu,t}$ are independent normally distributed white noise, and $cubs_t$ is the CUBS indicator at time t. CUBS is a composite indicator of capacity utilisation and business survey data. It is built from survey data on capacity utilisation in industry, as well as business sentiment indicators for services and construction sectors, as no capacity utilisation data are available for them. This indicator serves as a proxy of the true level of capacity utilisation in the economy.

Both the NAWRU and potential TFP are estimated using the Kalman filter. Maximum likelihood is used for NAWRU, while Bayesian approach is used for TFP.

3.3.3 Medium-term Forecasts

The input time series contain forecasts made by Directorate General for Economic and Financial Affairs (DG ECFIN) until $T+2$. This way the EUCAM methodology produces a short-term forecast using the aforementioned process. To produce the medium-term forecast, the potential output and capital stock are modelled jointly using the following system of equations.

$$\begin{aligned}\bar{Y}_t &= \exp(srk_t)\bar{L}_t^{\alpha}K_t^{1-\alpha} \\ K_t &= I_t + (1-\delta_t)K_{t-1} \\ I_t &= (\bar{IY}_t/100)\cdot\bar{Y}_t \\ Y_t &= \bar{Y}_t(1+\hat{Y}_t/100),\end{aligned} \tag{3.9}$$

where srk_t it the trend of log TFP, I_t is real gross fixed capital formation, $\bar{IY}_t$ is the investment to potential output ratio, and δ_t is the depreciation rate.

After performing this forecast, the output gap is assumed to close between $T+3$ and $T+5$ as

$$\begin{aligned}\hat{Y}_{t+3} &= \frac{2}{3}\hat{Y}_{t+2} \\ \hat{Y}_{t+4} &= \frac{1}{3}\hat{Y}_{t+2} \\ \hat{Y}_{t+5} &= 0.\end{aligned} \tag{3.10}$$

Thus regardless of the input series, output gap is forecast to close in five years. In practice this assumption could be criticized. For example, the average output gap in Finland is roughly -1% in the latest vintage releases across all observation years. Figure 3.7 presents the average output gap for a shorter sample period, such that all included countries have the same observation years in the sample. Countries were included based on data availability.

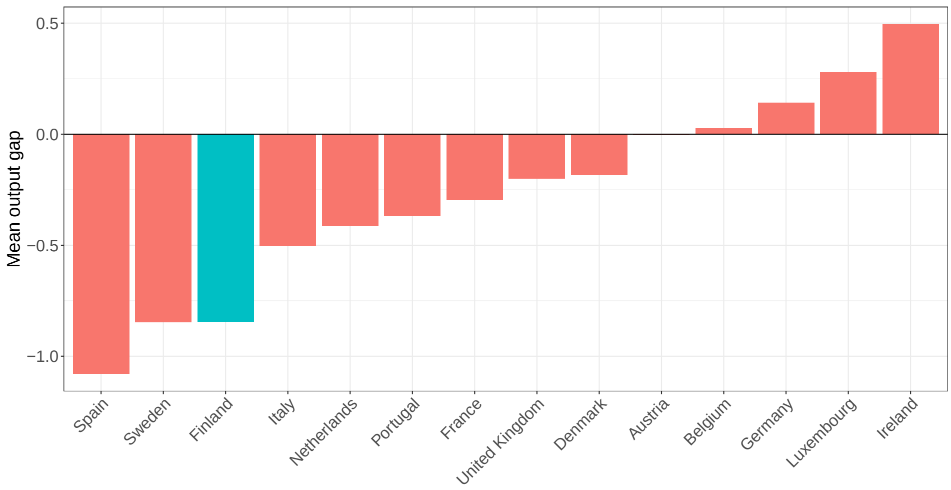

Figure 3.7: *Mean output gaps over the period 1991-2018, calculated using the 2023 spring vintage release of EC's economic forecast.*

4 Our Results

To produce the following results we have used the EUCAM software which is freely available from the Output Gaps Working Group (OGWG) library of CIRCABC[3] website, which is European Commission's hub for sharing information across users. The version of the software in use is the 2023 spring revision. The results we have obtained using this software version and parametrisation will be called ”EPC baseline” results, while the original releases by EC will be called ”original” results.

Our approach allows us to investigate how the potential output estimates behave across data vintages while keeping the methodology constant. Thus we can eliminate differences in potential output estimates arising from methodological changes across the vintage releases.

For the vintages starting in 2011, the data is collected from the AMECO archive releases. The previous data vintages are obtained from the releases of the OGWG at CIRCABC. This allows us to investigate vintages from spring 2005 to spring 2023. Furthermore, data on consumer price inflation was downloaded from OECD, as it was not available in all vintage releases of EC.

We use the number of 15–64 year olds as the working age population. This is not consistent with the current EC methodology, but it allows us to use the original Eurostat population projections as well as AMECO working age population values in our estimation.

For the new EU member states, data on capital stock is not available prior to 2012 vintage releases. Instead we calculate the capital stock using the perpetual inventory method with 5% yearly depreciation rate and a starting point for capital defined as twice the level of GDP in 1995, except for the Czech Republic, whose starting point is thrice the level of GDP. This approach is consistent with the historical methodology used to calculate the capital stock for EUCAM estimations for these countries, but it tends to produce slightly larger capital stock growth rates than the current data releases do.

4.1 Baseline Results

Figure 4.1 presents output gap estimates for Finland from the original AMECO releases as well as the ones calculated by us, starting from 1990. As can be seen, the general dynamics of the estimates are quite similar. Our estimates tend to be lower than the original estimates. The largest discrepancies occur in 2009–2011, with our real-time estimates being much lower than the real-time releases of the European Commission. Interestingly this discrepancy is similar in nature to the one between between EC and OECD estimates in figure 3.3. As the real GDP series are the same across the specifications, graph 4.1 also implicitly presents the potential GDP estimates.

