CESifo Working Papers
10245
2023
January 2023

Estimation of the TFP Gap for the Largest Five EMU Countries

Kai Carstensen, Felix Kießner, Thies Rossian

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10245*

Estimation of the TFP Gap for the Largest Five EMU Countries

Abstract

In this paper we augment the Bayesian unobserved components model of the EU Commission to estimate the cyclical component of total factor productivity (TFP gap) with a factor structure to include a wide array of business cycle indicators. We demonstrate that this model extension considerably stabilizes the estimate of the of the TFP gap. Specifically, consider the usual autumn forecast of the EU Commission in October of a year T. For the last two “in-sample” years $T-2$ and $T-1$, and for the “now-cast” year T, the year-to-year revisions can be reduced by up to 30 percent. Improvements for the two “out-of-sample” years $T+1$ and $T+2$ also considered relevant by the EU Commission are quantitatively smaller (up to 10 percent) but still relevant. The results do vary across countries but are qualitatively robust with respect to different indicator sets, model specifications or vintages considered.

JEL-Codes: C320, E370.

Keywords: trend-cycle decomposition, unobserved components model, factor model, Bayesian estimation, total factor productivity, EU Commission.

*Kai Carstensen**
University of Kiel / Germany
carstensen@stat-econ.uni-kiel.de

Felix Kießner
University of Kiel / Germany
kiessner@stat-econ.uni-kiel.de

Thies Rossian
University of Kiel / Germany
thies.rossian@gmx.de

*corresponding author

January 15, 2023
We thank Jens Boysen-Hogrefe, Alessandro Rossi, Werner Roeger, Anna Thum-Thysen, Francesca Maria D’Auria and the participants of the meeting of the Output Gap Working Group in Bucharest on 17th May 2019 for helpful discussion and comments. Financial support by the German Federal Ministry for Economics and Energy for the research project 34/17 (IC4 - 80 14 38/034) is gratefully acknowledged.

1 Introduction

As part of the European Semester, the EU Commission annually assesses the fiscal plans of its member countries. A cornerstone of the procedure is to calculate the structural government budgets, i.e., budgets adjusted for the business cycle situation. To this end, a reliable estimate of the output gap is needed that exhibits only small revisions in order to keep fiscal (and political) over- and under-reactions as small as possible.

The EU Commission estimates potential output from a Cobb-Douglas production function determining output as a function of labour, capital and total factor productivity (TFP), see Havik et al. (2014). Assuming the capital stock is largely non-cyclical, the output gap is determined the cyclical components of labour and TFP where the latter is quantitatively more important, see for a discussion Dovern and Zuber (2019), which is why we concentrate on it.

In this paper we augment the Bayesian unobserved components model of the EU Commission, henceforth referred to as the "EU model", to estimate the cyclical component of TFP (TFP gap) with a factor structure to include a wide array of business cycle indicators. Thereby, we intend to stabilize the estimate of the TFP gap. We demonstrate that this model extension considerably improves the nowcast of the TFP gap. To this end, we mimic the usual autumn forecast of the EU Commission that is executed in October of a year T and thus based on incomplete knowledge of this year's macroeconomic situation. For the last two fully known in-sample years $T-2$ and $T-1$, the year-to-year revisions can be reduced by up to 30 percent. For the partly unknown nowcast year T, the reduction in the revision size is still around 20 percent. Improvements for the two fully unknown out-of-sample years $T+1$ and $T+2$, which are also considered relevant by the EU Commission, are quantitatively smaller (up to 10 percent) but still relevant.

A strength of our approach is that it does not change the general setup of the EU model. Both the basic model structure and the Bayesian estimation method remain unaltered. Technically, we simply add a block of factor model equations to improve the stability of the TFP gap estimates. Thereby, it should be easier to convince the EU Commission and the member states to adopt the extension, while they might be much more hesitant to radically change the modeling philosophy.[1]

Our paper relates to the large literature on output gap estimation. One strand of this literature shows that it is beneficial in terms of estimation precision and economic interpretability to extract the output gap by means of multivariate unob-

[1]In fact, we were invited to present our approach at the meeting of the Output Gap Working Group in Bucharest on 17th May 2019 and have been in contact to the EU Commission since then. Our impression has generally been that both member states and the EU Commission are highly interested in modifications of the current EU model that reduce the revisions of the TFP gap.

served component (UC) models applied to GDP plus a low-dimensional vector of business cycle indicators such as inflation and unemployment, see Kuttner (1994), Planas et al. (2008), Basistha and Startz (2008), and Fleischman and Roberts (2011) for the US, Valle e Azevedo et al. (2006) and Jarociński and Lenza (2018) for the euro area, Melolinna and Toth (2019) for the UK, and Blagrave et al. (2015) for a large set of countries.[2]

Another strand of the literature extracts cyclical factors from high-dimensional panels of indicator variables. Aastveit and Trovik (2014) apply a dynamic factor model to a set of 54 US indicators in order to obtain GDP nowcasts that reduce the end-of-sample problem of the HP filter documented by Orphanides and van Norden (2002); Mise et al. (2005). Weiske (2018) smoothes the first principal component of a set of 37 (stationary) business cycle indicators and shows that it is a reliable real-time output gap estimator for the four largest euro area countries. Barigozzi and Luciani (2019) extract the common cycle from a set of 103 (nonstationary) US variables and show that it is a competitive output gap estimator exhibiting only moderate revisions over time.

This paper integrates these two strands by including a factor structure in an otherwise standard UC model of the output gap as used by the EU Commission. Specifically, we assume that TFP and our panel of business cycle indicators share the same common cycle. We estimate this cycle in a joint model which has the advantage that extraction of the common cycle is supervised in the sense that it needs to account for both the cyclical variation in the indicator panel and in TFP.

The paper is structured as follows. Section 2 introduces the EU model and our extensions of it. In Section 3 discusses the priors, Section 4 presents the data sets used, and Section 5 describes the estimation procedure. Section 6 presents the results of a recursive revision analysis for our preferred specifications, while Section 7 conducts a sensitivity analysis. Section 8 discusses the plausibility of the estimated trend and cycle components and Section 9 concludes.

2 Trend-cycle models

In this section we first describe the UC model currently used by the EU Commission to estimate the TFP gap. Subsequently, we introduce and discuss our extensions.

[2]Another way to increase the information set is to use geographically disaggregate data such as state-level GDP for the US output gap (González-Astudillo, 2019b) or country-level data for the euro area output gap (González-Astudillo, 2019a; Huber et al., 2020). We do not follow this approach as we are interested in output gap estimates for individual member states of the euro area. However, we include aggregate euro area indicators in our factor approach for each country.

2.1 The EU model

The EU Commission currently uses an unobserved components model to estimate the TFP gap. Log TFP, y_t, is decomposed linearly into trend, p_t, and cycle, c_t, by means of the observation equation

$$y_t = p_t + c_t. \tag{1}$$

The trend is assumed to follow a random walk with AR(1) slope, μ_t,

$$\Delta p_t = \mu_{t-1} \tag{2}$$

$$\mu_t = \omega(1-\rho) + \rho\mu_{t-1} + a_{\mu t}, \qquad a_{\mu t} \overset{\text{iid}}{\sim} N(0, V_\mu), \tag{3}$$

where $|\rho| < 1$ to ensure stationarity of μ_t. The first observation μ_1 is assumed to be drawn from its unconditional distribution (we use the same assumption for all stationary AR processes discussed below). The cycle is modeled as a mean zero AR(2) process

$$c_t = 2A\cos\left(\frac{2\pi}{\tau}\right)c_{t-1} - A^2 c_{t-2} + a_{ct}, \qquad a_{ct} \overset{\text{iid}}{\sim} N(0, V_c), \tag{4}$$

where the parameterization with periodicity $\tau > 0$ and amplitude $0 < A < 1$ ensures stationarity and cyclical behavior.

To reduce trend revisions at the sample end, Havik et al. (2014) relate the TFP cycle to the Capacity-Utilization-and-Business-Sentiment (CUBS) indicator, u_t, a weighted average of survey-based capacity utilization in the manufacturing sector and business sentiment indicators in the construction and services sectors that are derived from the Business Tendency Surveys of the EU Commission.[3] This gives rise to the second observation equation

$$u_t = \mu_u + \beta c_t + e_{ut}. \tag{5}$$

To allow for somewhat persistent deviations between CUBS and the cycle, the error is modeled as a mean-zero AR(1) process:

$$e_{ut} = \delta_u e_{u,t-1} + a_{ut}, \quad a_{ut} \overset{\text{iid}}{\sim} N(0, V_u), \tag{6}$$

where $|\delta_u| < 1$ to ensure stationarity.

Altogether, the EU model consists of the observation equations (1) and (5), and the state equations (2), (3), (4), and (6). It is cast in state space form and estimated with Bayesian methods using a Metropolis-within-Gibbs sampler. The priors are country-specific and discussed below.

[3] The reliability of capacity utilization as an indicator of the output gap and TFP gap is documented by Graff and Sturm (2010), D'Auria et al. (2010), Planas et al. (2013), and Turner et al. (2016).

2.2 Extension 1: Two-step factor model

Our first extension of the EU model simply adds an externally estimated business cycle factor, f_t, by means of another observation equation that links it to the TFP cycle:

$$f_t = \alpha c_t + e_{ft}, \tag{7}$$

where we neglect an intercept because the factor will be standardized and the cycle has a mean of zero by construction.

We again allow for persistence in the error, hence we model it as the AR(1) process

$$e_{ft} = \delta_f e_{f,t-1} + a_{ft} \quad a_{ft} \overset{\text{iid}}{\sim} N(0, V_f), \tag{8}$$

where $|\delta_f| < 1$ to ensure stationarity.

Furthermore, we make the following restriction regarding the smoothness of the trend in (2)

$$|\Delta_2 p_t| \leq \zeta_p, \tag{9}$$

where ζ_p is a threshold derived as the minimal smoothness of the trend growth rate of all countries and vintages of an HP-filter estimate with $\lambda = 6.25$. We introduce this restriction to keep the trend from taking up business cycle fluctuations.

Estimation proceeds in two steps. First, we estimate the factor from a set of business cycle indicators described below. To this end, we either use principle components (PCA) or a Bayesian factor model (Bay) presented in Appendix A.1 and A.2, respectively. Then we feed the factor in the extended model which consists of the observation equations (1), (5), and (7), and the state equations (2), (3), (4), (6), (8) and (9). Depending on which factor we use, we denote the model by 2-Step-PCA or 2-Step-Bay. We estimate it with Bayesian methods.

2.3 Extension 2: One-step factor model

In our second extension of the EU model, we integrate factor estimation and trend-cycle decomposition into one model which we estimate in a single step. This one-step approach has the advantage that parameter estimation and factor extraction is "supervised" in the sense that both the co-movement of the factor and the cyclical component of TFP is taken into account.

To facilitate one-step estimation, we add a set of k measurement equations to the EU model. Defining the $k \times 1$ vector of observed yearly standardized business cycle indicators, x_t, we assume the factor structure

$$x_t = \Lambda f_t + e_{xt}, \tag{10}$$

where $\Lambda = (\lambda_1, \dots \lambda_k)'$ is a $k \times 1$ vector. The errors follow the diagonal VAR(1) process

$$e_{xt} = \Gamma e_{x,t-1} + a_{xt} \quad a_{xt} \stackrel{\text{iid}}{\sim} N(0, V_x), \tag{11}$$

where V_x is a diagonal $k \times k$ covariance matrix with coefficients v_{ii}, $i = 1, \dots, k$, on the main diagonal, and Γ is a diagonal $k \times k$ matrix with coefficients $|\Gamma_{ii}| < 1$, $i = 1, \dots, k$, on the main diagonal to ensure stationarity. Altogether, this model, henceforth denoted by 1-Step, consists of the observation equations (1), (5), and (10), and the state equations (2), (3), (4), (6), (7), (8), (9) and (11). We estimate it with Bayesian methods.

2.4 Hodrick-Prescott (HP) filters

We compare the EU model and its extensions to the HP filter both with standard settings and with estimated signal-to-noise ratio. The standard HP filter is applied with two parameterizations. While a usual choice of the smoothing parameter for annual data is $\lambda = 100$, Ravn and Uhlig (2002) recommend $\lambda = 6.25$.

To estimate the signal-to-noise ratio, we consider the state space form of the HP filter (Harvey and Trimbur, 2008):

$$y_t = \tilde{p}_t + \tilde{c}_t \qquad \tilde{c}_t \stackrel{\text{iid}}{\sim} N(0, \tilde{\sigma}_c^2), \tag{12}$$

$$\Delta \tilde{p}_t = \tilde{\mu}_{t-1} \tag{13}$$

$$\tilde{\mu}_t = \tilde{\mu}_{t-1} + \tilde{\varepsilon}_t, \qquad \tilde{\varepsilon}_t \stackrel{\text{iid}}{\sim} N(0, \tilde{\sigma}_\varepsilon^2). \tag{14}$$

While the underlying assumption of an iid cycle $\tilde{c}_t$ is certainly unrealistic, it is routinely made in all papers using the HP filter. We estimate the two variances $\tilde{\sigma}_c^2$ and $\tilde{\sigma}_\varepsilon^2$, and thus implicitly the signal-to-noise ratio $\tilde{\lambda}^{-1} = \tilde{\sigma}_\varepsilon^2 / \tilde{\sigma}_c^2$, with Bayesian methods. We denote this estimated HP model by HP-EST.

3 Priors

We estimate the models with Bayesian methods, using a Metropolis-within-Gibbs algorithm, for the five largest EU countries Germany (DE), France (FR), Italy (IT), Spain (ES), and the Netherlands (NL). To replicate the results of the EU Commission, we use the EU model with the priors proposed by Havik et al. (2014). The priors for ω and ρ of the slope equation (3) are independently normal, the prior for the variance V_μ is independently inverse Gamma. A natural conjugate prior does not exist because we draw the first observation from its unconditional distribution. The priors for A and τ of the cycle equation (4) are independently

Beta, the prior for the variance V_c is independently inverse Gamma. For the parameters μ_u, β, and V_u of the CUBS equations (5) and (6) we assume a natural conjugate Normal-inverse Gamma (NiG) prior implying marginal t distributions for μ_u and β. The persistence parameter δ_u is a priorily independently normal. Details of the priors is shown in Table 1. The parameterization (including truncation) for the coefficients of the EU model, ω, ρ, V_μ, τ, A, V_c, β, μ_u, δ_u, and V_u, are taken from Havik et al. (2014).

The two-step factor models add the observation equation (7) and the AR(1) error equation (8) which altogerther feature three additional parameters. The prior for δ_f is truncated normal with mean zero, fairly large standard deviation of 1, and stationarity-preserving truncation bounds. The prior for α and V_f is NiG. The prior parameters are empirical in the sense that they are taken from a preliminary regression of the factor, f_t, on an initial estimate of the cycle obtained by applying the HP filter to log TFP. Specifically, the prior means of α and V_f are set to their OLS estimates, the prior standard deviations to four times their OLS standard errors.

