Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Measuring the Euro Area Output Gap

Matteo Barigozzi, Claudio Lissona, and Matteo Luciani

2024-099

Please cite this paper as:
Barigozzi, Matteo, Claudio Lissona, and Matteo Luciani (2024). "Measuring the Euro Area Output Gap," Finance and Economics Discussion Series 2024-099. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2024.099.

Measuring the Euro Area Output Gap*

Matteo Barigozzi
University of Bologna
matteo.barigozzi@unibo.it

Claudio Lissona
University of Bologna
claudio.lissona2@unibo.it

Matteo Luciani
Federal Reserve Board
matteo.luciani@frb.gov

Last update: October 29, 2024

Abstract

We measure the Euro Area (EA) output gap and potential output using a non-stationary dynamic factor model estimated on a large dataset of macroeconomic and financial variables. From 2012 to 2023, we estimate that the EA economy was tighter than the European Commission and the International Monetary Fund estimate, suggesting that the slow EA growth is the result of a potential output issue, not a business cycle issue. Moreover, we find that credit indicators are crucial for pinning down the output gap, as excluding them leads to estimating a lower output gap in periods of debt build-up and a higher gap in periods of deleveraging.

1 Introduction

The decomposition of GDP in potential output and the output gap—defined as the percentage deviation of GDP from its potential—is a fundamental task for policymakers. These quantities are essential for the common Euro Area (EA) monetary policy and the fiscal policy of individual countries. The output gap helps assess the cyclical position of the economy and, thus, potential inflationary pressures (e.g., Bańbura and Bobeica, 2023). Potential output and the output gap are among the main pillars of the EA fiscal surveillance framework, ultimately affecting the fiscal capacity of each member country (European Commission, 2018). However, since both quantities are unobserved, policymakers need a model that can extract them from the data.

This paper proposes a new measure of potential output and the output gap for the EA based on a non-stationary dynamic factor model estimated on a large dataset of EA macroeconomic and financial variables that incorporates relevant EA data features. Compared to the prevailing literature, which focuses on theoretical structural models with few variables of interest, we adopt a

*We would like to thank for helpful comment Travis Berge, Danilo Cascaldi Garcia, Antonio Conti, Thiago Ferreira, Manuel Gonzalez-Astudillo, Michele Lenza, Giovanni Pellegrino, and Riccardo Trezzi. This paper has benefited also from discussions with seminar participants at the Federal Reserve Board and with participants of several conferences. M. Barigozzi and C. Lissona gratefully acknowledge financial support from MIUR (PRIN2020, Grant 2020N9YFFE). Of course, any error is our responsibility.

distinct approach because we let the data speak by leveraging a large information set conditional on a few key macroeconomic priors—for example, the long-run slowdown in output growth (Cette et al., 2016).

There are at least two reasons why using a rich information set is crucial when estimating the output gap. First, cross-sectional aggregation of a large number of series allows us to consistently disentangle the co-movements in the data from idiosyncratic dynamics. Second, it is now widely recognized that potential output and the output gap, along with many other macroeconomic latent policy variables, commonly referred to as "stars," are driven by a large number of shocks; hence, we need a large information set to obtain meaningful estimates of these quantities (Buncic and Pagan, 2022). To the best of our knowledge, this is the first paper that estimates the EA output gap with such a rich information set.

We conduct our analysis on a new large-dimensional dataset comprising 119 EA economic indicators from 2001:Q1-2023:Q4. Three main results emerge from our analysis: first, our output gap estimate is in line with those published by the European Commission (EC) and the International Monetary Fund (IMF) in that the dating of the turning points perfectly coincides. The three estimates closely align until the 2011–2012 Sovereign Debt Recession (henceforth, SDR), after which our output gap measure suggests that the EA economy was tighter than estimated by the EC and the IMF. Moreover, we estimate that potential output growth decelerated after the 2008–2009 Global Financial Crisis (henceforth, GFC), the SDR further compounded this deceleration, and as of the end of 2023, potential output growth has yet to return to the pre-GFC pace. In other words, our results suggest that the EA has a potential output issue, not a business cycle issue. Hence, if the goal is to achieve better economic conditions in the EA, European countries should implement supply-side structural reforms that have long-run effects, while policies aiming at stimulating aggregate demand will have only short-term effects at best.

Second, we find that our model takes considerable signal for estimating the output gap from labor market and inflation variables, as, on average, it satisfies the Okun's law relationship and exhibits Phillips correlation, despite the fact we do not impose either of them. Moreover, we find that our output gap measure better predicts inflation than those obtained using commonly adopted methods. These results confirm that our data-driven measure is economically meaningful.

Third, in line with the argument of Borio et al. (2017), who argue that incorporating financial indicators in the dataset is necessary to obtain meaningful estimates of the business cycle, we find that credit indicators (mainly household liabilities) are crucial for pinning down the cyclical position of the economy. Specifically, we find that excluding household liabilities leads to estimating a lower output gap in periods of debt build-up (e.g., before the GFC) and a higher gap in periods of deleveraging (e.g., from 2013 to 2016), as the boost to GDP from increased leveraging is offset in periods of deleveraging. In other words, our results show that growth financed through household debt is not sustainable in the long run.