Figure 4.2 shows the baseline estimation level of potential output, as well as the latest estimate of real output, for Finland, Germany, and Sweden. As can be seen, the forecast errors in potential GDP before the financial crisis were especially large for Finland. In

[3]Communication and Information Resource Centre for Administrations, Businesses and Citizens

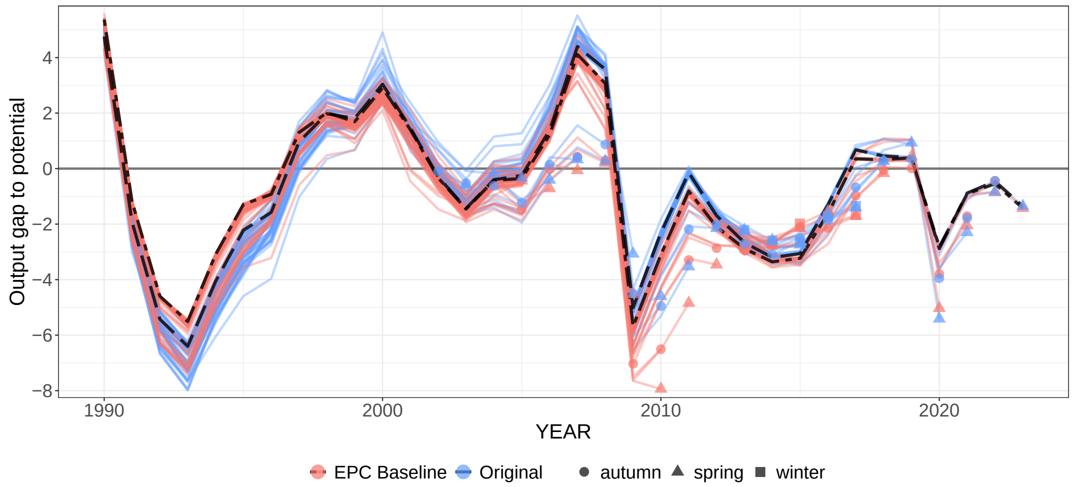

Figure 4.1: *Output gaps for Finland. Blue lines correspond to original releases from AMECO database. Red lines are output gaps calculated by us using the 2023 spring version of EUCAM software. Disks, triangles, and squares correspond to the real-time estimates.*

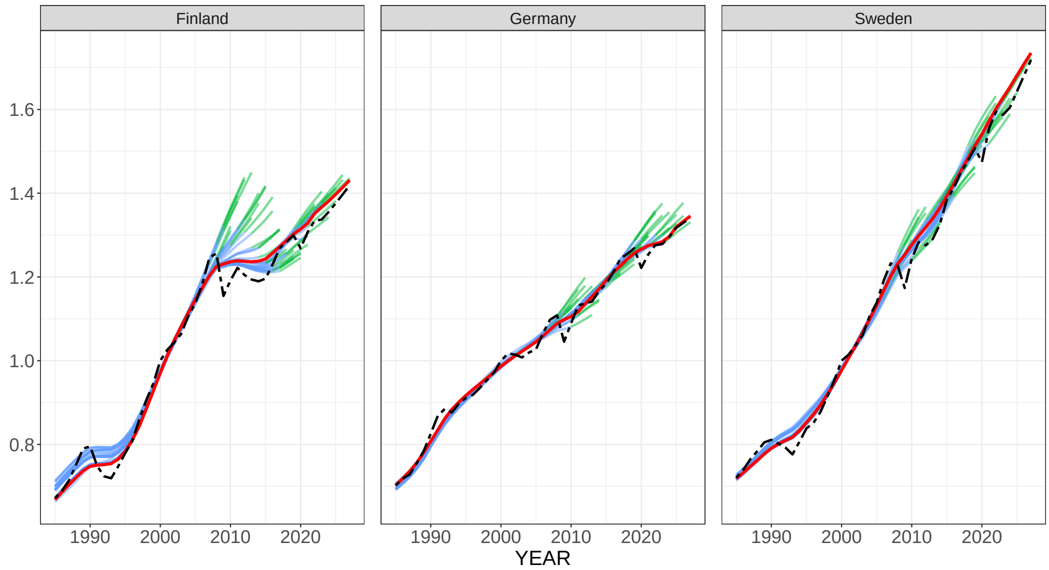

Figure 4.2: *EPC baseline estimation results for vintages of potential output (solid blue lines), latest vintage of potential output (thick red line), latest vintage of real GDP (dashed line), and forecasts of potential output (green) for Finland, Germany, and Sweden. All values are normalized such that value in 2000 is equal to 1. Potential GDP series is shifted by the value of output gap in 2000.*