The one-step factor model adds k observable business cycle indicators and treats the factor, f_t, as an unobservable state variable. To achieve identification, we set $\alpha = 1$, while the inverse Gamma prior V_f is parameterized empirically as described for the two-step factor model. Concerning the factor equations (10), the diagonal structure of Γ and V_x allows to specify equationwise priors for λ_i, v_{ii}, and Γ_{ii}. For the persistence parameters Γ_{ii} we use a truncated normal prior with mean zero, fairly large standard deviation of 1, and stationarity-preserving truncation bounds. For λ_i and v_{ii}, we apply an empirical NiG prior. It is specified by regressing an initial estimate of the factor, obtained by applying the HP filter to log TFP, on each business cycle indicator. Specifically, the prior means of λ_i and v_{ii} are set to their OLS estimates, the prior standard deviations to four times their OLS standard errors.

Table 1: Priors

Parameter	marginal pdf	mean	s.d.	range
Parameters of the EU model				
ω	N	.015	.01	(0,.03)
ρ	N	.8	.24	(0,.99)
V_μ ($\times 10^{-6}$)	iG	2.4	2.4	(0,∞)
τ	B	8.0	3.5	(2,32)
A	B	.42	.17	(0,1)
V_c ($\times 10^{-4}$)	iG	3.0	3.0	(0,∞)
β	t	1.4	.7	(0,5)
μ_u	t	0	.03	(-1,1)
δ_u	N	0	.4	(0,.99)
V_u ($\times 10^{-3}$)	iG	4.5	4.5	(0,∞)
Additional parameters of the 2-step factor model				
α	t	empirical		$(-\infty, \infty)$
δ_f	N	0	1	(-.99,.99)
V_f	iG	empirical		$(0, \infty)$
Additional parameters of the 1-step factor model, $i = 1, \dots, k$				
λ_{ii}	t	empirical		$(-\infty, \infty)$
Γ_{ii}	N	0	1	(-.99,.99)
υ_{ii}	iG	empirical		$(0, \infty)$

Notes: N denotes the normal density, iG the inverse Gamma density, B the four-parameter Beta density, and t the three-parameter t density. Note that the prior mean and standard deviation of V_μ, V_c, and V_u presented in Havik et al. (2014) are partly incorrect. We use the values implemented in the actual software used by the EU Commission which is available at the 'Output Gap' interest group of the CIRCABC website of the EU Commission. In addition, the values for some countries deviate from the ones in the Table above. Specifically, prior mean and standard deviation of V_μ for France are .3 and .2, for Spain .5 and .5, and for the Netherlands .4 and .4. The prior mean and standard deviation of V_c for Spain are 2.5 and 2. The prior mean and standard deviation of V_u for France are 2.5 and 2.5, for Italy 3 and 3, for Spain 1 and .4, and for the Netherlands 1 and 1. Finally, to achieve identification, the prior of α has mean 1 and standard deviation zero in the one-step factor model.

4 Data

We run the models with annual data from 1980 to 2021. Since we perform a recursive revision analysis that mimics the autumn projections of the EU Commission, we take the real-time nature of the data into account as far as possible.

Log TFP is calculated as

$$\log(TFP) = \log(GDP) - 0.65\log(L) - 0.35\log(K),$$

where L and K are labor and capital, respectively. It is provided by the EU Commission as vintage data that reflect the information known at the end of October each year. Since TFP data for year T are not fully available in the autumn, the EU Commission uses its economic projections to fill up the missing quarters. It even provides TFP forecasts of the years $T+1$ and $T+2$ to alleviate the usual end-point problem of gap estimates. We follow this approach and integrate these forecasts into our estimation approach.

The CUBS indicator is also published by the EU Commission and is available in vintages from 2009 onwards when it was introduced. We construct pseudo real time vintages for earlier years.

The business cycle factor of each country is constructed from a set of 42 indicators that are similar across countries and mostly available since 1980. They consist of three parts: domestic survey indicators, domestic hard indicators, and international indicators. We use 12 domestic survey indicators, all of which are taken from the harmonized EU surveys and provided by the EU Commission. They include, inter alia, capacity utilization, economic sentiment, appreciation of new orders, and production expectations. The domestic hard indicators cover the areas production and sales (8 indicators), labor market (3 indicators), prices (3 indicators), income and consumption (4 indicators), and finance (3 indicators). The international indicators cover the euro area (5 indicators), and the US (4 indicators). For details see Appendix B.

We convert all indicators to annual frequency before we make them stationary by essentially applying the transformations recommended by Stock and Watson (2002) for US data. While we only have revised data, we take the usual publication lags into account. This entails that at the sample end we include only observations that are known at the end of October of that year. While it would be optimal to have real-time vintages, only a few of our indicators are revised markedly at the monthly or quarterly frequency and these revisions typically average out to a large extent once they are aggregated to the annual frequency. In addition, Marcellino and Musso (2011) document that data revisions are a minor source of real-time output gap revisions for the euro area and Aastveit and Trovik (2014) show that using a factor model can strongly reduce the sensitivity of output gap estimates to data revisions. Hence, a quasi real-time approach to our indicators appears sufficient to compare output gap revisions across different models.

To study the relevance of the indicators, we run the factor models with three different sets: only the domestic survey indicators (henceforth denoted by SUR), all domestic indicators (DOM), and all indicators (BIG).

To account for the possibility that an indicator leads or lags the TFP cycle, we proceed as follows. We estimate a preliminary cycle by HP filtering, shift each indicator up to four quarters forward and backwards, and compute the correlation between the preliminary cycle and the shifted indicator. We include in the indicator panel the lead or lag which results in the largest absolute correlation. The indicator sets based on the optimally time-shifted indicators are labeled SURo, DOMo, and BIGo, respectively, while the indicator sets without time shift are denoted as SUR, DOM, and BIG.

5 Estimation

In the following we describe estimation of the one-step factor model because it encompasses both the EU model and the two-step factor model. We follow the Bayesian approach of Havik et al. (2014) but apply it to our extended model.

5.1 Likelihood

Since the one-step factor model is linear and the variables are normally distributed, it is straightforward to write it in state space form and derive the likelihood function.

5.2 Posterior simulation

The structure of the model suggests to apply a Metropolis-within-Gibbs simulation of the posterior distribution. We briefly comment on each Gibbs step.

Sampling the unobserved components. To sample trend, cycle, factor, and the missing observations in the CUBS indicator and the business cycle indicators conditional on the parameters, we directly draw from the conditionally normal posterior distribution imposing a diffuse initial condition for the trend.

Sampling ω, ρ, V_μ. Conditional on the trend, p_t, these are the coefficients of a stationary AR(1) model with the first observation drawn from its unconditional distribution. The posterior of ω given ρ and V_μ is truncated normal and the posterior of V_μ given ρ and ω is inverse Gamma, while the posterior of ρ given ω and V_μ needs to be sampled by means of a Metropolis-Hastings (MH) step. The proposal distribution is a truncated normal distribution as suggested by Havik et al. (2014).

Sampling A, τ, V_c. Conditional on the cycle, c_t, these are the coefficients of a stationary AR(2) model with the first two observations drawn from their unconditional distribution. Due to the specific parameterization, only the posterior of V_c given A and τ is analytical (inverse Gamma). To sample A we apply a MH step with the asymptotic normal distribution of the maximum likelihood estimator truncated to the unit interval as proposal pdf. To sample τ we again apply a MH step with the asymptotic normal distribution of a crude maximum likelihood estimator truncated to the prespecified range as proposal pdf.

Sampling μ_u, β, V_u. Conditional on the CUBS indicator u_t, the cycle c_t, and the AR(1) parameter δ_u, these are coefficients of a regression model with known autoregressive error structure. Given the NiG prior, the posterior is again NiG from which we can sample easily.

Sampling δ_u. Conditional on the residuals of the CUBS equation obtained in the previous step, this is the coefficient of an AR(1) model with known variance, where the first observation is drawn from its unconditional distribution. To sample δ_u we apply a MH step with the proposal pdf being the truncated normal posterior of an AR(1) model that conditions on the first observation.

Sampling Λ *and* V_x. By construction of the model, we can sample the elements λ_i and v_{ii} equationwise. Conditional on the factor and the AR(1) parameters Γ_{ii}, these are coefficients of a regression model with known autoregressive error structure. Given the NiG prior, the posterior is again NiG from which we can sample easily.

Sampling Γ. Due to the diagonal structure of Γ, we can sample each element Γ_{ii} separately. Conditional on the residuals of the indicator equations obtained in the previous step, Γ_{ii} is the coefficient of an AR(1) model with known variance, where the first observation is drawn from its unconditional distribution. We apply a MH step with the proposal pdf being the truncated normal posterior of an AR(1) model that conditions on the first observation.

Sampling α *and* V_f. Conditional on the factor f_t, the cycle c_t, and the AR(1) parameter δ_f, these are coefficients of a regression model with known autoregressive error structure. Given the NiG prior, the posterior is again NiG from which we can sample easily.

Sampling δ_f. Conditional on the residuals of the factor equation obtained in the previous step, this is the coefficient of an AR(1) model with known variance, where the first observation is drawn from its unconditional distribution. To sample δ_f we apply a MH step with the proposal pdf being the truncated normal posterior of an AR(1) model that conditions on the first observation.

We take 10,000 draws from the Gibbs sampler of which we discard the first 2,000 as burn-in sample. Convergence is checked by a diagnostic test that compares the means of the first 10 percent and the last 40 percent of the retained draws (Geweke, 1992).

6 Recursive revision analysis

We conduct a recursive revision analysis of the vintage years $T = 2005, \ldots, 2021$ to quantify the year-on-year cycle revisions of the various estimators presented above. We first describe the setup and then report the results.

6.1 Setup

To closely mimic the situation of the EU Commission which estimates the cycle in the autumn of each year, we use only information that is available by the end of October in each of those years. In particular, all steps of the estimation procedures, such as indicator selection, factor estimation, and empirical prior specification, are repeated recursively.

Havik et al. (2014) include in each year T the EU Commission's TFP forecast of years $T+1$ and $T+2$ as if they were observations in order to alleviate the end-point problem of any filter. While it renders all results dependent on the quality of those forecasts, we nevertheless follow this approach to be as close as possible to the EU Commission. This allows us to study how changes in the estimation methodology affect the cycle revisions keeping all other modeling choices constant.

We measure the year-on-year revisions of the cycle by the root mean-squared error (RMSE). It is calculated as an average over each vintage year $T = 2005, \ldots, 2021$ of the horizons $T+h$, where $h = -2, \ldots, 2$. Let $c_{h,T}^{(C,M)}$ denote the TFP cycle estimate (in percentage points) of model M for country C and horizon h of vintage T. Then the revision RMSE for horizon h of model M and country C is defined as

$$R_h^{(C,M)} = \sqrt{\frac{1}{16} \sum_{T=2005}^{2021} \left(c_{h,T}^{(C,M)} - c_{h-1,T+1}^{(C,M)} \right)^2} \Big/ s_{2021}^{(C,M)}, \tag{15}$$

where we normalize by the standard deviation, $s_{2021}^{(C,M)}$, of the cycle estimated in the last vintage year. This normalization ensures that we do not prefer specifications that yield cycles of small amplitudes.

Below we report, for each country, the relative revision of model M compared to the EU model,

$$R_{h,rel}^{(C,M)} = R_h^{(C,M)} / R_h^{(C,\text{EU model})}.$$

Since the number of vintages analyzed is only 17 and thus quite small, we also report the average relative revision of model M across all five countries,

$$R_{h,rel}^{(M)} = \frac{1}{5} \sum_C R_{h,rel}^{(C,M)}.$$

Table 2: Revisions to next vintage, average over all five countries

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.39	0.49	0.57	0.98	0.92
	-		1.17	1.13	1.06	1.01	0.99
HP-EST	-		0.99	1.19	1.40	1.02	0.69
HP	-	$\lambda = 6.25$	0.82	1.04	1.28	**0.96**	**0.64**
	-	$\lambda = 100$	1.14	1.18	1.09	0.76	0.66
2-Step-Bay	SURo	$\delta_u = 0$	0.83	0.84	0.90	1.04	1.06
	DOMo	$\delta_u = 0$	0.81	0.81	0.82	**0.96**	0.90
	BIGo	$\delta_u = 0$	0.80	0.79	0.81	1.00	0.94
2-Step-PCA	SURo	$\delta_u = 0$	0.76	0.74	0.88	1.10	1.16
	DOMo	$\delta_u = 0$	0.78	0.76	**0.79**	1.05	1.15
	BIGo	$\delta_u = 0$	0.80	0.75	0.83	1.13	1.24
1-Step	SURo	$\delta_u = 0, \Gamma = 0$	**0.69**	0.72	0.84	1.00	0.88
	DOMo	$\delta_u = 0, \Gamma = 0$	**0.69**	**0.71**	0.82	1.02	0.86
	BIGo	$\delta_u = 0, \Gamma = 0$	0.70	0.73	0.84	1.01	0.83

Notes: The first row shows the revision, $R_h^{(C,M)}$, of the EU model averaged across countries. The remaining rows show the revision $R_{h,rel}^{(M)}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

6.2 Aggregate results of the revision analysis

The results of the revision analysis averaged over all countries are shown in Table 2. We treat the EU model with the parameter restriction $\delta_u = 0$ as baseline because it is used by the EU Commission for all five countries except France. The restriction precludes persistent deviations between the CUBS indicator and the TFP gap thus giving large weight to the indicator. For this baseline model we report the average normalized revision of the TFP gap in percentage points in row 1. For the backcast of $T-2$ and $T-1$, the average revision size of the EU model is 39 percent and 49 percent of the standard deviation of the cycle, while it is 57 percent for the nowcast of T. Not surprisingly, the revision size jumps upward for the out-of-sample forecasts $T+1$ and $T+2$, exhibiting values of almost 100 percent of the cycle.

The results for the EU model without a restriction on δ_u are shown for completeness in row 2 because this is the specification preferred by the EU Commission for France. Here and for all other models we display the average relative revision $R_{h,rel}^{(M)}$ compared to the baseline model. Relative revisions that are larger than 1 indicate that the baseline EU model exhibits, on average, smaller revisions. For example, the average revision of the unrestricted EU model at the sample end, $h = 0$, is 6 percent larger than the average revision of the baseline EU model. It turns out that it does not pay off to relax the restriction $\delta_u = 0$ although for the out-of-sample forecast horizons the EU model without parameter restriction is on par.

The HP model with estimated signal-to-noise ratio (HP-EST) is reported in row 3. It does not outperform the baseline model except for the two-year ahead forecast $h = 2$. The classical HP filter with $\lambda = 6.25$ and $\lambda = 100$ shown in rows 4 and 5 again lead – with one exception – to much larger revisions than the baseline model within the sample, i.e., for forecast horizons $h = -2$ to $h = 0$. However, it works surprisingly well for the out-of-sample periods $h = 1$ and $h = 2$ given that the end-of-sample problem of the HP filter is well-documented in the literature.

Let us now turn to our first extension of the EU model, the two-step factor model, where the factor is estimated separately – either with a Bayesian approach or by PCA – and then is fed into the trend-cycle model. To be as comparable as possible to our baseline EU model, we impose the restriction $\delta_u = 0$. It turns out that the two-step factor model improves markedly over the baseline EU model no matter how the factor is estimated and which indicator set is used. It reduces the average revisions by 20-25 percent for the backcasts ($h = -2$ and $h = -1$) and by 10-20 percent for the nowcast ($h = 0$), while it is comparable to the baseline EU model for the forecasts ($h = 1$ and $h = 2$).