To measure potential output and the output gap, we first estimate the non-stationary dynamic factor model by Quasi Maximum Likelihood using the EM algorithm (Doz et al., 2012; Barigozzi and Luciani, 2024). Then, we extract a common trend from the estimated common factors running a second EM algorithm compute the cyclical component by subtracting the common trend from the common factors. In practice, however, to account for the Covid shock, we run this procedure twice, as we first estimate the model using only pre-Covid data. Then, we estimate the effects induced by the Covid shock (level-shift and increased volatility). Lastly, we re-estimate the model on the full dataset after purging the data from Covid-induced dynamics. Having estimated the common trend and the common cyclical component, we measure potential output as the part of GDP explained by the common trends and the output gap as the part of GDP explained by the common cyclical component.

How to estimate potential output and the output gap has been a hotly debated topic for several decades. The literature has proposed two main approaches: a theoretical approach and a statistical approach. The theoretical approach uses theoretical models, such as production-function-based models used by the EC (Havik et al., 2014) and the IMF (De Masi, 1997) or New-Keynesian DSGE models (Justiniano et al., 2013; Burlon and D'Imperio, 2020; Furlanetto et al., 2021). The statistical approach uses (univariate or multivariate) statistical models, sometimes paired with some macroeconomic relationships of interest, e.g., the Phillips curve. For example, many papers rely on univariate models (e.g., Morley et al., 2003; Kamber et al., 2018; Hamilton, 2018; Phillips and Jin, 2021; Phillips and Shi, 2021; Hartl et al., 2022), while few others employ multivariate non-stationary methods, which are either low-dimensional models (Jarociński and Lenza, 2018; González-Astudillo, 2019; Tóth, 2021; Hasenzagl et al., 2022), or medium-size, but stationary, models (Aastveit and Trovik, 2014; Morley and Wong, 2020; Morley et al., 2023). Most of these works focus on the US, while only a few of them focus on the EA, the most recent being Morley et al. (2023).

Our model is large-dimensional and non-stationary, thus allowing us to capture well-established co-movements in macroeconomic variables while retaining data in levels. Moreover, being an enhanced version of the model Barigozzi and Luciani (2023) used to measure the US output gap—enhanced because it does not require a long sample to identify the trend and deals with the Covid pandemic—our model follows the statistical approach and belongs to the class of unobserved component models. This class of models is among the most indicated ones for extracting the transitory component of GDP (Canova, 2022).

The rest of the paper is structured as follows. In Section 2, we present the data used. In Section 3, we present the model, the estimation strategy, and the model setup. Section 4 presents the main results. Next, Sections 5 and 6 shed light on from which variables the model takes more signal for the output gap estimation, with Section 5 focusing on labor market and inflation variables, and Section 6 on credit indicators. In Section 7, we assess the ability of our output gap estimate to forecast inflation. Section 8 concludes. This paper also contains an online appendix providing additional details and robustness analysis. Specifically, Appendix A provides full details on the dataset, while Appendixes B, D, and E provide additional details about the model. Next, Appendixes F, H, and I provide robustness analysis. Lastly, Appendixes C, G, and J provide additional results that did not make the cut to be in the main text.

2 A large Euro Area dataset

We construct a large macroeconomic dataset of $n = 119$ EA series, observed from 2001:Q1 to 2023:Q4, ($T = 92$). The dataset contains a wide range of macroeconomic indicators, including national account statistics, industrial production and turnover indicators, labor market and compensation indicators, price indexes, oil prices, natural gas prices, house prices, exchange rates, interest rates, a stock market index, monetary aggregates, non-financial assets and liabilities, and confidence indexes.

In terms of broad categories of data, we include in the dataset the usual suspects normally considered for high-dimensional macroeconomic analysis (see, for example, McCracken and Ng, 2016, 2020).

In terms of which and how many series to include for each category, we use a mix of economic and statistical reasoning. On the one hand, to identify the common factors driving the co-movement in the data, it is crucial to pool information from many indicators; hence, a larger information set should be preferred. On the other hand, a key assumption of the model is the presence of mild cross-sectional correlation among the idiosyncratic components: violating this assumption leads to a deterioration of the model's performance (Boivin and Ng, 2006; Luciani, 2014). Thus, when building a dataset for factor analysis, we face a trade-off between the need for a larger information set and the risk of introducing too much idiosyncratic correlation. For this reason, we selected the variables to include in the dataset to maximize economic signal while limiting the noise, with exceptions motivated by economic reasoning. For instance, we include GDP and its components since their informational content justifies a relatively high level of idiosyncratic correlations. Similarly, consumption and employment are decomposed according to durability and sectoral composition, respectively, while assets and liabilities are decomposed by ownership. In contrast, we keep the consumer price index for energy while dropping the producer price index for energy, as they carry the same signal (their correlation is greater than 0.95). Likewise, we drop the consumer price index for industrial goods because it has a correlation greater than 0.95 with the goods consumer price index.

As for the treatment of the series, we take logarithms for all variables except for confidence indicators and those already expressed in percentage points. We keep all variables in levels except for price indicators, for which we take first differences; i.e., we work with inflation rates. This is a common approach used in the literature to avoid spurious dynamics resulting from the potential $I(2)$ behavior in price indexes (Stock and Watson, 2016; McCracken and Ng, 2016, 2020). Appendix A provides the complete list of variables included in the dataset, along with their sources and treatment.