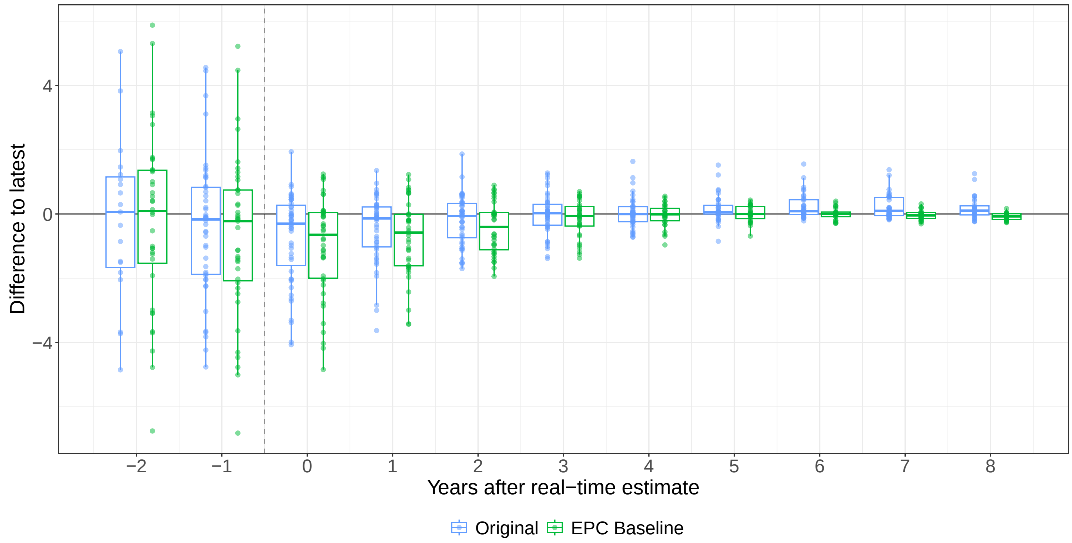

Figure 4.3: *Gaps to the latest vintage of output gap estimate by years after the real-time estimate for Finland. Dots are the observed values. Boxes present the 25th percentile, median, and 75th percentile. Whiskers extend to the furthest value no more than 1.5 times the interquartile range from the boxes. Positive values indicate that the estimate is above the latest estimate.*

fact, the overly optimistic view of the level of potential output persists all the way into the 2014 spring vintage of the data. The reason our results contain a longer forecast period than the ones in figure 3.1 is that the AMECO database only reports forecasts up to $t + 2$.

In figure 4.3 the differences between the current estimates and the latest available estimate are plotted, this time for the original AMECO data and our baseline specification estimates. It can be seen that at the longer horizons the gaps to latest estimate are lower in the baseline specification than in original releases. This is expected, as in our estimations there are no methodological changes between the vintages.

On the other hand, our estimates fare worse at short horizons. This is most likely due to the fact that the original estimates produced by EC can be produced by different parametrisations of the models. Thus expert information can be incorporated into the model, producing more accurate real-time estimates. This is something that our approach does not do, thus suggesting that the changes in parametrisations across the years have helped in real-time performance of the original estimates.

Figure 4.4 presents the mean forecast values of potential and real output, as well as their revisions, relative to the value of real GDP in the year preceding the forecast exercise. Solid lines represent the forecasts, dashed lines represent the revised values. The grey area connecting the lines of different colour can be interpreted as the mean output gap. As explained earlier, we can see that the output gap is always forecast to close at $t + 5$. In the graph this is seen as the two solid lines connecting at the last forecast horizon. Germany does in fact manage to close and reach a positive output gap during between t and $t + 2$, after which the revised values again become negative.

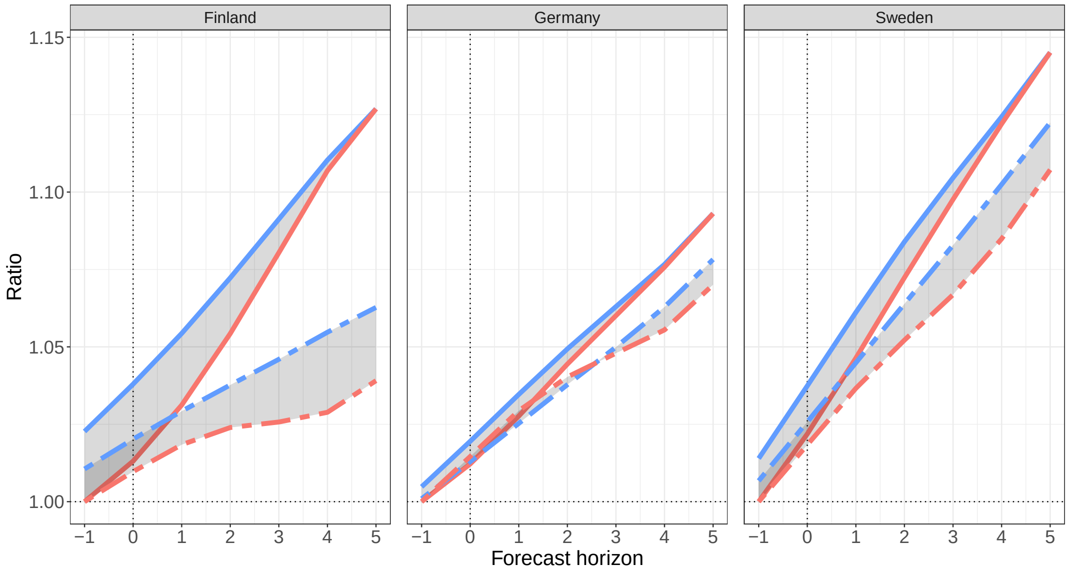

Figure 4.4: *EPC baseline estimates of potential (blue) and real (red) output together with their revised estimates (dashed lines) relative to output in year $t-1$. Revised values are taken to be the values from vintages five years after the initial estimate. The values are means over the vintages.*

Finland and Sweden do not manage to close the output gap at any forecast horizon. In Finland this is the case even though the potential output is on average revised heavily downwards, which is seen from the dashed blue line being significantly below the solid blue line.

4.1.1 Components of Revisions

To get a better understanding of the drivers of the large revisions in Finnish potential output (see figure 4.2), equation (3.2) allows the disaggregation of revisions into their subcomponents. We can write out the forecast error as

$$\bar{y}^{t}_{t+s} - \bar{y}^{t+s+i}_{t+s} = \bar{a}^{t}_{t+s} - \bar{a}^{t+s+i}_{t+s} + (1-\alpha)\left(k^{t}_{t+s} - k^{t+s+i}_{t+s}\right) + \alpha\left(\bar{\ell}^{t}_{t+s} - \bar{\ell}^{t+s+i}_{t+s}\right), \tag{4.1}$$

where the lower case letters are the logarithms of corresponding upper case letters, t is the time period when the forecast is made, and s is the forecast horizon. Subscripts denote the observation years, and superscripts the vintage years. The ”realized” value of the variable is taken to be the value in year $t+s$, as estimated in the vintage of $t+s+i$, naturally with $i \geq 0$.