Our second extension of the EU model is the one-step factor model which estimates the factor jointly with the trend-cycle decomposition. We again impose the restriction $\delta_u = 0$ as in the baseline EU model. In a similar vein, we restrict $\Gamma = 0$ and thus exclude persistent deviations between the factor and the indicators.[4] We find that this model is superior to both the baseline EU model and the two-step model, no matter which indicator set is used. It reduces the average revisions by around 30 percent for the backcasts ($h = -2$ and $h = -1$), almost 20 percent for the nowcast ($h = 0$), and around 15 percent for the two-step ahead forecast ($h = 2$), while it is on par with the baseline EU model for the one-step ahead forecast ($h = 1$).

We note that the differences between the three indicator sets are fairly small. We thus prefer the survey indicator set, SURo, because the data are easy to obtain and not revised by construction. This implies that a fully-fledged real-time analysis would deliver the same results. In contrast, the revision size using the indicator sets DOMo and BIGo might increase somewhat with truly real-time data even though probably not much as discussed above.

6.3 Country by country results of the revision analysis

Country by country results are shown in Tables 3 to 7. In the following we will discuss, for each country, the most important differences to the results averaged over all five countries presented above.

Germany. The baseline EU model works comparably to the five-country

[4]We discuss the effects of relaxing these restriction in Section 7.1.

Table 3: Revisions to next vintage, Germany

Model	Indicator set	Parameter restrictions	Relative revision $R^{(C,M)}_{h,rel}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.34	0.40	0.47	0.79	0.61
	-		1.06	1.08	1.05	0.90	0.93
HP-EST	-		0.82	0.85	1.03	0.73	**0.54**
HP	-	$\lambda = 6.25$	**0.70**	0.80	1.06	0.84	0.55
	-	$\lambda = 100$	1.02	1.07	1.03	**0.65**	0.59
2-Step-Bay	SURo	$\delta_u = 0$	0.94	0.99	0.93	1.01	1.00
	DOMo	$\delta_u = 0$	0.94	0.94	0.89	1.02	0.88
	BIGo	$\delta_u = 0$	0.94	0.91	0.87	1.02	0.89
2-Step-PCA	SURo	$\delta_u = 0$	0.90	**0.77**	**0.68**	1.15	1.39
	DOMo	$\delta_u = 0$	0.85	0.79	**0.68**	1.13	1.35
	BIGo	$\delta_u = 0$	0.91	0.80	0.70	1.14	1.38
1-Step	SURo	$\delta_u = 0$, $\Gamma = 0$	0.93	1.03	1.02	0.91	0.83
	DOMo	$\delta_u = 0$, $\Gamma = 0$	0.85	0.85	0.88	1.03	0.92
	BIGo	$\delta_u = 0$, $\Gamma = 0$	0.83	0.83	0.88	0.99	0.88

Notes: The first row shows the average revision, $R^{(C,M)}_h$, of the EU model. The remaining rows show the revision $R^{(C,M)}_{h,rel}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

average for the backcasts and the nowcast, while it exhibits much smaller revisions for the forecasts (see row 1 of Table 3). For example, the two-year ahead forecast revision is only 61 percent of the cycle standard deviation in Germany, while it is 92 percent in the five-country average shown in Table 2. Another notable result is that the HP filter generally outperforms the baseline EU model by far, reducing, e.g., two-year ahead forecast revision to almost 50 percent.

The two-step factor models behave similarly to the five-country average: It clearly reduces the revisions for the backcasts and nowcast, while it is comparable to the baseline EU model for the forecasts. Extracting the factor by PCA from the survey indicators set (SURo) is a particularly well suited to nowcast the TFP gap with revisions being reduced by almost one third.

The one-step factor model also improves over the baseline EU model but slightly less markedly than in the five-country average. This may reflect the extremely good performance of the baseline EU model for Germany. While the full indicator set (BIGo) is preferable, using the survey indicators set (SURo) is nevertheless a good choice, particularly with respect to forecasting the TFP gap. At the two-year horizon ($h = 2$), the revisions are reduced by 17 percent.

France. For France, the EU Commission uses the EU model without the restriction $\delta_u = 0$, i.e., they allow for a persistent error term in the CUBS equation. We find that this specification (row 2 in Table 4), performs slightly worse across all forecast horizons compared to the EU baseline model which we use as baseline. As in the case of Germany, the HP filter again outperforms the baseline EU model

Table 4: Revisions to next vintage, France

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.48	0.76	0.74	0.94	0.90
	-		1.11	1.08	1.03	1.01	1.02
HP-EST	-		0.82	0.75	0.77	0.72	0.54
HP	-	$\lambda = 6.25$	0.84	0.80	0.79	0.72	0.55
	-	$\lambda = 100$	0.86	0.80	0.69	**0.50**	**0.50**
2-Step-Bay	SURo	$\delta_u = 0$	0.76	0.68	0.59	0.81	0.89
	DOMo	$\delta_u = 0$	0.76	0.64	0.54	0.72	0.79
	BIGo	$\delta_u = 0$	0.77	0.65	0.57	0.86	0.97
2-Step-PCA	SURo	$\delta_u = 0$	0.81	0.76	0.76	1.02	1.22
	DOMo	$\delta_u = 0$	0.80	0.71	0.65	0.88	1.15
	BIGo	$\delta_u = 0$	0.90	0.76	0.82	1.20	1.54
1-Step	SURo	$\delta_u = 0, \Gamma = 0$	**0.65**	**0.56**	0.50	0.80	0.75
	DOMo	$\delta_u = 0, \Gamma = 0$	0.66	0.57	**0.47**	0.80	0.66
	BIGo	$\delta_u = 0, \Gamma = 0$	0.68	0.58	0.49	0.80	0.65

Notes: The first row shows the average revision, $R_h^{(C,M)}$, of the EU model. The remaining rows show the revision $R_{h,rel}^{(C,M)}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

by far. In particular, the HP filter with smoothing parameter $\lambda = 100$ cuts the revisions sizes by 50 percent for forecast horizons $h = 1$ and $h = 2$.

Of the two-step factor models, the one using Bayesian techniques to extract the factor in the first step is particularly well suited. It strongly reduces the revision sizes, no matter which indicator set is used.

The one-step factor model dominates all competitors for the backcast and the nowcast, reducing the revision sizes by up to 53 percent. It is also very well-suited for forecasting the TFP gap. While the best indicator set for forecasting is the full set (BIGo), the survey set (SURo) follows close behind and is optimal for backcasting.

Italy. For Italy, the EU Commission uses the baseline EU model. We find that relaxing the restriction $\delta_u = 0$ yields considerably smaller revision sizes (row 2 in Table 5). Again, the HP filter improves of the baseline EU model, however this time only for forecasting but not for backcasting and nowcasting the TFP gap.

Of the two-step factor models, again the one using Bayesian techniques to extract the factor in the first step is preferable, especially if combined with the survey indicator set (SURo). It strongly reduces the revision sizes for backcasting and nowcasting the TFP gap but also outperforms the baseline model when forecasting the TFP gap.

The one-step factor model is very well suited for backcasts and nowcasts, for which it reduces the reivsion sizes by 25-40 percent over the baseline EU model. However, it does not outperform the benchmark in terms of forecast revisions, no

Table 5: Revisions to next vintage, Italy

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.28	0.39	0.40	0.96	1.03
	-		0.98	0.86	0.76	0.87	0.85
HP-EST	-		0.95	1.28	1.52	0.84	0.64
HP	-	$\lambda = 6.25$	1.02	1.25	1.51	0.79	**0.60**
	-	$\lambda = 100$	1.57	1.36	1.25	**0.71**	0.68
2-Step-Bay	SURo	$\delta_u = 0$	0.76	0.69	0.62	0.90	0.79
	DOMo	$\delta_u = 0$	0.69	0.67	0.60	0.99	0.87
	BIGo	$\delta_u = 0$	**0.66**	0.63	**0.57**	1.09	0.97
2-Step-PCA	SURo	$\delta_u = 0$	0.69	0.61	0.67	1.01	0.93
	DOMo	$\delta_u = 0$	0.67	0.57	0.67	1.20	1.17
	BIGo	$\delta_u = 0$	**0.66**	**0.53**	0.64	1.31	1.30
1-Step	SURo	$\delta_u = 0$, $\Gamma = 0$	0.70	0.60	0.65	1.12	1.03
	DOMo	$\delta_u = 0$, $\Gamma = 0$	0.75	0.61	0.62	1.20	1.10
	BIGo	$\delta_u = 0$, $\Gamma = 0$	0.75	0.65	0.62	1.14	1.03

Notes: The first row shows the average revision, $R_h^{(C,M)}$, of the EU model. The remaining rows show the revision $R_{h,rel}^{(C,M)}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

matter which indicator set is used. Nevertheless, the survey set (SURo) is still a reasonable choice as it leads to forecast revisions that are comparable to those of the baseline EU model.

Spain. For Spain it is impossible to beat the baseline EU model when it comes to nowcasting or forecasting (see Table 6). In particular, both the HP filters and the two-step factor models are clearly outperformed at almost all horizons and for all indicator sets.

The one-step factor model that uses the survey indicator set (SURo) reduces the revision size of the TFP backcasts considerably over the EU baseline model but exhibits larger revisions for $h \geq 1$. This is surprising because it stands in contrast to the results for the other countries. We argue below that the revision stability of the baseline EU model comes at a cost: It reflects the difficulty to estimate a cycle for Spain that appears sensible at all.

Table 6: Revisions to next vintage, Spain

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.32	0.34	**0.42**	**0.70**	**0.60**
	-		1.66	1.62	1.47	1.33	1.25
HP-EST	-		1.50	1.90	2.79	2.17	1.38
HP	-	$\lambda = 6.25$	0.89	1.50	2.22	1.78	1.16
	-	$\lambda = 100$	1.41	1.65	1.71	1.40	1.20
2-Step-Bay	SURo	$\delta_u = 0$	0.95	1.06	1.49	1.70	1.86
	DOMo	$\delta_u = 0$	1.01	1.09	1.35	1.48	1.44
	BIGo	$\delta_u = 0$	0.96	1.04	1.32	1.43	1.35
2-Step-PCA	SURo	$\delta_u = 0$	0.64	0.75	1.35	1.48	1.41
	DOMo	$\delta_u = 0$	0.98	1.00	1.24	1.43	1.44
	BIGo	$\delta_u = 0$	0.89	0.95	1.24	1.38	1.33
1-Step	SURo	$\delta_u = 0, \Gamma = 0$	**0.54**	**0.74**	1.33	1.33	1.13
	DOMo	$\delta_u = 0, \Gamma = 0$	0.60	0.82	1.41	1.40	1.08
	BIGo	$\delta_u = 0, \Gamma = 0$	0.62	0.86	1.48	1.43	1.06

Notes: The first row shows the average revision, $R_h^{(C,M)}$, of the EU model. The remaining rows show the revision $R_{h,rel}^{(C,M)}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

Netherlands. The baseline EU model exhibits particularly large forecast revision of up to 150 percent of the cycle standard deviation for $h = 1$. The HP filter can reduce these revisions considerably, particularly if the smoothing parameter $\lambda = 100$ is chosen.

The two-step factor models strongly outperform the baseline EU model at all horizons, no matter which indicator set is used and how the factor is extracted from it. The advantage is particularly pronounced for the Bayesian factor estimation which yields reductions of the revision sizes between 25 and almost 50 percent.

The one-step factor model is comparable to the two-step factor models. It is best for backcasting and nowcasting the TFP gap but also well-suited for forecasting. Using the survey indicator set (SURo) is once again a reasonable choice as it yields excellent back- and nowcasts and also very good forecasts.

Table 7: Revisions to next vintage, Netherlands

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.51	0.58	0.84	1.50	1.44
	-		1.03	1.02	0.99	0.94	0.88
HP-EST	-		0.84	1.16	0.87	0.63	0.36
HP	-	$\lambda = 6.25$	0.65	0.87	0.83	0.65	**0.34**
	-	$\lambda = 100$	0.85	1.01	0.77	**0.53**	**0.34**
2-Step-Bay	SURo	$\delta_u = 0$	0.74	0.80	0.87	0.79	0.75
	DOMo	$\delta_u = 0$	0.65	0.73	0.74	0.60	0.53
	BIGo	$\delta_u = 0$	0.65	0.74	0.73	0.60	0.52
2-Step-PCA	SURo	$\delta_u = 0$	0.75	0.83	0.92	0.84	0.87
	DOMo	$\delta_u = 0$	0.62	0.71	0.73	0.62	0.63
	BIGo	$\delta_u = 0$	0.62	0.72	0.74	0.63	0.65
1-Step	SURo	$\delta_u = 0, \Gamma = 0$	0.62	**0.68**	**0.68**	0.82	0.65
	DOMo	$\delta_u = 0, \Gamma = 0$	**0.59**	0.69	0.70	0.68	0.55
	BIGo	$\delta_u = 0, \Gamma = 0$	0.61	0.72	0.71	0.68	0.52

Notes: The first row shows the average revision, $R_h^{(C,M)}$, of the EU model. The remaining rows show the revision $R_{h,rel}^{(C,M)}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

7 Sensitivity analysis

In this Section, we discuss how the results of both the two-step and the one-step factor models presented above depend on the specification choices made. We discuss various parameter restrictions (or relaxations thereof) in Section 7.1, analyze the effect of using the indicators without shifting them forward or backward in Section 7.2, assess whether preselecting only a few indicators improves the results in Section 7.3, and exclude the Covid-19 period from the sample in Section 7.4. We present and discuss the main results in the following and present more detailed results in the Tables in Appendix C.

7.1 Parameter restrictions

While we generally maintain the parameterization of the baseline EU model which is the core also of the extended models, we experiment with restrictions of the AR(1) parameters. δ_u, δ_f, and Γ. Detailed results can be found in Tables 13-18 in Appendix C.1.

7.1.1 The two-step factor models

Let us first consider the two-step factor models. Their baseline specification reported above imposes $\delta_u = 0$ and thus precludes persistent deviations of the cycle from the CUBS indicator. To check whether this restriction is too strong, we relax it and estimate completely unrestricted two-step factor models. It turns out that they exhibit larger revision sizes than our baseline specification for all countries, forecast horizons and indicator sets (except the one-year ahead forecast for Germany). Therefore, imposing $\delta_u = 0$ and thereby mimicking the preferred specification of the EU Commission clearly pays off.

We also analyze whether a more restrictive model can reduce the revision sizes. Specifically, we add to the baseline specification the restriction that $\delta_f = 0$ in the factor equation (8). It entails that the deviations of the factor from the cycle is non-persistent and thus appears quite strong because the factor is extracted from the indicators in the first step in an unsupervised manner, i.e., without any relationship to the TFP. Nevertheless, the results indicate that this restriction reduces the revision sizes of the backcasts and nowcasts of the TFP gap both for the five-country average and for (most of) the individual countries, while the revision sizes of the forecasts mostly increase. These results hold — with few exceptions — for both the Bayesian and the PCA approach and for all three indicator sets.

Overall, our takeaway from these sensitivity checks is that the baseline restriction $\delta_u = 0$ is clearly advantageous, while specifications with the additional restriction $\delta_f = 0$ of the factor equation lead to mixed results that do not uniformly dominate the baseline specification.