3 Methodology

In Section 3.1, we outline the model and its main features—we discuss all the details, formal assumptions, and further comments in Appendix B. In Section 3.2, we sketch how we estimate the model, while referring the reader to Appendix D for a step-by-step guide on how estimation is carried out in practice, and to Appendix E for the bootstrap procedure used to measure uncertainty around our estimates. Finally, in Section 3.3 we discuss the model setup.

3.1 The model

We denote the observed i-th time series at a given quarter t as y_{it}, with $1 \leq i \leq 119$ and $2001 : Q1 \leq t \leq 2023 : Q4$. In our non-stationary dynamic factor model, each variable is the sum of: (i) a secular component D_{it}, which is treated either as deterministic or stochastic, (ii) q common factors $\mathbf{f}_t = (f_{1t} \cdots f_{qt})'$, which capture the macroeconomic long- and short-run co-movements and have a dynamics governed by a VAR, (ii) an idiosyncratic component ξ_{it}, which captures local dynamics or measurement errors and is possibly correlated across i and t.

We partition the n series according to two features. First, according to the nature of the secular component, that is whether D_{it} is a stochastic or a deterministic process. Second, according to the nature of the idiosyncratic component, that is whether ξ_{it} has a stochastic trend or it is stationary.

In particular, we model D_{it} as a local linear trend for GDP to account for the well-documented slowdown in productivity (Cette et al., 2016) and for households' financial liabilities (HHLB) and households' long-term loans (HHLB.LLN), which are more than 80% of total household's liabilities, whose average growth rate has slowed down consistently since the GFC. Moreover, we model D_{it} as a local level model for the unemployment rate (UNETOT) to account for relevant labor market features, such as the reallocation of employees across sectors, which contributed to the slowdown in the EA productivity growth (Cette et al., 2016), and for overall and core inflation (HICPOV and HICPNEF, respectively) to account for the slowdown in inflation occurred after the GFC.[1] For all other series, D_{it} is either a linear trend with a constant slope, in which case we say that $i \in \mathcal{I}_b$, or D_{it} is just a constant equal to D_{i0}.

As for the idiosyncratic components, if $\xi_{it} \sim I(1)$, then we say that $i \in \mathcal{I}_1$ and we model ξ_{it} as a random walk, while if $\xi_{it} \sim I(0)$, we say that $i \notin \mathcal{I}_1$ and we leave its dynamics unspecified to avoid over-parametrization of the model.

Furthermore, we capture the effect of the Covid shock, which generated a large shift both in the levels (Ng, 2021; Maroz et al., 2021), and in the volatility (Carriero et al., 2022; Lenza and Primiceri, 2022) of most macroeconomic EA series, through an additional common factor g_t (Maroz et al., 2021), and a scalar s_t scaling the conditional volatility of the latent factors (Lenza and Primiceri, 2022). While the former has an impact on all series only in 2020 and 2021, the latter has an effect that persists even after the recovery from the pandemic. This parametrization is different from the one normally used for the US, where the Covid factor's effect persists only through the summer of 2020 (Maroz et al., 2021), and the increase in volatility lasts for a shorter amount of time (Lenza and Primiceri, 2022). We use this different parametrization because mobility restrictions and lockdowns in the EA have been on and off until early 2022, while in the US were enforced only at the beginning of the pandemic.

Formally, the model reads as follows:

$$y_{it} = \mathrm{D}_{it} + \boldsymbol{\lambda}_i'\mathbf{f}_t + \gamma_i g_t \mathbb{I}_{2020:\mathrm{Q}1 \le t \le 2021:\mathrm{Q}4} + \xi_{it}, \qquad 1 \le i \le 119, \quad 2001{:}\mathrm{Q}1 \le t \le 2023{:}\mathrm{Q}4, \tag{1}$$

$$\mathrm{D}_{it} = \mathrm{D}_{it-1} + b_{i,t-1}\mathbb{I}_{i \in \mathcal{I}_b} + \epsilon_{it}, \qquad \epsilon_{it} \overset{i.i.d.}{\sim} (0, \sigma^2_{\epsilon_i}\mathbb{I}_{i=\mathrm{UNETOT,HICPOV,HICPNEF}}), \tag{2}$$

$$b_{it} = b_{it-1} + \eta_{it}, \qquad \eta_{it} \overset{i.i.d.}{\sim} (0, \sigma^2_{\eta_i}\mathbb{I}_{i=\mathrm{GDP,HHLB,HHLB.LLN}}), \tag{3}$$

$$\mathbf{f}_t = \sum_{j=0}^{p-1} \mathbf{A}_j \mathbf{f}_{t-j} + \{s_t \mathbb{I}_{t \ge 2020:\mathrm{Q}1} + \mathbb{I}_{t < 2020:\mathrm{Q}1}\}\mathbf{u}_t, \qquad \mathbf{u}_t \overset{i.i.d.}{\sim} (\mathbf{0}, \boldsymbol{\Sigma}_u), \tag{4}$$

$$\xi_{it} = \xi_{it-1}\mathbb{I}_{i \in \mathcal{I}_1} + e_{it}, \qquad e_{it} \sim (0, \sigma^2_{e_i}), \tag{5}$$

where $\mathbb{I}_A = 1$, if A is true, and $\mathbb{I}_A = 0$, otherwise.