Unfortunately, due to benchmark revisions of the national accounts series, it is not possible to use the aforementioned decomposition directly, as even in the real terms observations for a given year are not necessarily comparable across vintages. To alleviate (though not completely eliminate) the problem of different accounting practices, we can investigate growth rates instead of levels of the variables. This suggests the following decomposition of the prediction errors of growth rates:

$$\Delta\bar{y}^{t}_{t+s} - \Delta\bar{y}^{t+s+i}_{t+s} = \Delta\bar{a}^{t}_{t+s} - \Delta\bar{a}^{t+s+i}_{t+s} + (1-\alpha)(\Delta k^{t}_{t+s} - \Delta k^{t+s+i}_{t+s}) + \alpha(\Delta\bar{\ell}^{t}_{t+s} - \Delta\bar{\ell}^{t+s+i}_{t+s}). \tag{4.2}$$

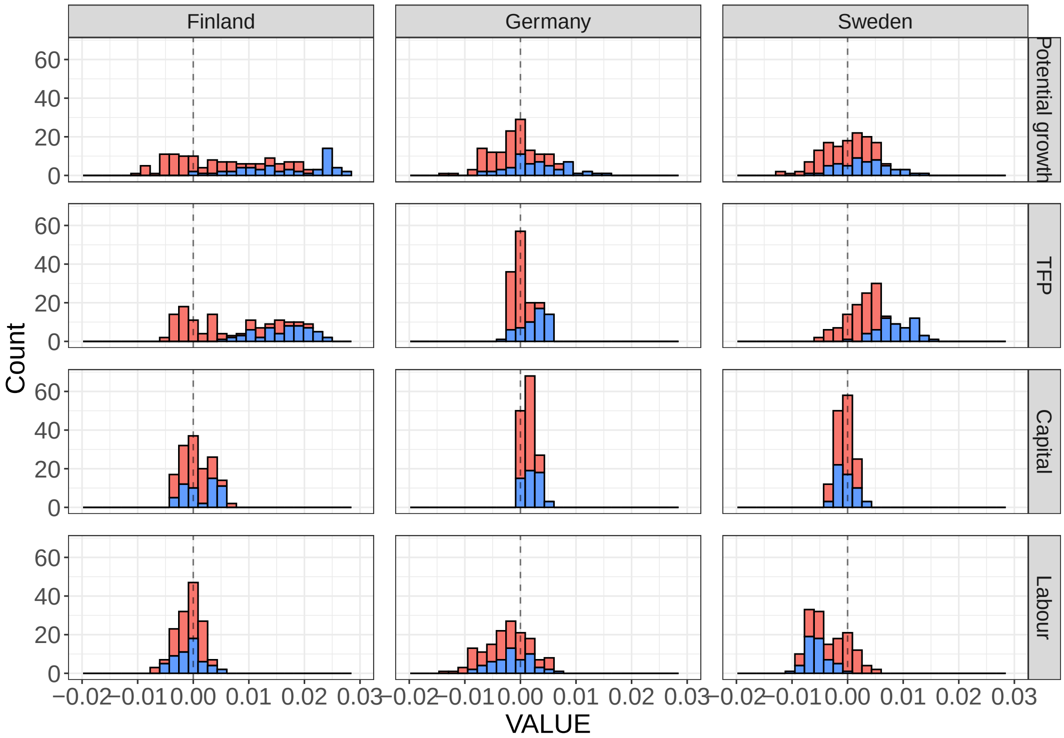

Figure 4.5: *Componenets of the revisions in growth rates of potential output. Red bars correspond to vintages starting from 2010, blue ones to vintages up to 2009. Forecast horizons from 0 to 5 included.*

Figure 4.5 presents the components of the revisions in growth rates for Finland, Germany, and Sweden, coloured by whether the vintage is before 2010. As can be seen, the forecasts of the growth rates of potential output perform the worst for Finland, with the forecasts being highly overly optimistic before 2010. Furthermore, we can see that the majority of these revisions is due to revisions in the growth rates of TFP. This is the case in all three countries, but the effect is the strongest in the case of Finland. An interesting observation is that labour growth rates tend to on average contribute negatively to the forecast errors of the output growth rate for Germany and Sweden, but not as much in Finland.

This approach also allows us to decompose the errors in the forecasts of potential labour input growth. Specifically we have

$$\Delta\bar{\ell}^t_{t+s} - \Delta\bar{\ell}^{t+s+i}_{t+s} = \Delta\ln(hperehp^t_{t+s}) - \Delta\ln(hperehp^{t+s+i}_{t+s}) + \Delta\ln(popw^t_{t+s}) - \Delta\ln(popw^{t+s+i}_{t+s}) + \Delta\ln(parts^t_{t+s}) - \Delta\ln(parts^{t+s+i}_{t+s}) + \Delta\ln(1-nawru^t_{t+s}) - \Delta\ln(1-nawru^{t+s+i}_{t+s}). \tag{4.3}$$

where *hperehp* is the HP-filtered trend of hours worked per person employed, *popw* is the population of working age, and *parts* is the participation rate.