7.1.2 The one-step factor models

The baseline specification of the one-step factor model imposes both $\delta_u = 0$ and $\Gamma = 0$. The latter restrictions ensures that the deviations of the individual indicators from the factor is non-persistent and appears to be strong because of the diversity of the indicators. The results indicate that relaxing one of the restrictions or both of them jointly by and large does not reduce the revision sizes.

As a final exercise, we add to the baseline specification the restriction that $\delta_f = 0$ in the factor equation (8). The results suggest that this extra restriction leads to smaller revisions than the baseline specification for the backcasts and, for some countries, also for the nowcasts of the TFP gap. However, forecasting performance is worse, partly by a large margin. Allowing for a persistent error term in the factor equation, as in our baseline, seems to balance the in-sample and out-of-sample fit well.

We conclude that our baseline specification of the one-step factor model dominates alternative specifications in most cases.

7.2 Using unshifted indicators

Furthermore we check whether shifting the indicators forward or backward to maximize their correlation with the cycle as applied in our baseline specification is sensible, since it might lead to in-sample gains at the cost of a deteriorated out-of sample performance due to overfitting. However, the results shown in Tables 19 to 24 in Appendix C.2 suggest that this is not an issue. In fact, using the indicators without a time shift and then applying the one-step or two-step factor models in the baseline specification lower the revision sizes of the backcasts and nowcasts somewhat but worsen the performance of the forecasts.

We conclude that time-shifting the indicators as in our baseline specification is beneficial with respect to forecasting the TFP gap but can be omitted especially for backcasting. We nevertheless prefer the baseline specification because it already strongly outperforms the baseline EU model in terms of backcasting but is less far ahead in terms of forecasting.

7.3 Preselecting indicators

We also considered preselecting indicators as for example suggested by Boivin and Ng (2006), Bai and Ng (2008), Schumacher and Breitung (2008) or Fuentes et al. (2015). To this end, we followed Bai and Ng (2008) and Carstensen et al. (2020) and applied, at each forecast origin, the elastic net of Zou and Hastie (2005) to select the 5, 10, 15 and 20 most relevant indicators. However, we found strong evidence that this approach does not improve revision performance.[5]

The result is not surprising as we have chosen all our indicators because they have typically be proved useful for business cycle analysis and thus are relevant cyclical indicators. Now, repeatedly applying the elastic net for each forecast origin leads to changes in the indicators selected which in turn transmits to larger revision sizes, while the potential benefit of the elastic net to exclude largely irrelevant indicators does not apply in our case.

7.4 Excluding the Covid-19 period

Finally, the results of the revision analysis might also depend on the sample considered, especially in case of huge shocks that potentially induce a structural break. In fact, our revision sample discussed above includes the beginning of the COVID-19 pandemic in 2020 which led to unprecedented slumps in output throughout the EU. Since the European Commission still needs to estimate the TFP gaps of the member states, we check how the models reacted to this event. To this end, we repeat the revision analysis but this time choose 2019 as the last vintage year. We

[5]Detailed results are available upon request.

use our baseline setup discussed in section 6.2 with the only exception that we now normalize our revision measure (15) by the standard deviation of the estimated TFP cycle as of 2019, $s_{2019}^{(C,M)}$.[6].

The result for the five-country average are shown in Table 25 in Appendix C.3. The baseline EU model (row 1) exhibits a revision size for the backcasts and the nowcast of the TFP gap that is about the same compared to the complete vintage analysis discussed in section 6.2. Not surprisingly, however, revisions are smaller for both out-of-sample forecast horizons. The relative performance of the HP filters deteriorates. Nevertheless, it yields smaller revisions than the baseline EU model for two-step forecasts of the TFP gap.

The advantage of both the two-step and the one-step factor models over the baseline EU model with respect to backcasting and nowcasting remains qualitatively but diminishes in size, regardless which indicator set is used. However, they are outperformed in terms of forecasting one year ahead, while especially the one-step factor model is roughly on par with the baseline EU model.

The revision results for the individual countries is displayed in Tables 26 to 30. They reveal that the one-step factor model still outperforms the baseline EU model with respect to backcasting and nowcasting the TFP gap for Germany, France, Italy, and the Netherlands. The exception is Spain which now — in contrast to the full vintage sample — rather favors the baseline EU model. If it comes to forecasting, the one-step factor model yields smaller revisions than the baseline EU model for Germany, France, and the Netherlands, while it is slightly worse for Italy and severely worse for Spain. In fact, once we exclude Spain, the one-step factor model still outperforms the baseline EU model on average over Germany, France, Italy, and the Netherlands.

Overall, comparing the vintage samples with and without the COVID-19 period suggests that the factor models are more flexible than the baseline EU model to accommodate a large shock which is why they increase their advantage in the full sample. However, Spain appears to be an outlier as, e.g., the one-step factor model exhibits more than 100 percent larger revisions than the baseline EU model. We suspect that this has to do with the specific business cycle history of Spain as we discuss in the next section.

[6]Note however that the standard deviation of the cycle estimate of 2019 is very similar to the one of 2021 in almost all specifications. Therefore, using $s_{2021}^{(C,M)}$ instead renders our results qualitatively unchanged.

8 Plausibility of the estimated trend and cycle components

So far we measured and discussed only the revision RMSE as an indicator of model performance. While cycle and trend are unknown, some properties of trend and cycle may nevertheless be known from external sources, for example recession periods, and estimation results should be consistent over time, i.e., over different vintages. In the following we will discuss and compare our factor model to the outcomes of the EU baseline model in these regards.

8.1 Decomposition of trend and cycle

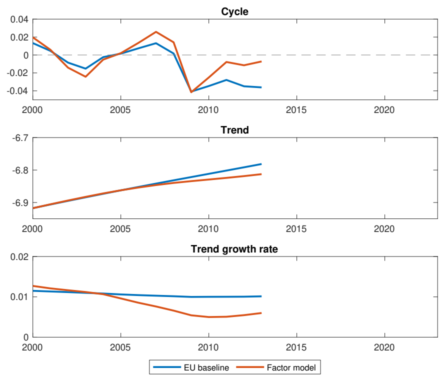
One key difference of all factor models and HP filters compared to the baseline EU model is that the former tend to predict a more optimistic state of the TFP cycle at the last 5 to 10 years of a sample, including the two-year ahead TFP gap forecasts. Conversely, the trend growth rate of the baseline EU model tends to be higher. Consequently, the baseline EU model attributes low TFP growth at the sample end more likely to the cycle than to the trend, while the factor models' estimate of the trend reacts faster.

In Figure (1) we illustrate this behavior by comparing a one-step factor model (SURo, baseline specification with $\delta_u = 0$ and $\Gamma = 0$, left panels) and the HP-filter with $\lambda = 100$ (right panels) to the EU-model for the vintages 2011 (upper panels) and 2021 (lower panels) for the Netherlands as an example. In each subfigure we report the estimated cycle (to facilitate the comparison, the cycles of the factor model and HP-filter are rescaled to have the same variance as the baseline EU model), the trend as well as the trend growth rate from 2000 to the respective sample end including the two-year ahead forecasts. The horizontal dashed grey line corresponds to a TFP gap of zero.

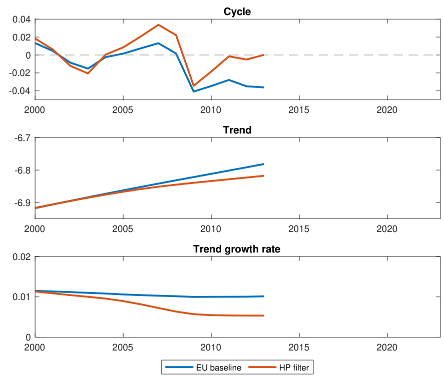
For vintage 2011 we observe a large discrepancy between the cycle estimates of the baseline EU model (blue line) and the factor model or HP-filter (red line), especially for the years 2010-2013. The baseline EU model predicts a substantially negative cycle, while the other two models predict a lower trend growth rate. Therefore, trend estimates of the competing models are strikingly different.

By comparing the results to the 2021 vintage shown in the lower panels, we find that the trend growth rate estimated by the baseline EU model eventually also drops considerably. In fact, in the 2021 vintage, it is almost identical to the trend growth rates of the factor model and the HP filter. Since this slowdown of the trend growth rate is pervasive across the EU, it leads to relatively large revisions of the baseline EU model, particularly for France, Italy, and the Netherlands. In fact, the factor models signal this development — which is probably a consequence of the Great Recession — already in 2011, while the baseline EU model capture it much later, beginning with vintage 2015.

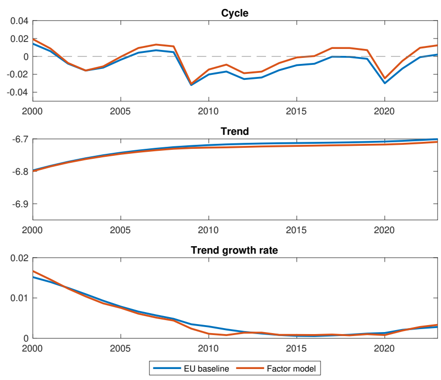
Our interpretation is that the CUBS indicator alone does not sufficiently strongly indicate in early vintages that it is not the cycle but the trend that needs to react. In contrast, our factor models include more information and attribute a larger weight to it so that they react much earlier. For the period after the Great Recession described here, this would have had strong policy implications. If the EU Commission would have found a lower trend growth and a less negative output gap, government would perhaps urged much more towards structural policies than to business cycle stabilization.



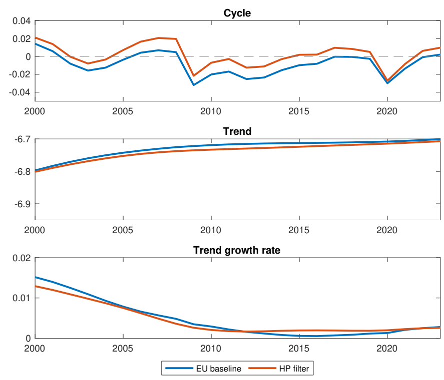
(a) SURo, vintage 2011



(b) $\lambda = 100$, vintage 2011



(c) SURo, vintage 2021



(d) $\lambda = 100$, vintage 2021

Figure 1: Cycle and trend estimates of factor model (left) and HP filter (right) compared to EU approach for the Netherlands

8.2 Cycle length

In Figure (2) we plot the estimated cycle of each country applying both the EU baseline model (blue) and our one-step factor model (SURo, baseline specification with $\delta_u = 0$ and $\Gamma = 0$) using the information available in 2021. The horizontal dashed grey line corresponds to a TFP gap of zero, and the shaded areas indicate recessions as defined by the OECD.[7]

For Germany both approaches lead to a fairly similar cycle even if again the TFP gap of the factor model is slightly more positive towards the sample end. Much larger differences at the sample end occur in the case of France, the Netherlands as well as Italy. For the latter we note that the periodicity of the cycle using the EU commission approach is larger (12 years on average) compared to our factor extension (9 years on average). Especially after 2008, the EU model always predicts a negative cycle.

The most striking difference however can be observed for Spain for which the estimated cycles differ substantially. The periodicity of the EU baseline is much larger (15 years) than that of the factor model (10 years on average). In particular, the EU model predicts a positive TFP gap for all twenty years from 1985 to 2004 and a consistently negative TFP gap thereafter. This roughly coincides with the long housing boom and subsequent bust in Spain. Our interpretation is that the CUBS indicator did not strongly enough signal that the bust was to a relevant extent structural — after all, many resources of the construction sector became unproductive — as opposed to only cyclical. Again using a broader information base appears to deliver a more realistic picture of the business cycle stance.

[7] FRED data, derived from the OECD Composite Leading Indicators: Reference Turning Points and Component Series.

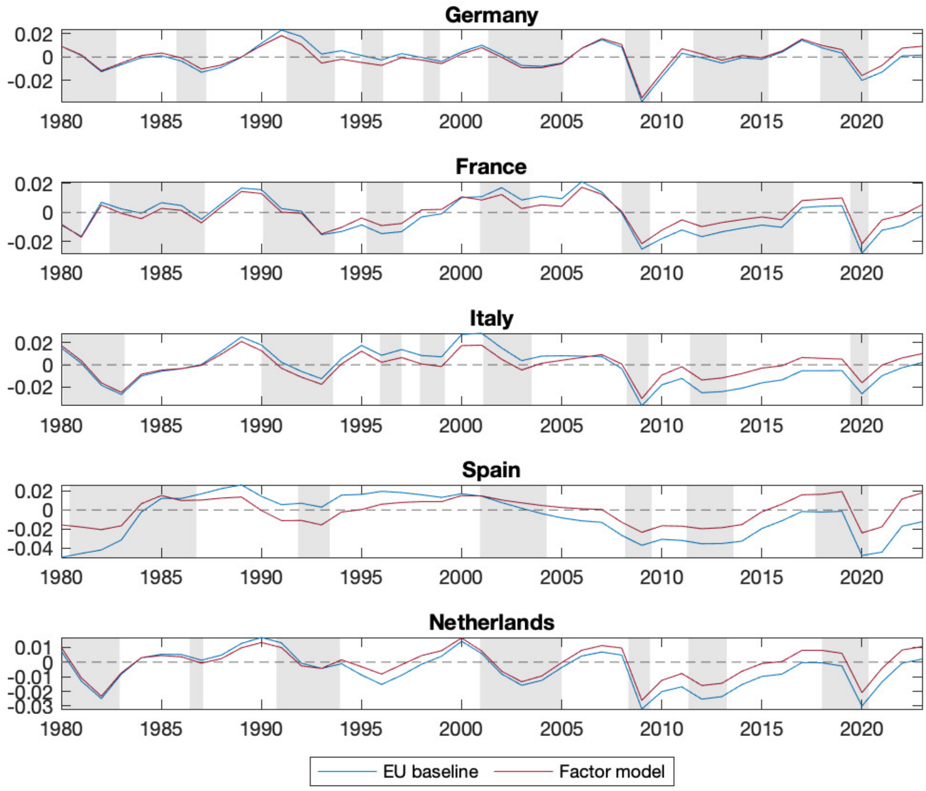


Figure 2: Estimated cycles for vintage 2021

9 Conclusion

In this paper we showed that estimating and forecasting the TFP gap of the largest five EU countries — and presumably also for the smaller countries — can benefit considerably from including a set of business cycle indicators via a factor model extension in the otherwise unchanged unobserved components model of the EU Commission. Based on a vintage of data as of October of a year T, the advantage is particularly pronounced for the last two in-sample years $T-2$ and $T-1$, for which the year-to-year revisions can be reduced by up to 30 percent, and for the nowcast year T, for which the reduction is around 20 percent. Improvements for the two out-of-sample years $T+1$ and $T+2$ are also feasible but quantitatively smaller (around 10 percent) and dependent on the country specifics.

We also argued that the EU model produces implausible results for Spain. Especially for this country, its estimate of the trend growth rate is very stable over vintages leading to small revisions and thereby outperforming our factor models.

However, its cycle length is implausibly long indicating that Spain went trough only one upswing and one recession since 1985. A factor constructed from a set of business cycle indicators indicates that this is not a complete description of the Spanish business cycle history which is why the factor models yield much smaller cycle lengths.

In a sensitivity analysis we showed that the results of the baseline specification of our factor models yield very good results compared to relevant alternatives. Overall, the results do not depend overly on the specification of the factor model or the indicator set chosen — it appears sufficient to include "enough" relevant information.