Furthermore, we assume that one common trend, τ_t, which we model as a random walk, drives

[1] Appendix I shows robustness results when removing the time-variation in the secular trends.

the non-stationarity in the factors $\mathbf{f}_t$ (Del Negro et al., 2007). We remain agnostic on the law of motion of the residual, i.e., the stationary cyclical component, which we denote as $\boldsymbol{\omega}_t$. Specifically, we consider the decomposition:

$$\mathbf{f}_t = \boldsymbol{\psi}\tau_t + \boldsymbol{\omega}_t, \qquad \boldsymbol{\omega}_t \sim (\mathbf{0}, \boldsymbol{\Sigma}_\omega), \tag{6}$$

$$\tau_t = \tau_{t-1} + \nu_t, \qquad \nu_t \sim (0, \sigma^2_\nu). \tag{7}$$

We can then identify the common trend τ_t by properly initializing its variance σ^2_ν, which is analogous to controlling the signal-to-noise ratio. This approach is equivalent to Morley et al. (2023) approach, as it has the same goal (smoothing the trend) but achieves it differently: we smooth the trend using the Kalman smoother properly initialized by setting σ^2_ν; Morley et al. (2023) smooth the trend using the estimated parameters of an ARIMA model. A comparison between the two approaches is in Appendix F.

In addition, by properly initializing its variance σ^2_ν, our estimation strategy for the common trend is consistent with the results of Kim and Kim (2022), who show that extracting a common trend modeled with an unconstrained random walk component overfits GDP, thus producing a potential output that fluctuates too much and generating the so-called "pile-up" problem. Moreover, Kim and Kim (2022) show that the common practice of modeling the stationary component as an AR(2) poses identification problems since the same state-space representation can also be obtained with an ARMA(2,1) specification. For this reason, we do not impose any parametric model for $\boldsymbol{\omega}_t$.

The model we just described is a modified version of the model Barigozzi and Luciani (2023) (BL) used to estimate the output gap in the US. We modified BL's model to overcome two important limitations. First, their model relies on estimating cointegrating relationships to identify the common trends, thus requiring long time series to get a reliable estimate. As such, BL's model can be estimated only on US macroeconomic data for which more than 50 years of quarterly data are available. Second, BL estimate their model on pre-Covid data; thus, to incorporate more recent observations, some modification is needed to handle the different co-movements brought about by the Covid pandemic. In this paper, we solve both limitations by introducing (6)-(7), which we can estimate even on short samples and by incorporating recently proposed methods to handle the Covid period in the estimation strategy.

Combining Equations (1) and (6), we obtain the decomposition of each observed variable:

$$y_{it} = \mathrm{D}_{it} + \boldsymbol{\lambda}_i'\boldsymbol{\psi}\tau_t + \boldsymbol{\lambda}_i'\boldsymbol{\omega}_t + \gamma_i g_t \mathbb{I}_{2020:\mathrm{Q1} \leq t \leq 2021:\mathrm{Q4}} + \xi_{it}.$$

Focusing on GDP, we define potential output, PO_t, and the output gap, OG_t, as:

$$\mathrm{PO}_t = \mathrm{D}_{\mathrm{GDP},t} + \boldsymbol{\lambda}_{\mathrm{GDP}}'\boldsymbol{\psi}\tau_t, \tag{8}$$

$$\mathrm{OG}_t = \boldsymbol{\lambda}_{\mathrm{GDP}}'\boldsymbol{\omega}_t. \tag{9}$$

Hence, in our framework, potential output is the sum of the time-varying secular trend of GDP ($\mathrm{D}_{\mathrm{GDP},t}$), which captures the long-run decline in EA output growth, and the part of GDP driven by the common trend component (τ_t); the output gap is the part of GDP driven by the stationary cyclical component ($\boldsymbol{\omega}_t$). What is left out are the idiosyncratic component, $\xi_{\mathrm{GDP},t}$ and the Covid component $\gamma_{\mathrm{GDP}} g_t$. While the idiosyncratic component is likely to be just a measurement error

(Aruoba et al., 2016), hence, it is clear why we are leaving it out; the exclusion of the Covid shock deserves an explanation.

The Covid component represents the co-movements from 2020:Q1 to 2021:Q4 that neither potential output nor the output gap captures. In principle, this component could be allocated to the output gap, which would be equivalent to assuming that the productive capacity of the EA "froze" due to the lockdowns. While this view is commonly accepted by European institutions (Thum-Thysen et al., 2022), it is still unclear whether, and by what amount, the EA productive capacity has been affected by the Covid shock. Thus, we remain agnostic on the allocation of the Covid component between potential output and the output gap, and we will present it as a standalone component.

In summary, our setup lets the data speak by leveraging a large number of variables. Moreover, it allows us to reconcile a purely statistical factor model with some key macroeconomic priors that are likely important to obtain meaningful estimates of potential output and the output gap. Lastly, it allows us to go beyond the common practice of pre-transforming data to work with stationary and centered variables, thus being able to capture essential features that would be inevitably lost if we were to difference the data to achieve stationarity (Ng, 2018).

3.2 Estimating the model

To estimate the model in (1)-(5), we need to extract the latent states $\mathbf{f}_t$, g_t, $\mathrm{D}_{\mathrm{GDP},t}$, $\mathrm{D}_{\mathrm{HHLB},t}$, $\mathrm{D}_{\mathrm{HHLB.LLN},t}$, $\mathrm{D}_{\mathrm{UNETOT},t}$, $\mathrm{D}_{\mathrm{HICPOV},t}$, $\mathrm{D}_{\mathrm{HICPNEF},t}$, and ξ_{it} (if $i \in \mathcal{I}_1$), and estimate the parameters $\boldsymbol{\lambda}_i$, γ_i, $\mathbf{A}_j$, s_t, $\boldsymbol{\Sigma}_u$, $\sigma^2_{e_i}$, a_i, and b_i. To do so, we use a three-step estimation procedure that we summarize below.