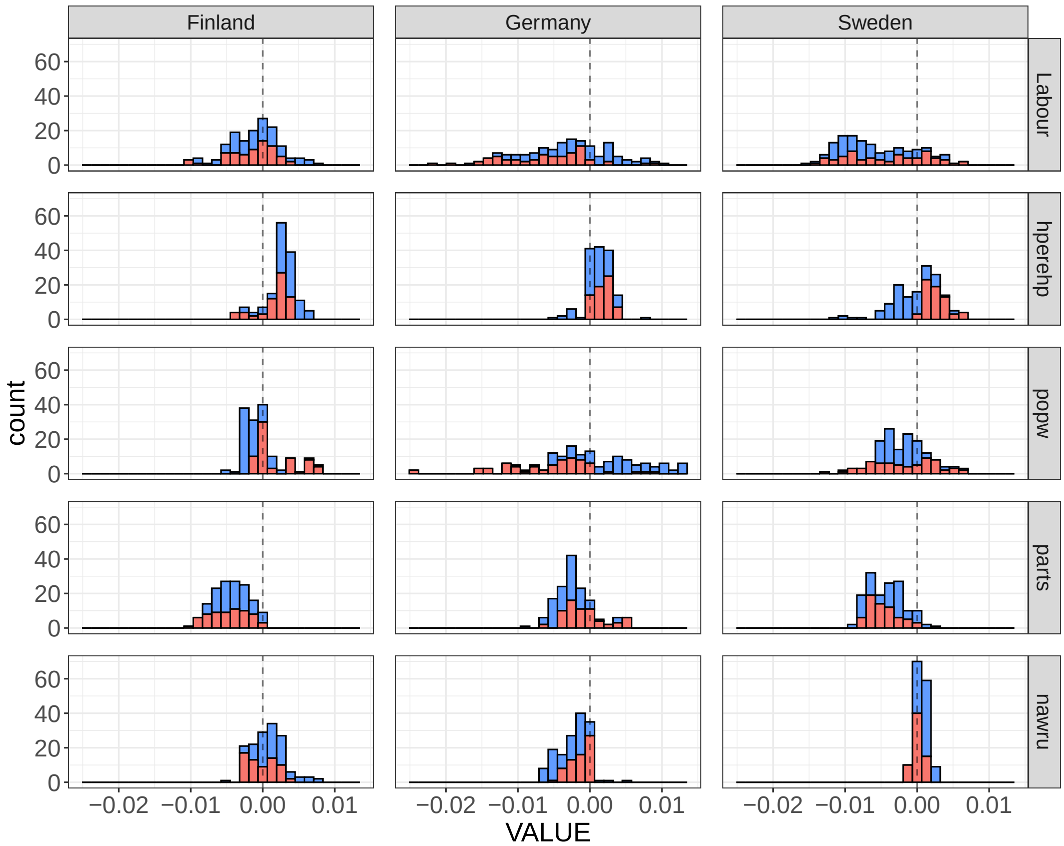

Figure 4.6: *Components of the revisions in growth rates of potential labour input. Error is defined as the difference between the forecast and the latest available estimate. Only observation years up to 2018 are included in the sample. Red bars correspond to vintages 2011–2015. Blue bars correspond to the rest of the vintages. Row names are as defined after equation* (4.3).

Figure 4.6 presents the components of these revisions for Finland, Germany, and Sweden. From the second panel from the top of the figure it can be seen that the forecasts of hours worked per person employed (hperehp) tend to be overly optimistic, with the effect appearing the strongest for Finland.

The forecasts of NAWRU are stable in comparison to the other components, with just a few outlier observations for Finland and Germany. This stability is in line with the goals and results presented in Hristov et al. (2017). The revisions of participation rate are quite stable, although biased downwards.

Finally somewhat surprisingly, the forecasts errors of working age population series in the middle panel have a large volatility, with German population projections faring the worst. Much of the German forecast errors come from vintages 2011–2015, which are coloured red in the figure. 2015 was the peak year of the migrant crisis in the EU, with Germany experiencing an especially large increase in immigration.[4] This seems like a plausible explanation for the large revisions in the contribution of population of working age.

4.2 Further Results

In addition to the baseline specification, we have also estimated two other specifications. One sets $\lambda = 1000$ in the HP-filter of participation rate (denoted "HP1000"). This decision was based on the considerations by Huovari et al. (2017), who notice that this setting produces a more convincing trend of the participation rate for Finland.

Another specification we have estimated is one in which we add a time trend to the AR extension of the series of hours worked per person employed. This choice was based on the persistent downward trend in the Finnish series.

Figure 4.7 displays the HP-filtered trends of hours worked per person employed for Finland and Sweden. As can be seen, the baseline specification, which does not contain a time trend, appears to suffer from a mean-reversion of the forecast values for Finland, implying a consistent downward revisions of the estimated trends. Including a time trend in the model fixes some of this mean-reversion, especially towards the latter years of the forecasts. However the inclusion of the time trend still does not alleviate the real-time optimism of the estimates. For Sweden including a time trend appears to create larger errors in the forecast series, as hours worked have stayed somewhat stable in the past two decades.

The mean-reversion of the baseline specification is problematic from the point of view of forecasting the potential level of output. The estimates of the trend of hours worked are systematically revised downwards, implying a downwards revision in the value of potential output, thus an upwards revision of the output gap. However although consistent, the change of the level of projected potential output arising from this bias is not too large. The change in projection for five years after the forecast exercise is roughly one percentage point. This is seen in figure 4.10.

The errors in the forecasts of the growth rates of hours worked per person employed are plotted in figure 4.8. The figure plots the errors for forecast horizon $s = 4$. It can be seen that an inclusion of the time trend works quite well for Finland, where the density of the errors gets centred around zero. However for Germany and Sweden this result does not hold, instead the errors get larger.

[4] According to Eurostat data, immigration to Germany was 77.5% higher in 2015 than in 2014. In Sweden it increased by 5.7% over the same period, and by 28.3% between 2014 and 2016. Immigration to Finland was 8.8% lower in 2015 than in 2014, and 10.8% higher in 2016 than in 2014.