It remains to be analyzed which role the prior choices made by the EU commission play. In an attempt to make our factor models as comparable as possible to the baseline EU model, we have refrained from optimizing on this margin.

References

Aastveit, K. A. and Trovik, T. (2014). Estimating the output gap in real time: A factor model approach. *The Quarterly Review of Economics and Finance*, 54(2):180–193.

Bai, J. and Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2):304–317.

Barigozzi, M. and Luciani, M. (2019). Measuring the Output Gap Using Large Datasets. Technical report, Available at SSRN.

Basistha, A. and Startz, R. (2008). Measuring the NAIRU with reduced uncertainty: a multiple-indicator common-cycle approach. *The Review of Economics and Statistics*, 90(4):805–811.

Blagrave, P., Garcia-Saltos, M. R., Laxton, M. D., and Zhang, F. (2015). *A simple multivariate filter for estimating potential output.* Number 15-79. International Monetary Fund.

Boivin, J. and Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132:169–194.

Carstensen, K., Heinrich, M., Reif, M., and Wolters, M. H. (2020). Predicting Ordinary and Severe Recessions with a Three-State Markov-Switching Dynamic Factor Model — An Application to the German Business Cycle. *International Journal of Forecasting*, forthcoming.

D'Auria, F., Denis, C., Havik, K., Morrow, K. M., Planas, C., Raciborski, R., Röger, W., and Rossi, A. (2010). The production function methodology for calculating potential growth rates and output gaps. Technical Report Economic Papers 420, European Economy.

Dovern, J. and Zuber, C. (2019). Revisions of potential output estimates in the EU after the Great Recession. *CESifo Working Paper*, (7681).

Fleischman, C. and Roberts, J. (2011). From many series, one cycle: improved estimates of the business cycle from a multivariate unobserved components model. Finance and Economics Discussion Series 2011-46, Board of Governors of the Federal Reserve System (U.S.).

Fuentes, J., Poncela, P., and Rodríguez, J. (2015). Sparse partial least squares in time series for macroeconomic forecasting. *Journal of Applied Econometrics*, 30(4):576–595.

Geweke, J. (1992). *Bayesian statistics*, volume 4, chapter Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, pages 169–193. Oxford University Press.

González-Astudillo, M. (2019a). An output gap measure for the euro area: Exploiting country-level and cross-sectional data heterogeneity. *European Economic Review*, 120(C).

González-Astudillo, M. (2019b). Estimating the U.S. output gap with state-level data. *Journal of Applied Econometrics*, 34(5):795–810.

Graff, M. and Sturm, J.-E. (2010). The information content of capacity utilisation rates for output gap estimates. KOF Working papers 10-269, KOF Swiss Economic Institute, ETH Zurich.

Harvey, A. and Trimbur, T. (2008). Trend estimation and the Hodrick-Prescott filter. *Journal of the Japanese Statistical Society*, 380(1):41–49.

Havik, K., Morrow, K. M., Orlandi, F., Planas, C., Raciborski, R., Röger, W., Rossi, A., Thum-Thysen, A., and Vandermeulen, V. (2014). The Production Function Methodology for Calculating Potential Growth Rates & Output Gaps. Technical report, EU Commission Economic Papers 535.

Huber, F., Pfarrhofer, M., and Piribauer, P. (2020). A multi-country dynamic factor model with stochastic volatility for euro area business cycle analysis. *Journal of Forecasting*, n/a(n/a).

Jarociński, M. and Lenza, M. (2018). An inflation-predicting measure of the output gap in the euro area. *Journal of Money, Credit and Banking*, 50(6):1189–1224.

Kuttner, K. N. (1994). Estimating potential output as a latent variable. *Journal of Business & Economic Statistics*, 12(3):361–368.

Marcellino, M. and Musso, A. (2011). The reliability of real-time estimates of the euro area output gap. *Economic Modelling*, 28(4):1842 – 1856.

Melolinna, M. and Toth, M. (2019). Output gaps, inflation and financial cycles in the uk. *Empirical Economics*, 56(3):1039–1070.

Mise, E., Kim, T.-H., and Newbold, P. (2005). On suboptimality of the hodrick–prescott filter at time series endpoints. *Journal of Macroeconomics*, 27(1):53 – 67.

Orphanides, A. and van Norden, S. (2002). The unreliability of output-gap estimates in real time. *The Review of Economics and Statistics*, 84(4):569–583.

Planas, C., Roeger, W., and Rossi, A. (2013). The information content of capacity utilization for detrending total factor productivity. *Journal of Economic Dynamics and Control*, 37(3):577 – 590.

Planas, C., Rossi, A., and Fiorentini, G. (2008). Bayesian analysis of the output gap. *Journal of Business & Economic Statistics*, 26:18–32.

Ravn, M. O. and Uhlig, H. (2002). On adjusting the Hodrick-Prescott filter for the frequency of observations. *Review of Economics and Statistics*, 84(2):371–380.

Schumacher, C. and Breitung, J. (2008). Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting*, 24(3):386–398.

Stock, J. H. and Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.

Turner, D., Cavalleri, M. C., Guillemette, Y., Kopoin, A., Ollivaud, P., and Rusticelli, E. (2016). An investigation into improving the real-time reliability of OECD output gap estimates. OECD Economics Department Working Papers 1294, OECD Publishing.

Valle e Azevedo, J., Koopman, S. J., and Rua, A. (2006). Tracking the business cycle of the euro area: A multivariate model-based bandpass filter. *Journal of Business & Economic Statistics*, 24:278–290.

Weiske, S. (2018). Indicator-based estimates of the output gap in the euro area. Technical Report Working Paper No. 12/2018, Sachverständigenrat zur Begutachtung der Gesamtwirtschaftlichen Entwicklung.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.

A Estimation of factor models

A.1 Factor model estimated by principal components

To estimate a single business cycle factor from a set of k standardized indicators by principal component analysis, we apply the EM algorithm of Stock and Watson (2002) to cope with missing data at the beginning of the sample. It is initialized by inserting zeros where observations are missing in the $T \times k$ data matrix $\mathbf{X}$. In the first step, the factor and its loading vector are estimated by standard principal component analysis. In the second step, the missing observations in $\mathbf{X}$ are updated by setting them to factor times loading coefficient. We iterate these two steps until convergence to obtain a final estimate of the factor.

A.2 Factor model estimated with Bayesian techniques

A.2.1 The model

The purpose of the model is to extract a single factor, f_t, from a $k \times 1$ vector of observed business cycle indicators, x_t, which are standardized. The measurement equation is

$$x_t = \Lambda f_t + v_t, \quad t = 1, \dots, T, \quad v_t \overset{\text{iid}}{\sim} N(0, \Sigma_v), \tag{16}$$

where Σ_v is diagonal with diagonal elements $\sigma_1^2, \dots, \sigma_k^2$. The factor is assumed to follow an AR(2) process,

$$f_t = 2\varphi \cos(2\pi/\varrho) f_{t-1} - \varphi^2 f_{t-2} + w_t, \quad t = 3, \dots, T, \quad w_t \overset{\text{iid}}{\sim} N(0, 1), \tag{17}$$

with the first two observations being drawn from the unconditional distribution. The parameterization follows Havik et al. (2014). Assuming $0 < \varrho < 1$ and $2 < \varphi < 32$ ensures stationarity.

Defining $\Lambda = (\lambda_1, \dots, \lambda_k)'$, identification of the factor is achieved by setting the first element to one, i.e., $\lambda_1 = 1$.

The parameters of interest are defined as $\theta = (\lambda_2, \dots, \lambda_k, \sigma_1^2, \dots, \sigma_k^2, \varrho, \varphi)'$.

A.2.2 Priors

Conditional on the factor, (16) defines a set of k independent regression equations of which the first one is special because its mean parameter is restricted to one. To obtain closed-form solutions, we use conjugate priors. For the first regression equation we thus use an inverse Gamma[8] prior,

$$\sigma_1^2 \sim iG(\underline{s}_1^{-2}, \underline{\nu}_1), \tag{18}$$

[8]We parameterize the pdf of the inverse Gamma distribution, $y \sim iG(s^{-2}, \nu)$, as $p(y) = c_{iG} y^{-\nu/2-1} \exp(-0.5\nu s^2/y)$, where c_{iG} is an integration constant.

and for the remaining equations we use Normal-inverse Gamma priors,

$$\lambda_i|\sigma_i^2 \sim N(\underline{\lambda}_i, \sigma_i^2\underline{V}), \quad \sigma_i^2 \sim iG(\underline{s}_i^{-2}, \underline{\nu}_i). \tag{19}$$

For the AR(2) parameters we follow Havik et al. (2014) and use

$$\varphi \sim B(\underline{\alpha}_1, \underline{\beta}_1, 0, 1) \tag{20}$$

and

$$\varrho \sim B(\underline{\alpha}_2, \underline{\beta}_2, 2, 32), \tag{21}$$

where $B(\alpha, \beta, a, b)$ is the Beta distribution with parameters α and β rescaled to the interval (a, b).

The implied prior for f_1, f_2 is

$$[f_1, f_2]'|\varrho, \varphi \sim N(0, \Upsilon(\varrho, \varphi)), \tag{22}$$

where $\Upsilon(\varrho, \varphi)$ is the unconditional 2×2 variance matrix of an AR(2) process and thus a function of ϱ and φ. For $f_3, \ldots, f_T$ the prior can be written as

$$f_t|f_{t-1}, f_{t-2}, \varrho, \varphi \sim N(2\varrho \cos(2\pi/\varphi) f_{t-1} - \varphi^2 f_{t-2}, 1). \tag{23}$$

A.2.3 Posterior simulation

The structure of the model suggests to apply a Metropolis-within-Gibbs simulation of the posterior distribution. We briefly comment on each Gibbs step.

Sampling unobserved components. To sample factor and the missing observations in the indicators conditional on the parameters, we directly draw from the conditionally normal posterior distribution.

Sampling φ and ϱ. Conditional on the factor, f_t, these are the coefficients of a stationary AR(2) model with the first two observations drawn from their unconditional distribution. To sample ϱ we apply a MH step with the asymptotic normal distribution of the maximum likelihood estimator truncated to the unit interval as proposal pdf. To sample φ we again apply a MH step with the asymptotic normal distribution of a crude maximum likelihood estimator truncated to the prespecified range as proposal pdf.

Sampling $\lambda_2, \ldots, \lambda_k$ and $\sigma_1^2, \ldots, \sigma_k^2$. Conditional on everything else the likelihood is the standard likelihood of k independent regressions. As prior we choose the natural conjugate Normal-Inverse-Gamma prior. We have to make the following case distinction:

- Due to identification we set $\lambda_1 = 1$ which implies an inverse-Gamma posterior for σ_1^2.
- For $i = 2, \ldots, k$ the posterior of λ_i, σ_i^2 is also Normal-Inverse-Gamma.

B Data

Table 8: Indicators for Germany

Indicator	Freq.	Provider	Source	Mnemonic	Tr.	Qu.
Hard indicators						
Total industrial production	m	FRED	OECD MEI	DEUPROINDMISMEI	3	3
Manufacturing production	m	FRED	OECD MEI	DEUPROMANMISMEI	3	3
Investment goods production	m	FRED	OECD MEI	DEUPRMNVG01IXOBSAM	3	3
Construction	m	FRED	OECD MEI	DEUPROCONMISMEI	3	3
Retail sales	m	FRED	OECD MEI	DEUSARTMISMEI	3	3
Passenger car registrations	m	FRED	OECD MEI	DEUSLRTCR03IXOBSAM	3	3
Building permits	m	FRED	OECD MEI	DEUODCNPI03MLSAM	3	3
Unemployment rate	m	FRED	OECD MEI	LMUNRRTTDEM156S	2	3
Number of employees	q	FRED	OECD MEI	LFEMTTTTDEQ647S	3	3
Hours worked	q	Eurostat	Eurostat		3	3
CPI	m	FRED	OECD MEI	DEUCPIALLMINMEI	3	3
PPI	m	FRED	OECD MEI	DEUPPDMMINMEI	3	3
House prices	q	FRED	BIS	QDEN628BIS	3	3
GDP	q	FRED	OECD MEI	NAEXKP01DEQ661S	3	2
Income	q	FRED	OECD QNA	DEUCOMPQDSNAQ	3	2
Hourly earnings	q	FRED	OECD MEI	DEUHOUREAQISMEI	3	2
Private consumption	q	FRED	OECD MEI	NAEXKP02DEQ661S	3	2
Interest rate spread	m	FRED	OECD	DEULOCOSIORSTM	1	3
Stock market index	m	FRED	OECD MEI	SPASTT01DEM661N	3	3
REER	m	FRED	BIS	RNDEBIS	3	3
Survey indicators						
Capacity utilization	q	DG EcFin	DG EcFin	INDU.DE.TOT.13.QPS.Q	1	3
New Orders, total	q	DG EcFin	DG EcFin	INDU.DE.TOT.11.BS.Q	1	3
New Orders, Intermediate goods	q	DG EcFin	DG EcFin	INDU.DE.INTM.11.BS.Q	1	3
New Orders, Investment goods	q	DG EcFin	DG EcFin	INDU.DE.INVE.11.BS.Q	1	3
Assessment of order-book levels	m	DG EcFin	DG EcFin	INDU.DE.TOT.2.BS.M	1	3
Production expectations	m	DG EcFin	DG EcFin	INDU.DE.TOT.5.BS.M	1	3
Industrial confidence	m	DG EcFin	DG EcFin	DE.INDU	1	3
Services confidence	m	DG EcFin	DG EcFin	DE.SERV	1	3
Consumers confidence	m	DG EcFin	DG EcFin	DE.CONS	1	3
Retail confidence	m	DG EcFin	DG EcFin	DE.RETA	1	3
Construction confidence	m	DG EcFin	DG EcFin	DE.BUIL	1	3
Economic sentiment	m	DG EcFin	DG EcFin	DE.ESI	1	3
International indicators						
Economic sentiment, Euro Area	m	DG EcFin	DG EcFin	EA.ESI	1	3
Unemployment rate, Euro Area	m	FRED	OECD MEI	LRHUTTTTEZM156S	1	3
Eurocoin	m	CEPR	CEPR	Eurocoin	1	3
European Stock market index	m	FRED	OECD	SPASTT01EZM661N	3	3
Oil price / Brent	m	FRED	IMF	POILBREUSDM	3	3
Employment, US	m	FRED	BLS	PAYEMS	3	3
GDP, US	q	FRED	BEA	GDPC1	3	2
Real personal income, US	m	FRED	BEA	RPI	3	3
Federal Funds Rate, US	m	FRED	Fed	FEDFUNDS	2	3

Notes: m and q denote monthly and quarterly frequency. Tr denotes the stationarity-generating transformations: 1 = level, 2 = difference, 3= log difference. Qu denotes the last quarter of a year available at the end of October of a year.