Step 1: Estimate the model up to 2019:Q4 (pre-Covid step). We obtain a preliminary estimate of the parameters using non-stationary PCAs (Bai and Ng, 2004; Barigozzi et al., 2021; Onatski and Wang, 2021). Then, we run the EM algorithm, jointly with the Kalman smoother, as described in Doz et al. (2012) and Barigozzi and Luciani (2024) in the high-dimensional stationary case, and reconsidered in the non-stationary case by Barigozzi and Luciani (2019, 2023).

Step 2: Estimate the Covid factor and volatility (Covid step). Using the parameter estimated over the pre-Covid period, we apply the Kalman filter and smoother to extract the latent states using data up to the end of the sample. To address the influence of the Covid outliers on the estimates, we truncate the Kalman Smoother at 2020:Q1 and then we continue the backward iterations by re-initializing the Kalman smoother using the Kalman filter estimate for 2019:Q4. This truncation avoids any backward spurious effect from the Covid period to the pre-Covid estimates, but creates a structural break in the estimated secular components in 2020:Q1. For this reason, we adjust the level of the smoothed secular components by judgmentally allocating the break in 2020:Q1 to the idiosyncratic component, as suggested by Ahn and Luciani (2024).[2]

[2] An alternative approach we could have taken consists in adjusting for outliers on a series-by-series basis. However, we did not pursue this option because much of the additional volatility during the Covid shock is "economic" volatility induced by the pandemic, not measurement error (Ng, 2021). A univariate outlier adjustment method cannot distinguish between the two, so it likely removes economically relevant information. Indeed, if we do univariate outlier adjustment, we get that the Covid factor is essentially zero and that the output gap is flat as if nothing happened.

At this point, we have an estimate of the states over the entire sample given the information set prior to the Covid shock. This implies that the co-movements in the Covid period are left unaccounted for, and are captured by the idiosyncratic component ξ_{it}. Thus, since Covid was a common shock affecting most of (if not all) the series in the dataset, we estimate the Covid factor $\widehat{g}_t$ and its loadings $\widehat{\gamma}_i$ by PCA on the variance-covariance matrix of the idiosyncratic component for the period 2020:Q1-2021:Q4, following Maroz et al. (2021).

Lastly, we estimate the Covid volatilities $\widehat{s}_t$ for the period 2020:Q1-2023:Q4 by maximum likelihood and by using the factors extracted using pre-Covid parameters. We find that $\widehat{s}_t$ jumps from 1 to about 3.5 at the onset of the Covid pandemic, and then remain larger than 2 until the end of 2022, thus justifying our choice of imposing a time-varying volatility until the end of the sample. In Appendix H, we show how our measures would change if we do not model the effect of Covid explicitly, or if we use the exponential decay parametrization proposed by Lenza and Primiceri (2022).

Step 3: Full sample estimation. We estimate all the parameters and latent states up to the end of the sample, by using data net of the Covid component, i.e., with $y_{it} - \widehat{\gamma}_i\widehat{g}_t$. Specifically, by using the factors estimated over the whole sample in Step 2 rescaled by $\widehat{s}_t$ in the last part of the sample, we estimate the parameters, $\widehat{\boldsymbol{\lambda}}_i$, $\widehat{\mathbf{A}}_j$, $\widehat{\boldsymbol{\Sigma}}_u$, $\widehat{\sigma}^2_{e_i}$, $\widehat{a}_i$, and $\widehat{b}_i$, by maximizing the expected likelihood. Finally, with the estimated parameters in hand, we obtain a final estimate of the states, $\widehat{\mathbf{f}}_t$, $\widehat{\mathrm{D}}_{i,t}$, and $\widehat{\xi}_{it}$, through the Kalman smoother again truncated in 2020:Q1 and reinitialized before iterating backward, as explained in Step 2.

Having estimated the model parameters and unobserved states, we can now estimate the common trend. To this end, we estimate the state-space model in (6)-(7) using the EM algorithm by replacing the true factors with the estimated factors.

At convergence of the EM algorithm, we obtain a final estimate of the parameters, $\widehat{\boldsymbol{\psi}}$, $\widehat{\boldsymbol{\Sigma}}_\omega$, and $\widehat{\sigma}^2_\nu$, and using these estimates we have a final estimate of the trend $\widehat{\tau}_t$ and of the cyclical component $\widehat{\boldsymbol{\omega}}_t = \widehat{\mathbf{f}}_t - \widehat{\boldsymbol{\psi}}\widehat{\tau}_t$, obtained through the Kalman smoother. Given the estimates of the common trend and the cyclical common component, we compute our final estimates of potential output and the output gap according to (8) and (9), respectively.

The estimation procedure we just outlined delivers consistent estimators of all parameters and of the factors, provided that n and T grow to infinity. Furthermore, we neither have to impose the gaussianity assumption, nor we have to require uncorrelatedness of the idiosyncratic components (ξ_{it} if $i \notin \mathcal{I}_1$ or e_{it} if $i \in \mathcal{I}_1$) across i or t. Rather, we just have to impose mild moment conditions (see Doz et al., 2012, Bai and Li, 2016, and Barigozzi and Luciani, 2019, 2024, for details).