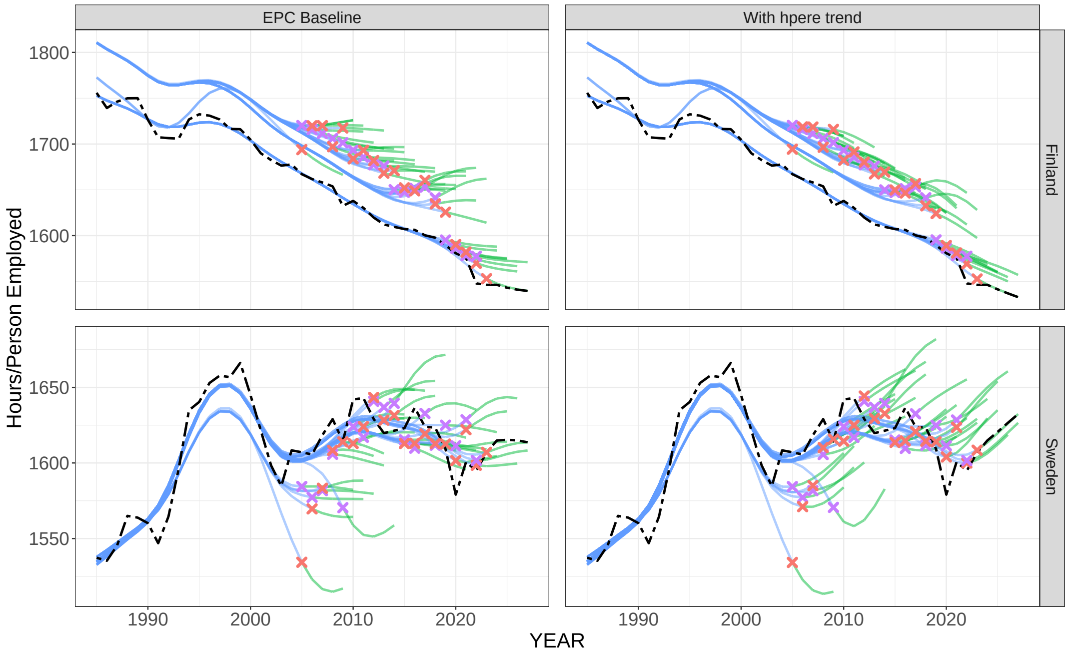

Figure 4.7: *HP-filtered trends (solid lines) of hours worked per person employed as reported by the EUCAM software for Finland and Sweden. Forecasts are marked with green. Crosses represent real-time estimates. Left panel displays the EPC baseline values, right panel displays the specification with a time trend added to the autoregression used to extend the hours worked series. Dashed line displays the raw series of the latest vintage.*

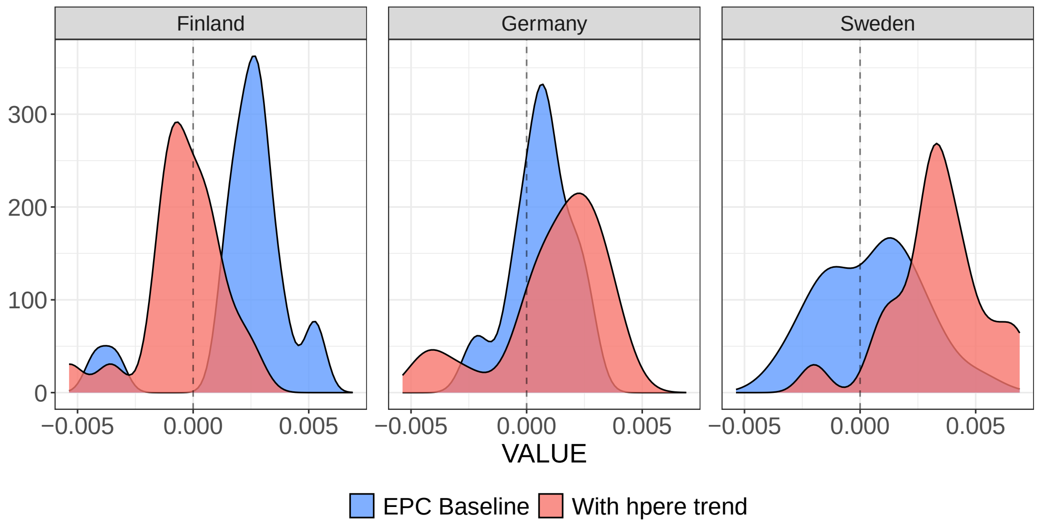

Figure 4.8: *Density of the errors in the forecasts of growth rates of hours worked per person employed, at forecast horizon of four years. Error is defined as the difference between the forecast and the latest available estimate. Only observation years up to 2018 are included in the sample.*

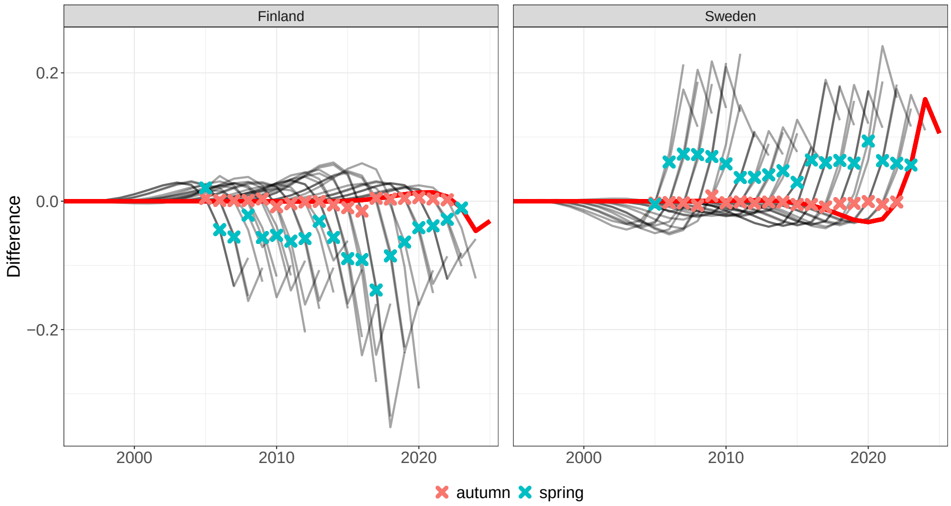

Figure 4.9: *Difference between the output gaps obtained in the EPC baseline specification and the specification with the time trend for hours worked. Crosses represent the real-time estimates. Red line corresponds to the latest available vintage.*

Figure 4.9 presents the differences in output gap estimates between the baseline specification and the hours worked time trend specification for Finland and Sweden. Negative values mean that output gap is more positive in the specification with a time trend. We can see that the real-time differences are concentrated in the spring vintages. This is most probably due to the fact that in spring releases the variables for current year are still forecasts. As can be expected from the differences between Finland and Sweden in figure 4.7, the sign of the differences is different between the two countries; inclusion of a time trend increases the output gap estimate for Finland, while it decreases for Sweden.