Table 9: Indicators for France

Indicator	Freq.	Provider	Source	Mnemonic	Tr.	Qu.
Hard indicators						
otal industrial production	m	FRED	OECD MEI	FRAPROINDMISMEI	3	3
Manufacturing production	m	FRED	OECD MEI	FRAPROMANMISMEI	3	3
Intermediate goods production	m	FRED	OECD MEI	FRAPRMNIG01IXOBSAM	3	3
Investment goods production	m	FRED	OECD MEI	FRAPRMNVG01IXOBSAM	3	3
Construction	m	FRED	OECD MEI	FRAPROCONMISMEI	3	3
Retail sales	m	FRED	OECD MEI	FRASARTMISMEI	3	3
Passenger car registrations	m	FRED	OECD MEI	FRASLRTCR03IXOBSAM	3	3
Building permits	m	FRED	OECD MEI	FRAPERMITMISMEI	3	3
Unemployment rate	m	FRED	OECD MEI	LRHUTTTTFRM156S	2	3
Number of employees	q	FRED	OECD MEI	LFEMTTTTFRQ647S	3	3
Hours worked	q	Eurostat	Eurostat		3	3
CPI	m	FRED	OECD MEI	FRACPIALLMINMEI	3	3
PPI	m	FRED	OECD MEI	FRAPPDMMINMEI	3	3
House prices	q	FRED	BIS	QFRN628BIS	3	3
GDP	q	FRED	OECD MEI	NAEXKP01FRQ661S	3	2
Income	q	FRED	OECD QNA	FRACOMPQDSNAQ	3	2
Hourly earnings	q	FRED	OECD MEI	LCEAMN01FRQ661S	3	2
Private consumption	q	FRED	OECD MEI	NAEXKP02FRQ661S	3	2
Interest rate spread	m	OECD	OECD		1	3
Stock market index	m	FRED	OECD MEI	SPASTT01FRM661N	3	3
REER	m	FRED	BIS	RNFRBIS	3	3
Survey indicators						
Capacity utilization	q	DG EcFin	DG EcFin	INDU.FR.TOT.13.QPS.Q	1	3
New Orders, total	q	DG EcFin	DG EcFin	INDU.FR.TOT.11.BS.Q	1	3
New Orders, Intermediate goods	q	DG EcFin	DG EcFin	INDU.FR.INTM.11.BS.Q	1	3
New Orders, Investment goods	q	DG EcFin	DG EcFin	INDU.FR.INVE.11.BS.Q	1	3
Assessment of order-book levels	m	DG EcFin	DG EcFin	INDU.FR.TOT.2.BS.M	1	3
Production expectations	m	DG EcFin	DG EcFin	INDU.FR.TOT.5.BS.M	1	3
Industrial confidence	m	DG EcFin	DG EcFin	FR.INDU	1	3
Services confidence	m	DG EcFin	DG EcFin	FR.SERV	1	3
Consumers confidence	m	DG EcFin	DG EcFin	FR.CONS	1	3
Retail confidence	m	DG EcFin	DG EcFin	FR.RETA	1	3
Construction confidence	m	DG EcFin	DG EcFin	FR.BUIL	1	3
Economic sentiment (ESI)	m	DG EcFin	DG EcFin	FR.ESI	1	3

Notes: m and q denote monthly and quarterly frequency. Tr denotes the stationarity-generating transformations: 1 = level, 2 = difference, 3= log difference. Qu denotes the last quarter of a year available at the end of October of a year. The international indicators are the same as shown for Germany and thus not repeated here.

Table 10: Indicators for Italy

Indicator	Freq.	Provider	Source	Mnemonic	Tr.	Qu.
Hard indicators						
Total industrial production	m	FRED	OECD MEI	ITAPROINDMISMEI	3	3
Manufacturing production	m	FRED	OECD MEI	ITAPROMANMISMEI	3	3
Intermediate goods production	m	FRED	OECD MEI	ITAPRMNIG01IXOBSAM	3	3
Investment goods production	m	FRED	OECD MEI	ITAPRMNVG01IXOBSAM	3	3
Construction	m	FRED	OECD MEI	ITAPRCNTO01IXOBSAM	3	3
Retail sales	m	FRED	OECD MEI	ITASARTMISMEI	3	3
Passenger car registrations	m	FRED	OECD MEI	ITASLRTCR03IXOBSAM	3	3
Building permits	q	ISTAT	ISTAT		3	3
Unemployment rate	m	FRED	OECD MEI	LRHUTTTTITM156S	2	3
Number of employees	q	FRED	OECD MEI	LFEMTTTTITQ647S	3	3
Hours worked	q	Eurostat	Eurostat		3	3
CPI	m	FRED	OECD MEI	ITACPIALLMINMEI	3	3
PPI	m	FRED	OECD MEI	ITAPPDMMINMEI	3	3
House prices	q	FRED	BIS	QITN628BIS	3	3
GDP	q	FRED	OECD MEI	LORSGPORITQ661S	3	2
Income	q	FRED	OECD QNA	ITACOMPQDSNAQ	3	2
Hourly earnings	q	FRED	OECD MEI	LCEAMN01ITQ661S	3	2
Private consumption	q	FRED	OECD MEI	NAEXKP02ITQ661S	3	2
Interest rate spread	m	FRED	OECD		1	3
Stock market index	m	FRED	OECD MEI	SPASTT01ITM661N	3	3
REER	m	FRED	BIS	RNITBIS	3	3
Survey indicators						
Capacity utilization	q	DG EcFin	DG EcFin	INDU.IT.TOT.13.QPS.Q	1	3
New Orders, total	q	DG EcFin	DG EcFin	INDU.IT.TOT.11.BS.Q	1	3
New Orders, Intermediate goods	q	DG EcFin	DG EcFin	INDU.IT.INTM.11.BS.Q	1	3
New Orders, Investment goods	q	DG EcFin	DG EcFin	INDU.IT.INVE.11.BS.Q	1	3
Assessment of order-book levels	m	DG EcFin	DG EcFin	INDU.IT.TOT.2.BS.M	1	3
Production expectations	m	DG EcFin	DG EcFin	INDU.IT.TOT.5.BS.M	1	3
Industrial confidence	m	DG EcFin	DG EcFin	IT.INDU	1	3
Services confidence	m	DG EcFin	DG EcFin	IT.SERV	1	3
Consumers confidence	m	DG EcFin	DG EcFin	IT.CONS	1	3
Retail confidence	m	DG EcFin	DG EcFin	IT.RETA	1	3
Construction confidence	m	DG EcFin	DG EcFin	IT.BUIL	1	3
Economic sentiment (ESI)	m	DG EcFin	DG EcFin	IT.ESI	1	3

Notes: m and q denote monthly and quarterly frequency. Tr denotes the stationarity-generating transformations: 1 = level, 2 = difference, 3= log difference. Qu denotes the last quarter of a year available at the end of October of a year. The international indicators are the same as shown for Germany and thus not repeated here.

Table 11: Indicators for Spain

Indicator	Freq.	Provider	Source	Mnemonic	Tr.	Qu.
Hard indicators						
Total industrial production	m	FRED	OECD MEI	ESPPROINDMISMEI	3	3
Manufacturing production	m	FRED	OECD MEI	ESPPROMANMISMEI	3	3
Intermediate goods production	m	FRED	OECD MEI	ESPPRMNIG01IXOBSAM	3	3
Investment goods production	m	FRED	OECD MEI	ESPPRMNVG01IXOBSAM	3	3
Construction	m	FRED	OECD MEI	ESPPROCONMISMEI	3	3
Retail sales	m	FRED	OECD MEI	ESPSARTMISMEI	3	3
Passenger car registrations	m	FRED	OECD MEI	ESPSLRTCR03MLSAM	3	3
Building permits	m	FRED	OECD MEI	ESPPERMITMISMEI	3	3
Unemployment rate	m	FRED	OECD MEI	LRHUTTTTESM156S	2	3
Number of employees	q	FRED	OECD MEI	LFEMTTTTESQ647S	3	3
Hours worked	q	Eurostat	Eurostat		3	3
CPI	m	FRED	OECD MEI	ESPCPIALLMINMEI	3	3
PPI	m	FRED	OECD MEI	ESPPPDMMINMEI	3	3
House prices	q	FRED	BIS	QESN628BIS	3	3
GDP	q	FRED	OECD MEI	LORSGPORESQ661S	3	2
Income	q	Eurostat	Eurostat		3	2
Hourly earnings	q	FRED	OECD MEI	LCEAMN01ESQ661S	3	2
Private consumption	q	FRED	OECD MEI	NAEXKP02ESQ661S	3	2
Interest rate spread	m	OECD	OECD		1	3
Stock market index	m	FRED	OECD MEI	SPASTT01ESM661N	3	3
REER	m	FRED	BIS	RNESBIS	3	3
Survey indicators						
Capacity utilization	q	DG EcFin	DG EcFin	INDU.ES.TOT.13.QPS.Q	1	3
New Orders, total	q	DG EcFin	DG EcFin	INDU.ES.TOT.11.BS.Q	1	3
New Orders, Intermediate goods	q	DG EcFin	DG EcFin	INDU.ES.INTM.11.BS.Q	1	3
New Orders, Investment goods	q	DG EcFin	DG EcFin	INDU.ES.INVE.11.BS.Q	1	3
Assessment of order-book levels	m	DG EcFin	DG EcFin	INDU.ES.TOT.2.BS.M	1	3
Production expectations	m	DG EcFin	DG EcFin	INDU.ES.TOT.5.BS.M	1	3
Industrial confidence	m	DG EcFin	DG EcFin	ES.INDU	1	3
Services confidence	m	DG EcFin	DG EcFin	ES.SERV	1	3
Consumers confidence	m	DG EcFin	DG EcFin	ES.CONS	1	3
Retail confidence	m	DG EcFin	DG EcFin	ES.RETA	1	3
Construction confidence	m	DG EcFin	DG EcFin	ES.BUIL	1	3
Economic sentiment (ESI)	m	DG EcFin	DG EcFin	ES.ESI	1	3

Notes: m and q denote monthly and quarterly frequency. Tr denotes the stationarity-generating transformations: 1 = level, 2 = difference, 3= log difference. Qu denotes the last quarter of a year available at the end of October of a year. The international indicators are the same as shown for Germany and thus not repeated here.

Table 12: Indicators for Netherlands

Indicator	Freq.	Provider	Source	Mnemonic	Tr.	Qu.
Hard indicators						
Total industrial production	m	FRED	OECD MEI	NLDPROINDMISMEI	3	3
Manufacturing production	m	FRED	OECD MEI	NLDPROMANMISMEI	3	3
Intermediate goods production	m	FRED	OECD MEI	NLDPRMNIG01IXOBSAM	3	3
Investment goods production	m	FRED	OECD MEI	NLDPRMNVG01IXOBSAM	3	3
Construction	m	FRED	OECD MEI	NLDPROCONMISMEI	3	3
Retail sales	m	FRED	OECD MEI	NLDSARTMISMEI	3	3
Passenger car registrations	m	FRED	OECD MEI	NLDSLRTCR03MLSAM	3	3
Building permits	m	FRED	OECD MEI	NLDPERMITMISMEI	3	3
Unemployment rate	m	FRED	OECD MEI	LRHUTTTTNLM156S	2	3
Number of employees	q	FRED	OECD MEI	LFEMTTTTNLQ647S	3	3
Hours worked	q	Eurostat	Eurostat		3	3
CPI	m	FRED	OECD MEI	NLDCPIALLMINMEI	3	3
PPI	m	FRED	OECD MEI	NLDPPDMMINMEI	3	3
House prices	q	FRED	BIS	QNLN628BIS	3	3
GDP	q	FRED	OECD MEI	LORSGPORNLQ661S	3	2
Income	q	Eurostat	Eurostat		3	2
Hourly earnings	q	FRED	OECD MEI	LCEAMN01NLQ661S	3	2
Private consumption	q	FRED	OECD MEI	NAEXKP02NLQ661S	3	2
Interest rate spread	m	OECD	OECD		1	3
Stock market index	m	FRED	OECD MEI	SPASTT01NLM661N	3	3
REER	m	FRED	BIS	RNNLBIS	3	3
Survey indicators						
Capacity utilization	q	DG EcFin	DG EcFin	INDU.NL.TOT.13.QPS.Q	1	3
New Orders, total	q	DG EcFin	DG EcFin	INDU.NL.TOT.11.BS.Q	1	3
New Orders, Intermediate goods	q	DG EcFin	DG EcFin	INDU.NL.INTM.11.BS.Q	1	3
New Orders, Investment goods	q	DG EcFin	DG EcFin	INDU.NL.INVE.11.BS.Q	1	3
Assessment of order-book levels	m	DG EcFin	DG EcFin	INDU.NL.TOT.2.BS.M	1	3
Production expectations	m	DG EcFin	DG EcFin	INDU.NL.TOT.5.BS.M	1	3
Industrial confidence	m	DG EcFin	DG EcFin	NL.INDU	1	3
Services confidence	m	DG EcFin	DG EcFin	NL.SERV	1	3
Consumers confidence	m	DG EcFin	DG EcFin	NL.CONS	1	3
Retail confidence	m	DG EcFin	DG EcFin	NL.RETA	1	3
Construction confidence	m	DG EcFin	DG EcFin	NL.BUIL	1	3
Economic sentiment (ESI)	m	DG EcFin	DG EcFin	NL.ESI	1	3

Notes: m and q denote monthly and quarterly frequency. Tr denotes the stationarity-generating transformations: 1 = level, 2 = difference, 3= log difference. Qu denotes the last quarter of a year available at the end of October of a year. The international indicators are the same as shown for Germany and thus not repeated here.