3.3 Setting-up the model

Before estimating the model, we need to choose the number of common factors q, the number of common trends and the number of lags p in the VAR for the common factors. As for the number of lags p, we select $p = 2$ based on the BIC criterion for a VAR on the estimated factors.

Assumptions (1)-(3), and (7) imply that the covariance matrix of the differenced data $\Delta\mathbf{y}_t = (\Delta y_{1t} \cdots \Delta y_{nt})'$ has at most q eigenvalues diverging as $n \to \infty$, with all the others staying bounded. Furthermore, the q-largest eigenvalues of the spectral density of $\Delta\mathbf{y}_t$ diverge as $n \to \infty$ at all frequencies except the zero-frequency, where the number of diverging eigenvalues corresponds to the number of common trends, while all the others stay bounded. This allows us to consistently recover the numbers of common factors and common trends via the information criteria by (Bai and Ng, 2002; Hallin and Liška, 2007; Barigozzi et al., 2021). We find evidence of $q = 4$ common factors and one common trend. Details are in Appendix C.

Finally, we need to choose which idiosyncratic component to model as random walk and for which variable to include a linear trend. For the idiosyncratic component, we employ the test proposed by Bai and Ng (2004) for the null hypothesis of an idiosyncratic unit root. To determine for which variable to include a linear trend, we test the significance of the sample mean of Δy_{it}. The results of these tests can be found in Appendix A.

4 Potential output and output gap of the Euro Area

Figure 1 presents our potential output estimate, both in 100 × log levels (left plot) and in year-on-year (YoY) growth rates (right plot). The right plot also shows estimates of potential output growth from EC and IMF. Comparing our estimates with those of the EC and IMF is useful because our approach differs from that of these institutions, as our estimate of the output gap is primarily data-driven, whereas the EC's and IMF's estimates are based mainly on theoretical macroeconomic models. Specifically, both the EC and IMF derive the output gap and potential output according to the so-called "production function approach" (Kiley, 2013): after having specified a neoclassical production function endowed with a Phillips curve and an Okun's law, the trend component of GDP is extracted, and the output gap is defined as the deviation from the trend (Havik et al., 2014).

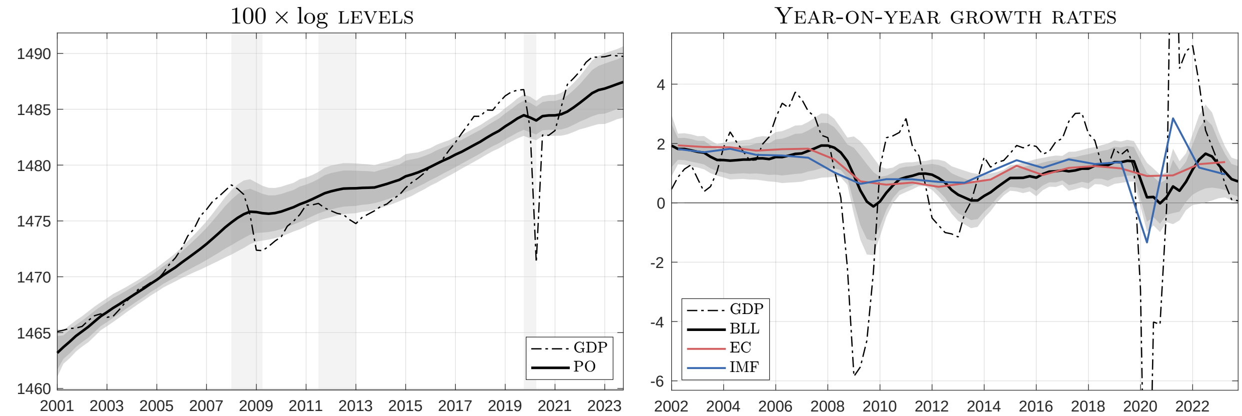
Three main results emerge from Figure 1. First, potential output growth decelerated after the GFC, and the SDR further compounded this deceleration. This result aligns with the prevailing institutional perspective, according to which the GFC had a persistent negative effect on economic capacity (Thum-Thysen et al., 2022)—potential output growth was 1.9% right before the GFC, and the subsequent peak was 1.4% right before the Covid pandemic.[3] Second, while the EC and IMF also estimate a slowdown in potential growth after the GFC (as shown by blue and red lines in the right plot), they do not estimate any effects of the SDR on potential growth. Third, in contrast with the GFC, the Covid recession had only a transitory effect on potential output growth, with potential growth averaging 1.6% in the second half of 2022 before slipping to 0.7% at the end of the sample.

Figure 2 presents the estimated output gap, together with the estimates from the EC and the IMF.[4] Our output gap estimate looks similar to those of the EC and IMF in that the dating of the turning points perfectly coincides. Moreover, the three estimates closely align until the SDR, as they all suggest a substantial overheating of the economy in the pre-GFC period, followed by a persistently negative output gap during the two recessions. However, our measure increased

[3] Our result is also consistent with Schmöller and Spitzer's (2021) interpretation that the decline in total factor productivity resulting from the GFC induced hysteresis effects which can explain the decline in potential growth.