Another observation is that the magnitude of the differences keeps increasing in time, until the point in which the mechanical closing rule of the output gap in equation (3.10) kicks in. This is again quite expected, as the magnitude of the difference in forecasts of the trend in hours worked grows in time between the specifications.

In figure 4.10 we again plot the forecast and revised values of potential output and real output for Finland. The data is relative to real output in the year preceding the forecast exercise. The panels display results from different model specifications. We include a model with HP-parameter $\lambda = 1000$ for the participation rate, and a model with time trend for hours worked per person employed. The selection of $\lambda = 1000$ was based on the discussion in the article by Huovari et al. (2017).

Two observations can be made from this figure. Compared with the EPC baseline, the average forecasts (represented by solid lines) of the model with a time trend for hours worked in the rightmost panel are slightly lower at further forecast horizons. Thus they are closer to the revised values of the potential and real output (represented by dashed lines). The other observation is that the higher smoothing of participation rate creates slightly lower real-time estimates of the potential output, which in the graph is seen as the solid blue line being lower in the middle panel at forecast horizon 0. This brings the real-time estimate closer to the revised value than it is in the baseline specification.

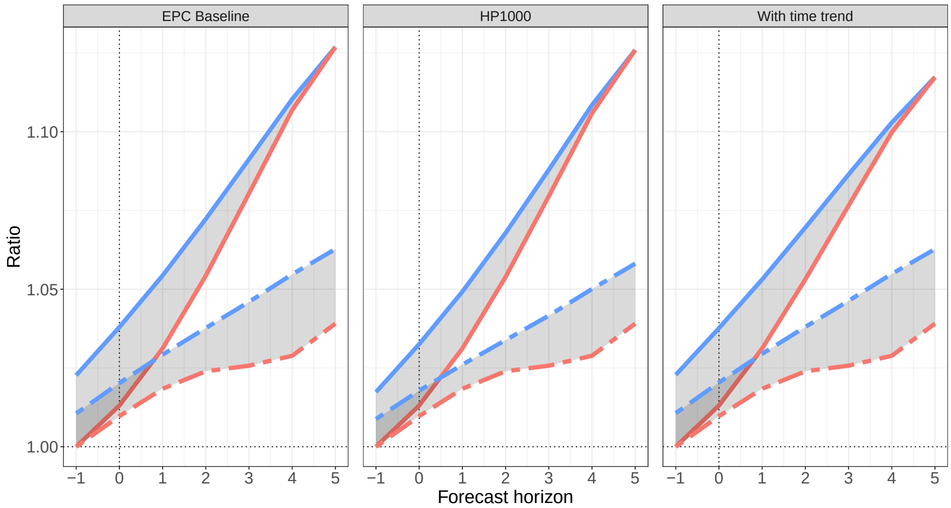

Figure 4.10: *Estimates of potential (blue) and real (red) output together with their revised estimates (dashed lines) for Finland, relative to output in year $t-1$. Revised values are taken to be the values from vintages five years after the initial estimate. The values are means over the vintages. "HP1000" corresponds to the model with $\lambda = 1000$ for the HP-filter of participation rate. "With time trend" denotes the model with a time trend in the extension of hours worked per person employed.*

Table 4.1 presents the root mean squared errors (RMSE) of the forecasts of the growth rate of potential output for the two alternate specifications of the model. The values are relative to the RMSE of the baseline specification. As can be seen, the specification with hours worked trend performs better for Finland across all forecast horizons. The specification with $\lambda = 1000$ perform even better at nowcasting, but loses performance as the forecast horizon increases.

The performance of the HP1000 specification across forecast horizons evolves differ-

		Forecast horizon					
	Specification	0	1	2	3	4	5
Finland	HP1000	0.96	0.98	0.99	1.01	1.03	1.03
	With hpere trend	0.99	0.97	0.95	0.94	0.94	0.93
Germany	HP1000	1.22	1.21	1.15	1.12	1.11	1.08
	With hpere trend	1.01	1.02	1.03	1.05	1.06	1.07
Sweden	HP1000	1.01	1.03	1.00	0.91	0.92	0.90
	With hpere trend	1.02	1.07	1.07	1.04	1.02	1.04

Table 4.1: *Root mean squared errors of the forecasts of potential output growth rate of different specifications relative to the root mean squared error of the EPC baseline specification. HP1000 refers to the specification with $\lambda = 1000$ for the HP-filter of participation rate. "hpere" denotes hours worked per person employed.*

ently in Finland than in Germany and Sweden. In Finland this specification loses forecasting power relative to the baseline specification as the horizon grows longer, becoming worse starting from a forecast horizon of three years. In Sweden, the opposite is the case, although the performance does not get better in each consecutive year. For Germany as the horizon increases the performance of the highly smoothed parametrization becomes better, but never surpasses the baseline.

For Germany, neither specification performs better than baseline. Specification with $\lambda = 1000$ performs especially bad at short horizons. For Sweden, neither specification is better at short term forecasts, but the highly smoothed specification outperforms baseline specification starting from forecast horizons of three years.