C Results of the sensitivity analysis

C.1 Alternative parameter restrictions

Table 13: Factor models with alternative parameter restrictions, revisions to next vintage, average over all five countries

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SURo		0.95	0.96	1.02	1.08	1.12
	DOMo		0.89	0.88	0.90	1.00	0.97
	BIGo		0.86	0.85	0.88	1.04	1.01
	SURo	$\delta_u=0$, $\delta_f=0$	**0.79**	**0.80**	**0.88**	1.06	1.10
	DOMo	$\delta_u=0$, $\delta_f=0$	**0.74**	**0.73**	**0.78**	1.01	0.93
	BIGo	$\delta_u=0$, $\delta_f=0$	**0.73**	**0.72**	**0.79**	1.05	1.01
2-Step-PCA	SURo		0.81	0.80	0.94	1.11	1.19
	DOMo		0.84	0.82	0.86	1.09	1.19
	BIGo		0.84	0.81	0.89	1.15	1.28
	SURo	$\delta_u=0$, $\delta_f=0$	**0.72**	**0.70**	**0.85**	1.12	1.17
	DOMo	$\delta_u=0$, $\delta_f=0$	**0.72**	**0.67**	**0.74**	1.09	1.15
	BIGo	$\delta_u=0$, $\delta_f=0$	**0.75**	**0.68**	**0.81**	1.17	1.27
1-Step	SURo		0.75	0.80	0.90	1.00	0.90
	DOMo		0.75	0.79	0.90	1.08	0.90
	BIGo		0.76	0.80	0.91	1.05	0.87
	SURo	$\delta_u=0$	0.70	0.74	0.84	**0.99**	0.88
	DOMo	$\delta_u=0$	0.69	0.72	0.83	1.05	0.89
	BIGo	$\delta_u=0$	0.71	0.73	0.85	1.02	0.85
	SURo	$\Gamma=0$	0.73	0.79	0.90	1.00	0.89
	DOMo	$\Gamma=0$	0.75	0.78	0.89	1.05	0.88
	BIGo	$\Gamma=0$	0.75	0.80	0.91	1.03	0.85
	SURo	$\delta_u=0$, $\Gamma=0$, $\delta_f=0$	**0.66**	**0.70**	0.85	1.01	0.88
	DOMo	$\delta_u=0$, $\Gamma=0$, $\delta_f=0$	**0.65**	**0.66**	0.83	1.10	0.93
	BIGo	$\delta_u=0$, $\Gamma=0$, $\delta_f=0$	**0.66**	**0.68**	0.84	1.08	0.89

Notes: All rows show the revision $R_{h,rel}^{(C,M)}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

Table 14: Factor models with alternative parameter restrictions, revisions to next vintage, Germany

Model	Indicator set	Parameter restrictions	Relative revision $R^{(C,M)}_{h,rel}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SURo		0.99	1.05	0.98	**0.96**	**0.98**
	DOMo		0.99	0.98	0.92	1.00	0.88
	BIGo		0.97	0.95	0.89	**1.01**	0.89
	SURo	$\delta_u=0,\ \delta_f=0$	0.95	**0.98**	**0.92**	1.07	1.18
	DOMo	$\delta_u=0,\ \delta_f=0$	**0.93**	**0.90**	**0.83**	1.09	0.98
	BIGo	$\delta_u=0,\ \delta_f=0$	**0.93**	**0.87**	**0.81**	1.09	0.99
2-Step-PCA	SURo		0.93	0.80	0.71	1.13	1.38
	DOMo		0.88	0.82	0.71	**1.12**	1.35
	BIGo		0.93	0.82	0.73	**1.12**	**1.36**
	SURo	$\delta_u=0,\ \delta_f=0$	0.90	**0.75**	**0.62**	1.19	1.44
	DOMo	$\delta_u=0,\ \delta_f=0$	**0.83**	**0.75**	**0.62**	1.18	1.41
	BIGo	$\delta_u=0,\ \delta_f=0$	**0.89**	**0.76**	**0.65**	1.20	1.45
1-Step	SURo		1.03	1.12	1.08	**0.85**	0.86
	DOMo		0.88	0.92	0.93	**1.02**	0.93
	BIGo		0.91	0.91	0.94	**0.95**	0.89
	SURo	$\delta_u=0$	0.99	1.07	1.04	**0.89**	0.87
	DOMo	$\delta_u=0$	**0.84**	0.88	0.91	1.04	0.94
	BIGo	$\delta_u=0$	0.86	0.87	0.90	**0.97**	0.88
	SURo	$\Gamma=0$	0.97	1.08	1.07	**0.88**	**0.82**
	DOMo	$\Gamma=0$	0.90	0.90	0.92	**1.01**	0.92
	BIGo	$\Gamma=0$	0.85	0.86	0.91	**0.97**	0.88
	SURo	$\delta_u=0,\ \Gamma=0,\ \delta_f=0$	**0.92**	1.05	1.07	**0.88**	**0.82**
	DOMo	$\delta_u=0,\ \Gamma=0,\ \delta_f=0$	**0.84**	**0.84**	**0.87**	1.07	0.93
	BIGo	$\delta_u=0,\ \Gamma=0,\ \delta_f=0$	**0.80**	**0.80**	**0.86**	1.01	0.88

Notes: All rows show the revision $R^{(C,M)}_{h,rel}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u=0$ and $\Gamma=0$, where the latter only applies to the one-step model.

Table 15: Factor models with alternative parameter restrictions, revisions to next vintage, France

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SURo		0.88	0.76	0.70	0.87	1.07
	DOMo		0.88	0.72	0.62	0.74	0.93
	BIGo		0.90	0.75	0.67	0.91	1.12
	SURo	$\delta_u=0,\ \delta_f=0$	**0.73**	**0.61**	**0.51**	0.87	0.88
	DOMo	$\delta_u=0,\ \delta_f=0$	**0.72**	**0.54**	**0.46**	0.83	0.79
	BIGo	$\delta_u=0,\ \delta_f=0$	0.80	**0.60**	**0.55**	0.97	1.11
2-Step-PCA	SURo		0.84	0.80	0.80	1.04	1.31
	DOMo		0.83	0.78	0.70	0.94	1.22
	BIGo		0.94	0.81	0.89	1.22	1.62
	SURo	$\delta_u=0,\ \delta_f=0$	0.84	**0.72**	**0.72**	1.06	1.24
	DOMo	$\delta_u=0,\ \delta_f=0$	**0.79**	**0.62**	**0.54**	0.93	**1.08**
	BIGo	$\delta_u=0,\ \delta_f=0$	0.98	**0.69**	**0.77**	1.25	1.59
1-Step	SURo		**0.62**	0.61	0.53	0.80	0.77
	DOMo		0.69	0.63	0.52	0.81	0.68
	BIGo		0.70	0.62	0.52	0.80	0.66
	SURo	$\delta_u=0$	**0.62**	0.56	**0.48**	**0.78**	**0.73**
	DOMo	$\delta_u=0$	0.67	0.59	0.48	0.81	0.68
	BIGo	$\delta_u=0$	**0.67**	0.58	0.49	0.80	0.66
	SURo	$\Gamma=0$	**0.64**	0.60	0.53	0.81	0.78
	DOMo	$\Gamma=0$	0.69	0.62	0.51	0.82	0.66
	BIGo	$\Gamma=0$	0.70	0.62	0.53	0.80	0.65
	SURo	$\delta_u=0,\ \Gamma=0,\ \delta_f=0$	**0.62**	**0.54**	**0.46**	0.80	**0.70**
	DOMo	$\delta_u=0,\ \Gamma=0,\ \delta_f=0$	0.68	0.58	0.48	0.86	0.71
	BIGo	$\delta_u=0,\ \Gamma=0,\ \delta_f=0$	0.69	0.58	**0.48**	0.85	0.70

Notes: All rows show the revision $R_{h,rel}^{(C,M)}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

Table 16: Factor models with alternative parameter restrictions, revisions to next vintage, Italy

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SURo		0.93	0.84	0.80	0.91	**0.77**
	DOMo		0.76	0.76	0.71	1.01	0.87
	BIGo		0.71	0.69	0.64	1.13	0.99
	SURo	$\delta_u = 0,\ \delta_f = 0$	**0.73**	**0.67**	**0.65**	0.94	0.82
	DOMo	$\delta_u = 0,\ \delta_f = 0$	0.69	**0.65**	0.60	1.02	0.89
	BIGo	$\delta_u = 0,\ \delta_f = 0$	0.66	**0.62**	0.59	1.13	1.02
2-Step-PCA	SURo		0.77	0.68	0.76	**1.00**	**0.91**
	DOMo		0.72	0.63	0.73	1.23	1.19
	BIGo		0.70	0.58	0.69	1.33	1.33
	SURo	$\delta_u = 0,\ \delta_f = 0$	**0.67**	0.61	0.69	1.02	0.93
	DOMo	$\delta_u = 0,\ \delta_f = 0$	0.67	**0.56**	0.67	1.23	1.21
	BIGo	$\delta_u = 0,\ \delta_f = 0$	0.68	0.54	0.66	1.32	1.33
1-Step	SURo		0.76	0.64	0.72	1.14	1.04
	DOMo		0.83	0.69	0.74	1.25	1.12
	BIGo		0.84	0.73	0.75	1.19	1.05
	SURo	$\delta_u = 0$	0.70	0.61	0.66	1.14	1.05
	DOMo	$\delta_u = 0$	0.75	0.61	0.65	1.22	1.11
	BIGo	$\delta_u = 0$	0.76	**0.64**	0.64	1.16	1.05
	SURo	$\Gamma = 0$	0.75	0.63	0.72	1.14	1.04
	DOMo	$\Gamma = 0$	0.83	0.69	0.72	1.24	1.11
	BIGo	$\Gamma = 0$	0.84	0.74	0.74	1.17	1.03
	SURo	$\delta_u = 0,\ \Gamma = 0,\ \delta_f = 0$	**0.68**	**0.59**	0.67	1.15	1.06
	DOMo	$\delta_u = 0,\ \Gamma = 0,\ \delta_f = 0$	**0.73**	0.61	0.67	1.30	1.18
	BIGo	$\delta_u = 0,\ \Gamma = 0,\ \delta_f = 0$	0.75	0.65	0.67	1.24	1.10

Notes: All rows show the revision $R_{h,rel}^{(C,M)}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

Table 17: Factor models with alternative parameter restrictions, revisions to next vintage, Spain

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SURo		1.16	1.26	1.71	1.87	2.05
	DOMo		1.15	1.18	1.50	1.62	1.63
	BIGo		1.08	1.12	1.46	1.56	1.51
	SURo	$\delta_u=0,\ \delta_f=0$	**0.80**	**0.95**	**1.44**	**1.61**	**1.79**
	DOMo	$\delta_u=0,\ \delta_f=0$	**0.86**	**1.00**	1.37	**1.45**	**1.39**
	BIGo	$\delta_u=0,\ \delta_f=0$	**0.76**	**0.92**	1.34	1.43	**1.30**
2-Step-PCA	SURo		0.75	0.85	1.46	1.52	1.44
	DOMo		1.15	1.16	1.42	1.55	1.55
	BIGo		1.02	1.07	1.39	1.47	1.42
	SURo	$\delta_u=0,\ \delta_f=0$	**0.50**	**0.65**	1.36	**1.46**	**1.37**
	DOMo	$\delta_u=0,\ \delta_f=0$	**0.87**	**0.93**	1.26	**1.42**	**1.37**
	BIGo	$\delta_u=0,\ \delta_f=0$	0.70	0.82	1.28	1.40	1.29
1-Step	SURo		0.68	0.87	1.44	1.41	1.21
	DOMo		0.74	0.97	1.58	1.53	1.19
	BIGo		0.72	0.96	1.63	1.57	1.18
	SURo	$\delta_u=0$	0.55	**0.73**	**1.30**	1.35	1.15
	DOMo	$\delta_u=0$	0.60	0.83	1.43	1.44	1.13
	BIGo	$\delta_u=0$	**0.61**	**0.84**	1.49	1.47	1.11
	SURo	$\Gamma=0$	0.67	0.88	1.48	1.41	1.20
	DOMo	$\Gamma=0$	0.70	0.94	1.55	1.49	1.13
	BIGo	$\Gamma=0$	0.74	0.99	1.64	1.53	1.15
	SURo	$\delta_u=0,\ \Gamma=0,\ \delta_f=0$	**0.51**	**0.70**	1.38	1.35	1.14
	DOMo	$\delta_u=0,\ \Gamma=0,\ \delta_f=0$	**0.51**	**0.74**	1.50	1.48	1.16
	BIGo	$\delta_u=0,\ \Gamma=0,\ \delta_f=0$	**0.56**	**0.79**	1.57	1.52	1.16

Notes: All rows show the revision $R_{h,rel}^{(C,M)}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u=0$ and $\Gamma=0$, where the latter only applies to the one-step model.

Table 18: Factor models with alternative parameter restrictions, revisions to next vintage, Netherlands

Model	Indicator set	Parameter restrictions	Relative revision $R^{(C,M)}_{h,rel}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SURo		0.78	0.89	0.92	0.79	0.75
	DOMo		0.65	0.75	0.77	0.62	0.54
	BIGo		0.65	0.76	0.76	0.61	0.54
	SURo	$\delta_u=0$, $\delta_f=0$	0.74	**0.78**	0.90	0.83	0.81
	DOMo	$\delta_u=0$, $\delta_f=0$	**0.49**	**0.55**	**0.65**	0.66	0.61
	BIGo	$\delta_u=0$, $\delta_f=0$	**0.50**	**0.61**	**0.67**	0.65	0.61
2-Step-PCA	SURo		0.76	0.88	0.95	0.84	0.89
	DOMo		0.62	0.73	0.76	0.62	0.64
	BIGo		0.63	0.76	0.77	0.63	0.65
	SURo	$\delta_u=0$, $\delta_f=0$	**0.71**	**0.77**	**0.88**	0.86	0.88
	DOMo	$\delta_u=0$, $\delta_f=0$	**0.46**	**0.51**	**0.62**	0.68	0.70
	BIGo	$\delta_u=0$, $\delta_f=0$	**0.52**	**0.60**	**0.67**	0.68	0.69
1-Step	SURo		0.68	0.78	0.73	0.82	**0.64**
	DOMo		0.63	0.73	0.71	0.80	0.59
	BIGo		0.64	0.78	0.72	0.76	0.55
	SURo	$\delta_u=0$	0.65	0.73	0.71	**0.78**	**0.61**
	DOMo	$\delta_u=0$	0.60	0.70	0.70	0.72	0.57
	BIGo	$\delta_u=0$	0.64	0.74	0.71	0.70	0.53
	SURo	$\Gamma=0$	0.63	0.76	0.72	**0.77**	**0.63**
	DOMo	$\Gamma=0$	0.61	0.74	0.73	0.69	0.56
	BIGo	$\Gamma=0$	0.61	0.77	0.73	**0.67**	0.54
	SURo	$\delta_u=0$, $\delta_f=0$, $\Gamma=0$	**0.57**	**0.63**	**0.67**	0.86	0.69
	DOMo	$\delta_u=0$, $\delta_f=0$, $\Gamma=0$	**0.47**	**0.53**	**0.62**	0.81	0.65
	BIGo	$\delta_u=0$, $\delta_f=0$, $\Gamma=0$	**0.49**	**0.58**	**0.60**	0.80	0.60

Notes: All rows show the revision $R^{(C,M)}_{h,rel}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

C.2 Contemporaneous indicator sets without time shift

Table 19: Including indicator sets without time shift, revisions to next vintage, average over all five countries

Model	Indicator set	Parameter restrictions	Relative revision $R^{(C,M)}_{h,rel}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SUR	$\delta_u = 0$	0.87	0.86	**0.89**	**1.01**	**0.98**
	DOM	$\delta_u = 0$	0.90	0.86	0.85	1.02	0.96
	BIG	$\delta_u = 0$	0.89	0.87	0.85	1.02	0.96
2-Step-PCA	SUR	$\delta_u = 0$	0.84	0.80	0.90	1.13	**1.15**
	DOM	$\delta_u = 0$	0.88	0.83	0.95	1.23	1.27
	BIG	$\delta_u = 0$	0.90	0.85	0.97	1.24	1.27
1-Step	SUR	$\delta_u = 0,\ \Gamma = 0$	**0.68**	**0.67**	**0.79**	1.03	0.98
	DOM	$\delta_u = 0,\ \Gamma = 0$	**0.68**	**0.67**	**0.79**	1.07	0.99
	BIG	$\delta_u = 0,\ \Gamma = 0$	**0.69**	**0.70**	**0.80**	1.03	0.93

Notes: All rows show the revision $R^{(C,M)}_{h,rel}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

Table 20: Including indicator sets without time shift, revisions to next vintage, Germany

Model	Indicator set	Parameter restrictions	Relative revision $R^{(C,M)}_{h,rel}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SUR	$\delta_u = 0$	0.98	1.00	1.02	1.02	**0.99**
	DOM	$\delta_u = 0$	**0.89**	**0.92**	0.95	1.05	0.92
	BIG	$\delta_u = 0$	**0.90**	0.93	0.97	1.05	0.95
2-Step-PCA	SUR	$\delta_u = 0$	0.96	0.94	0.95	**1.11**	**1.18**
	DOM	$\delta_u = 0$	0.90	0.88	0.91	1.25	1.42
	BIG	$\delta_u = 0$	0.93	0.91	0.93	1.28	1.47
1-Step	SUR	$\delta_u = 0,\ \Gamma = 0$	**0.81**	**0.76**	**0.79**	1.04	1.12
	DOM	$\delta_u = 0,\ \Gamma = 0$	**0.74**	**0.71**	**0.77**	1.11	1.24
	BIG	$\delta_u = 0,\ \Gamma = 0$	**0.74**	**0.71**	**0.78**	1.07	1.17