[4] In Appendix G, we show alternative estimates of the output gap based on: the HP filter, the filter by Hamilton (2018), the boosted HP filter Phillips and Shi (2021), the Butterworth filter as recommended by Canova (2022), and the large Bayesian VAR approach by Morley et al. (2023)—overall, the Butterworth filter estimate of the output gap seems to be the most similar to ours and the most stable one. Moreover, in Appendix J, we assess the real-time reliability of our output gap estimate.

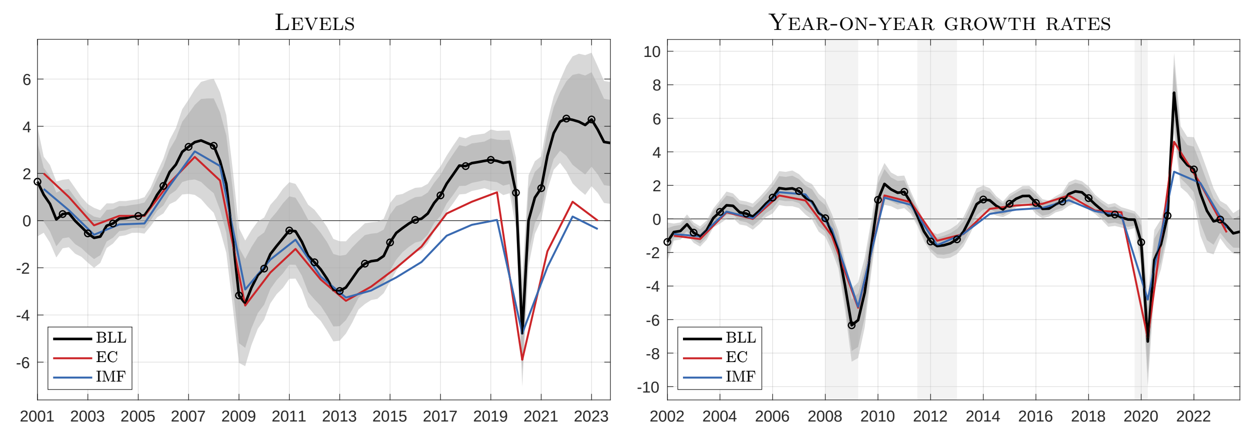
Figure 1: *Potential output.*



Notes: In all charts, the black solid line is our estimate of potential output, the grey shaded areas are the 68% and 84% confidence bands, and the dashed black line is GDP—we truncated the y-axis in the right chart for readability. In the right chart, the blue and red lines are the potential output estimates published by the European Commission and the IMF, respectively. The IMF estimate of YoY potential output growth reported in the right chart is the result of our own calculation. Indeed, the IMF publishes only an estimate of the output gap from which we backed out potential output. Thus, the blue line in the right chart does not account for any adjustment for Covid that the IMF might have done.

after the SDR, hovering at about 2% from 2017 to the Covid pandemic, thus signaling a much tighter economy than the EC or the IMF. As we will discuss in Section 6, household liabilities, whose growth rate picked up in 2014 after declining for seven years in a row, are key drivers of the output gap after the GFC (see also Borio et al., 2017).[5] This result highlights how considering many variables beyond the usual suspects (i.e., inflation indexes and labor market indicators) is important to estimate the cyclical position of the economy (Buncic and Pagan, 2022).

Figure 2: *Output gap.*



Notes: The black line is our estimate of the output gap (OG) in levels (left plot) and YoY growth rates (right plot)—the level of the output gap is the percentage deviation from potential, the YoY growth rates is $\mathrm{OG}_t - \mathrm{OG}_{t-4}$. Each black marker denotes one year (four quarters), starting from 2001:Q1. The grey shaded areas are the 68% and 84% confidence bands. The red and blue lines are the output gap estimates published by the European Commission and the IMF, respectively.

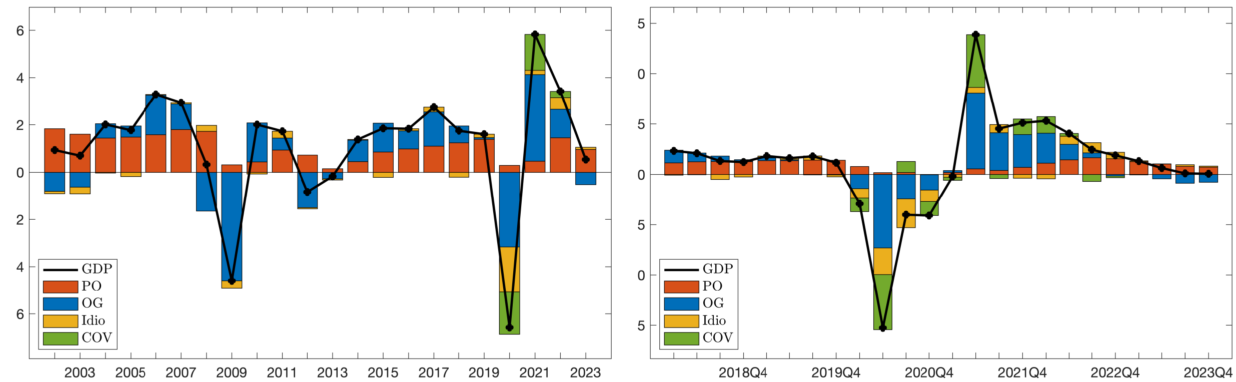
Figure 3 decomposes GDP growth. During the Covid pandemic, the output gap subtracted 7.3 percentage points (p.p.) from YoY GDP growth in 2020:Q2. At a YoY GDP growth rate of -15.3%, a -7.3 p.p. contribution from the output gap seems dubious. However, as explained in Section 3.2,

[5] In line with our result, Gambetti and Musso (2017) find that loan supply shocks have a significant effect on the EA business cycle.

our output gap estimate captures only standard co-movements, while the Covid factor captures the additional co-movements brought about by the Covid shock. As shown by the green bars in the right plot in Figure 3, the Covid factor accounts for an additional -5.5 p.p. of the YoY GDP plunge in 2020:Q2.