5 Conclusion

In this paper we have provided a descriptive analysis of the properties of the output gap estimates for Finland, as reported by EC, IMF, and OECD. In the case of Finland, Commission's methodology for estimating the potential output appears to reach its goal of reducing the amount of revisions in the estimates, as compared with the estimates of IMF and OECD. Specifically, the revisions are both smaller in magnitude, and have a lower standard deviation. The methodology tends to underestimate the output gap in real-time, as compared with the latest available revisions of the data. This is driven especially by the large errors in output gap estimates made before the financial crisis.

Comparing the EC estimates by country, the revision of output gap for Finland are larger than for comparison countries, Sweden and Germany. This pattern also holds for IMF's estimates, and part of OECD estimates.

An OLS regression of the revisions in output gap estimates against real GDP growth was also carried out, in order to gain insight into their cyclicality. The results suggest that with a couple outliers removed, EC's and IMF's estimates are roughly on par in cyclicality, and less cyclical than OECD's estimates. None of the estimates appear to be free of cyclicality.

For further analysis, we have used the 2023 spring revision of the software for calculating the potential output using EUCAM. This also allows to control for methodological changes. Our results suggest that the bulk of forecast errors in the growth rates of potential output in Finland come from overly optimistic views on the developments of productivity. These occurred mostly in the vintages prior to the financial crisis.

Our results also suggest that for Finland, small but systematic improvements of the forecast accuracy of the model can be achieved by simple reparametrisations. Specifically, addition of a time trend to the hours worked series improves the forecasts at a longer horizon, while higher smoothing of the participation rate improves short-term forecasts and real-time estimates. These results are not universal, as the addition of a time trend worsens the forecast performance for Germany and Sweden, while higher smoothing of participation rate only helps medium-term forecasts in Sweden.

References

Billmeier, A. (2004). Measuring a roller coaster: evidence on the Finnish output gap.

Blondeau, F., Planas, C., Rossi, A., et al. (2021). Output Gap Estimation Using the European Union's Commonly Agreed Methodology Vade Mecum & Manual for the EUCAM Software. *Directorate General Economic and Financial Affairs*.

Chalaux, T., & Guillemette, Y. (2019). The OECD potential output estimation methodology. *OECD Economics Department Working Papers*, (1563).

Chen, M. J., & Górnicka, L. (2020). Measuring Output Gap: Is It Worth Your Time? *IMF Working Papers*, (2020/024).

Coibion, O., Gorodnichenko, Y., & Ulate, M. (2017). The Cyclical Sensitivity in Estimates of Potential Output. *NBER Working Paper Series*, (23580).

D'Auria, F., Havik, K., Mc Morrow, K., Planas, C., Raciborski, R., Roger, W., Rossi, A., et al. (2010). The production function methodology for calculating potential growth rates and output gaps. *Economics Papers*, (420).

De Masi, M. P. (1997). IMF estimates of potential output: theory and practice. *IMF Working Papers*, (1997/177).

De Resende, C. (2014). An assessment of IMF medium-term forecasts of GDP growth. *IEO Background Paper No. BP/14/01*.

Denis, C., Grenouilleau, D., Mc Morrow, K., Röger, W., et al. (2006). Calculating potential growth rates and output gaps-A revised production function approach. *Economics Papers*, (247).

Denis, C., Mc Morrow, K., Röger, W., et al. (2002). Production function approach to calculating potential growth and output gaps-estimates for the EU Member States and the US. *Economics Papers*, (176).

European Commission. (2018). Vade Mecum on the Stability and Growth Pact. 2018 edition. *European Economy Institutional Paper*, (075).

Grigoli, F., Herman, A., Swiston, A., & Di Bella, G. (2015). Output gap uncertainty and real-time monetary policy. *Russian Journal of Economics*, *1*(4), 329–358.

Haavio, M. (2008). Tuotantokuilu Suomessa. *Publication of Bank of Finland*, (2008/4).

Havik, K., Mc Morrow, K., Orlandi, F., Planas, C., Raciborski, R., Röger, W., Rossi, A., Thum-Thysen, A., Vandermeulen, V., et al. (2014). The production function methodology for calculating potential growth rates & output gaps. *Economic Papers*, (535).

Hristov, A., Planas, C., Roeger, W., & Rossi, A. (2017). NAWRU Estimation Using Structural Labour Market Indicators. *European Economy Discussion Papers*, (069).

Huovari, J., Jauhiainen, S., & Kekäläinen, A. (2017). Assessment of the EC cyclical adjustment methodology for Finland–impact on budget balances. *PTT Working Papers*, *191*.

IMF. (2014). Finland: Selected issues. *IMF Country Report*, (2014/140).

Jysmä, S., Kiema, I., Kuusi, T., & Lehmus, M. (2019). The Finnish Potential Output: Measurement and Medium-term Prospects. *Publications of the Government's analysis, assessment and research activities*, (13/2019).

Kangur, M. A., Kirabaeva, K., Natal, J.-M., & Voigts, S. (2019). How Informative Are Real Time Output Gap Estimates in Europe? *IMF Working Papers*, (2019/200).

Melolinna, M. (2010). Euroalueen ja Suomen tuotantokuilu. *Publication of Bank of Finland*, (4/2010).

Orjasniemi, S., & Kuusela, A. (2021). Output Gap Estimates for Finland with Multivariate Filters. *Publications of the Finnish Econimic Policy Council*.

Orphanides, A., & Norden, S. v. (2002). The unreliability of output-gap estimates in real time. *Review of economics and statistics*, *84*(4), 569–583.

Ripatti, A., & Vilmunen, J. (2001). Declining labour share-Evidence of a change in underlying production technology? *Bank of Finland Discussion Papers*, (10/2001).

Valtiovarainministeriö. (2022). Taloudellinen katsaus Syksy 2022. *Valtiovarainministeriön julkaisuja*, *2022:58*.