Notes: All rows show the revision $R^{(C,M)}_{h,rel}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

Table 21: Including indicator sets without time shift, revisions to next vintage, France

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SUR	$\delta_u = 0$	0.78	0.69	**0.58**	**0.78**	**0.83**
	DOM	$\delta_u = 0$	0.91	0.75	0.59	0.83	0.92
	BIG	$\delta_u = 0$	0.92	0.77	0.59	0.87	0.97
2-Step-PCA	SUR	$\delta_u = 0$	0.95	0.83	**0.75**	**1.01**	**1.07**
	DOM	$\delta_u = 0$	1.04	0.87	0.74	1.11	1.19
	BIG	$\delta_u = 0$	1.12	0.94	0.89	**1.17**	**1.17**
1-Step	SUR	$\delta_u = 0,\ \Gamma = 0$	**0.63**	0.57	0.50	0.82	0.79
	DOM	$\delta_u = 0,\ \Gamma = 0$	**0.65**	0.57	**0.45**	0.83	0.73
	BIG	$\delta_u = 0,\ \Gamma = 0$	**0.65**	0.58	**0.45**	0.80	0.69

Notes: All rows show the revision $R_{h,rel}^{(C,M)}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

Table 22: Including indicator sets without time shift, revisions to next vintage, Italy

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
2-Step-Bay	SUR	$\delta_u = 0$	0.77	**0.65**	**0.57**	0.93	0.84
	DOM	$\delta_u = 0$	0.73	**0.60**	**0.55**	1.04	0.91
	BIG	$\delta_u = 0$	0.74	**0.62**	**0.54**	**1.01**	**0.88**
2-Step-PCA	SUR	$\delta_u = 0$	0.72	**0.56**	0.76	1.24	1.21
	DOM	$\delta_u = 0$	0.73	0.57	0.99	1.45	1.44
	BIG	$\delta_u = 0$	0.75	0.58	0.91	1.39	1.37
1-Step	SUR	$\delta_u = 0,\ \Gamma = 0$	0.76	**0.58**	0.71	1.19	1.13
	DOM	$\delta_u = 0,\ \Gamma = 0$	0.76	**0.55**	0.65	1.26	1.19
	BIG	$\delta_u = 0,\ \Gamma = 0$	0.75	**0.59**	**0.61**	1.20	1.12

Notes: All rows show the revision $R_{h,rel}^{(C,M)}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

Table 23: Including indicator sets without time shift, revisions to next vintage, Spain

Model	Indicator set	Parameter restrictions	$h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
			Relative revision $R_{h,rel}^{(C,M)}$				
2-Step-Bay	SUR	$\delta_u = 0$	1.09	1.20	**1.43**	**1.53**	**1.56**
	DOM	$\delta_u = 0$	1.28	1.31	1.40	**1.45**	1.48
	BIG	$\delta_u = 0$	1.19	1.26	1.42	1.45	1.46
2-Step-PCA	SUR	$\delta_u = 0$	0.83	0.85	**1.17**	**1.38**	**1.40**
	DOM	$\delta_u = 0$	1.04	1.11	1.33	1.46	1.49
	BIG	$\delta_u = 0$	0.96	1.02	1.31	1.46	1.47
1-Step	SUR	$\delta_u = 0,\ \Gamma = 0$	0.56	**0.73**	**1.24**	**1.26**	1.14
	DOM	$\delta_u = 0,\ \Gamma = 0$	0.63	0.84	**1.37**	**1.35**	1.11
	BIG	$\delta_u = 0,\ \Gamma = 0$	0.65	0.88	**1.41**	**1.36**	1.07

Notes: All rows show the revision $R_{h,rel}^{(C,M)}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

Table 24: Including indicator sets without time shift, revisions to next vintage, Netherlands

Model	Indicator set	Parameter restrictions	$h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
			Relative revision $R_{h,rel}^{(C,M)}$				
2-Step-Bay	SUR	$\delta_u = 0$	**0.72**	**0.78**	**0.83**	0.81	**0.69**
	DOM	$\delta_u = 0$	0.68	0.74	0.75	0.73	0.57
	BIG	$\delta_u = 0$	0.69	0.77	0.73	0.70	0.56
2-Step-PCA	SUR	$\delta_u = 0$	0.76	**0.80**	**0.87**	0.93	0.88
	DOM	$\delta_u = 0$	0.69	0.74	0.78	0.88	0.82
	BIG	$\delta_u = 0$	0.75	0.82	0.82	0.90	0.85
1-Step	SUR	$\delta_u = 0,\ \Gamma = 0$	0.63	0.69	0.72	0.84	0.72
	DOM	$\delta_u = 0,\ \Gamma = 0$	0.61	0.69	0.73	0.79	0.66
	BIG	$\delta_u = 0,\ \Gamma = 0$	0.64	0.74	0.73	0.73	0.61

Notes: All rows show the revision $R_{h,rel}^{(C,M)}$ relative to the baseline EU model. Improvements over this model are marked by shaded cells. Bold numbers indicate an improvement over the corresponding factor model with the same indicator set but with the baseline restrictions $\delta_u = 0$ and $\Gamma = 0$, where the latter only applies to the one-step model.

C.3 Vintages 2005-2019 excluding the COVID-19 years

Table 25: Vintage sample without COVID-19 years, revisions to next vintage, average over all five countries

Model	Indicator set	Parameter restrictions	Relative revision $R^{(C,M)}_{h,rel}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.39	0.51	0.58	0.80	0.78
	-		1.12	1.08	1.02	1.03	0.95
HP-EST	-		1.01	1.23	1.28	1.21	**0.80**
HP	-	$\lambda = 6.25$	1.00	1.17	1.15	1.10	0.82
	-	$\lambda = 100$	1.18	1.19	1.01	**0.87**	0.93
2-Step-Bay	SURo	$\delta_u = 0$	0.90	0.91	0.92	1.23	1.25
	DOMo	$\delta_u = 0$	0.89	0.89	0.79	1.01	1.06
	BIGo	$\delta_u = 0$	0.89	0.89	0.77	1.02	1.05
2-Step-PCA	SURo	$\delta_u = 0$	0.82	0.81	0.89	1.37	1.44
	DOMo	$\delta_u = 0$	0.86	0.83	**0.73**	1.19	1.42
	BIGo	$\delta_u = 0$	0.89	0.85	**0.73**	1.24	1.47
1-Step	SURo	$\delta_u = 0,\ \Gamma = 0$	**0.75**	**0.79**	0.80	1.18	1.08
	DOMo	$\delta_u = 0,\ \Gamma = 0$	0.78	0.80	0.76	1.17	1.08
	BIGo	$\delta_u = 0,\ \Gamma = 0$	0.78	0.82	0.77	1.17	1.05

Notes: The first row shows the average revision, $R^{(C,M)}_h$, of the EU model. The remaining rows show the revision $R^{(C,M)}_{h,rel}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

Table 26: Vintage sample without COVID-19 years, revisions to next vintage, Germany

Model	Indicator set	Parameter restrictions	Relative revision $R^{(C,M)}_{h,rel}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.34	0.43	0.49	0.70	0.50
	-		1.05	1.07	1.03	0.86	0.82
HP-EST	-		0.82	0.83	0.95	0.76	0.69
HP	-	$\lambda = 6.25$	**0.69**	0.78	0.96	0.86	0.67
	-	$\lambda = 100$	0.92	0.99	0.93	**0.65**	0.67
2-Step-Bay	SURo	$\delta_u = 0$	0.89	0.97	0.88	1.03	0.98
	DOMo	$\delta_u = 0$	0.90	0.92	0.85	1.06	1.00
	BIGo	$\delta_u = 0$	0.92	0.91	0.82	1.08	0.98
2-Step-PCA	SURo	$\delta_u = 0$	0.82	**0.74**	**0.59**	1.19	1.52
	DOMo	$\delta_u = 0$	0.79	0.77	0.60	1.18	1.55
	BIGo	$\delta_u = 0$	0.88	0.79	0.62	1.21	1.58
1-Step	SURo	$\delta_u = 0,\ \Gamma = 0$	0.93	1.04	1.03	0.84	**0.64**
	DOMo	$\delta_u = 0,\ \Gamma = 0$	0.82	0.86	0.87	1.02	0.87
	BIGo	$\delta_u = 0,\ \Gamma = 0$	0.80	0.84	0.87	0.98	0.82

Notes: The first row shows the average revision, $R^{(C,M)}_h$, of the EU model. The remaining rows show the revision $R^{(C,M)}_{h,rel}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

Table 27: Vintage sample without COVID-19 years, revisions to next vintage, France

Model	Indicator set	Parameter restrictions	Relative revision $R^{(C,M)}_{h,rel}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.52	0.83	0.81	0.83	0.89
	-		1.07	1.06	1.01	1.01	1.02
HP-EST	-		0.97	0.91	0.59	0.74	0.69
HP	-	$\lambda = 6.25$	0.93	0.90	0.62	0.67	0.65
	-	$\lambda = 100$	0.79	0.81	0.56	**0.31**	**0.53**
2-Step-Bay	SURo	$\delta_u = 0$	0.78	0.71	0.58	0.73	0.93
	DOMo	$\delta_u = 0$	0.79	0.66	0.51	0.57	0.85
	BIGo	$\delta_u = 0$	0.80	0.68	0.55	0.83	1.08
2-Step-PCA	SURo	$\delta_u = 0$	0.83	0.78	0.77	1.03	1.30
	DOMo	$\delta_u = 0$	0.81	0.72	0.63	0.78	1.24
	BIGo	$\delta_u = 0$	0.91	0.77	0.82	1.26	1.68
1-Step	SURo	$\delta_u = 0, \Gamma = 0$	**0.70**	**0.62**	0.46	0.80	0.86
	DOMo	$\delta_u = 0, \Gamma = 0$	0.72	0.64	**0.40**	0.74	0.76
	BIGo	$\delta_u = 0, \Gamma = 0$	0.75	0.65	**0.40**	0.72	0.75

Notes: The first row shows the average revision, $R^{(C,M)}_h$, of the EU model. The remaining rows show the revision $R^{(C,M)}_{h,rel}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

Table 28: Vintage sample without COVID-19 years, revisions to next vintage, Italy

Model	Indicator set	Parameter restrictions	Relative revision $R^{(C,M)}_{h,rel}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.27	0.39	0.41	0.68	0.69
	-		1.02	0.83	0.72	**0.82**	**0.76**
HP-EST	-		1.04	1.16	1.02	1.14	0.93
HP	-	$\lambda = 6.25$	1.09	1.09	1.01	1.02	0.83
	-	$\lambda = 100$	1.40	0.99	0.76	0.89	0.90
2-Step-Bay	SURo	$\delta_u = 0$	0.74	0.57	0.47	1.02	0.89
	DOMo	$\delta_u = 0$	**0.67**	0.58	0.43	1.17	1.07
	BIGo	$\delta_u = 0$	**0.67**	0.60	0.42	1.11	0.98
2-Step-PCA	SURo	$\delta_u = 0$	0.70	0.53	0.57	1.22	1.16
	DOMo	$\delta_u = 0$	**0.67**	**0.52**	0.57	1.41	1.46
	BIGo	$\delta_u = 0$	0.68	**0.52**	**0.41**	1.31	1.32
1-Step	SURo	$\delta_u = 0, \Gamma = 0$	0.74	0.55	0.49	1.27	1.18
	DOMo	$\delta_u = 0, \Gamma = 0$	0.79	0.59	0.44	1.26	1.13
	BIGo	$\delta_u = 0, \Gamma = 0$	0.79	0.61	0.42	1.22	1.08

Notes: The first row shows the average revision, $R^{(C,M)}_h$, of the EU model. The remaining rows show the revision $R^{(C,M)}_{h,rel}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

Table 29: Vintage sample without COVID-19 years, revisions to next vintage, Spain

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.28	**0.29**	**0.25**	**0.27**	**0.36**
	-		1.52	1.46	1.41	1.49	1.25
HP-EST	-		1.24	1.90	2.90	2.69	1.23
HP	-	$\lambda = 6.25$	1.62	2.17	2.38	2.29	1.53
	-	$\lambda = 100$	2.01	2.15	2.08	1.96	2.20
2-Step-Bay	SURo	$\delta_u = 0$	1.36	1.51	1.82	2.57	2.72
	DOMo	$\delta_u = 0$	1.44	1.54	1.40	1.63	1.83
	BIGo	$\delta_u = 0$	1.40	1.49	1.32	1.47	1.70
2-Step-PCA	SURo	$\delta_u = 0$	1.00	1.13	1.57	2.53	2.37
	DOMo	$\delta_u = 0$	1.40	1.40	1.09	1.92	2.21
	BIGo	$\delta_u = 0$	1.34	1.38	1.05	1.80	2.13
1-Step	SURo	$\delta_u = 0, \Gamma = 0$	**0.77**	1.02	1.34	2.18	2.10
	DOMo	$\delta_u = 0, \Gamma = 0$	0.94	1.20	1.35	2.12	2.07
	BIGo	$\delta_u = 0, \Gamma = 0$	0.94	1.25	1.41	2.22	2.09

Notes: The first row shows the average revision, $R_h^{(C,M)}$, of the EU model. The remaining rows show the revision $R_{h,rel}^{(C,M)}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.

Table 30: Vintage sample without COVID-19 years, revisions to next vintage, Netherlands

Model	Indicator set	Parameter restrictions	Relative revision $R_{h,rel}^{(C,M)}$ $h=-2$	$h=-1$	$h=0$	$h=1$	$h=2$
EU model	-	$\delta_u = 0$	0.55	0.63	0.92	1.50	1.45
	-		0.96	0.98	0.95	0.96	0.89
HP-EST	-		0.98	1.36	0.94	0.74	0.45
HP	-	$\lambda = 6.25$	0.68	0.93	0.80	0.68	0.40
	-	$\lambda = 100$	0.78	1.00	0.72	**0.53**	**0.33**
2-Step-Bay	SURo	$\delta_u = 0$	0.72	0.80	0.86	0.80	0.75
	DOMo	$\delta_u = 0$	0.66	0.76	0.75	0.61	0.53
	BIGo	$\delta_u = 0$	0.66	0.77	0.74	0.59	0.52
2-Step-PCA	SURo	$\delta_u = 0$	0.75	0.86	0.93	0.86	0.87
	DOMo	$\delta_u = 0$	0.65	0.75	0.76	0.65	0.64
	BIGo	$\delta_u = 0$	0.64	0.77	0.77	0.64	0.65
1-Step	SURo	$\delta_u = 0, \Gamma = 0$	**0.62**	**0.70**	**0.69**	0.83	0.64
	DOMo	$\delta_u = 0, \Gamma = 0$	**0.62**	0.73	0.73	0.69	0.55
	BIGo	$\delta_u = 0, \Gamma = 0$	0.64	0.77	0.74	0.69	0.52

Notes: The first row shows the average revision, $R_h^{(C,M)}$, of the EU model. The remaining rows show the revision $R_{h,rel}^{(C,M)}$ relative to the EU model. Improvements over the EU model are marked by shaded cells, the best model is highlighted by a bold number.