FIGURE 3: *Decomposition of GDP growth*

ANNUAL YEAR-ON-YEAR GROWTH RATE | QUARTERLY YEAR-ON-YEAR GROWTH RATES



NOTES: The black line with dot markers is GDP growth. The bars represent the contribution of each component to GDP growth rate of GDP. The left plot shows YoY growth for the full year, while the right plot shows YoY growth in each quarter.

In 2021 and 2022, GDP growth was driven by the output gap and the Covid factor, while in 2023, GDP growth was sustained by potential output. However, as shown in the left plot, potential output growth reached only about half its pre-GFC average pace of around 1.6%. This result, coupled with the left plot in Figure 2 showing the output gap at 3 p.p. above potential at the end of 2023, suggests that there is a potential output issue in the EA, not a business cycle issue. Consequently, European countries should implement supply-side structural reforms with long-term effects to boost growth, as policies aimed solely at stimulating aggregate demand are likely to yield short-lived effects.

5 What about the Okun's law and the Phillips curve?

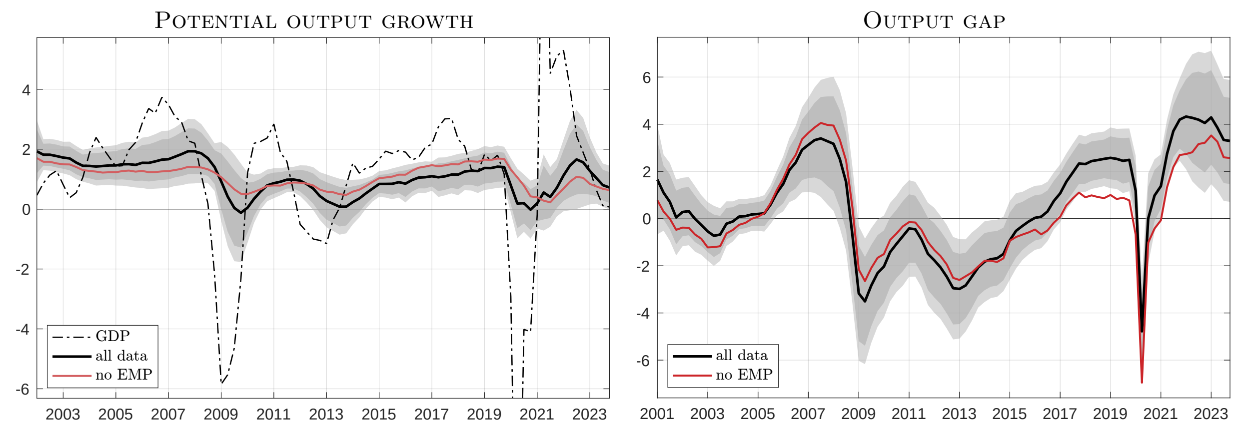
Our estimate of the output gap has a different meaning than the production-function-based model estimate. In production-function-based models, the output and unemployment gaps are related through the Okun's law. Thus, in these models, the labor and the goods and services markets are tightly related. As a result, the unemployment gap decreases whenever the output gap increases, and vice-versa. Likewise, in these models, the output gap is related to inflation through the Phillips curve. As a result, these models take considerable signal from inflation data to estimate the output gap: the absence of inflation typically suggests a negative output gap, while high inflation indicates a positive gap; similarly, a decrease in inflation generally corresponds to a reduction in the output gap, and vice-versa.

In our model (like any statistical model), there is no Okun's law or Phillips curve, just the data. The output gap is the part of GDP driven by the common cyclical component, and as such, it is about whether or not the recent GDP growth pace is sustainable in the long run. Thus, when we compare our estimate with those of the IMF or the EC, we are not really comparing apples to apples. Nonetheless, this comparison is useful because whenever these measures differ, inspecting why they differ can be very useful to better understand what happened (or is happening) in the economy. Thus, in this section and the next, we try to open the black box to understand what signals the model takes from the data and how these signals affect our estimate of the output gap. In this section, we will focus on labor market indicators (and so, the Okun's law) and on price inflation indicators (and so, the Phillips curve); in the next section, we will focus on credit variables, which production-function-based models usually do not consider.

5.1 The Okun's law

Figure 4 shows what our output gap estimate would be if we remove all 18 labor market indicators included in our dataset. As can be seen, our output gap estimate would be much lower than our baseline estimate between 2015 and the end of the sample. Indeed, since 2012, the unemployment rate has constantly decreased from just over 12% to just under 7%, except between 2020 and 2021 during the Covid lockdowns and mobility restrictions. If not fed with this information, the model would have concluded that the economy was less tight than our benchmark estimate.

FIGURE 4: *Output gap and potential output growth when excluding labor market indicators*



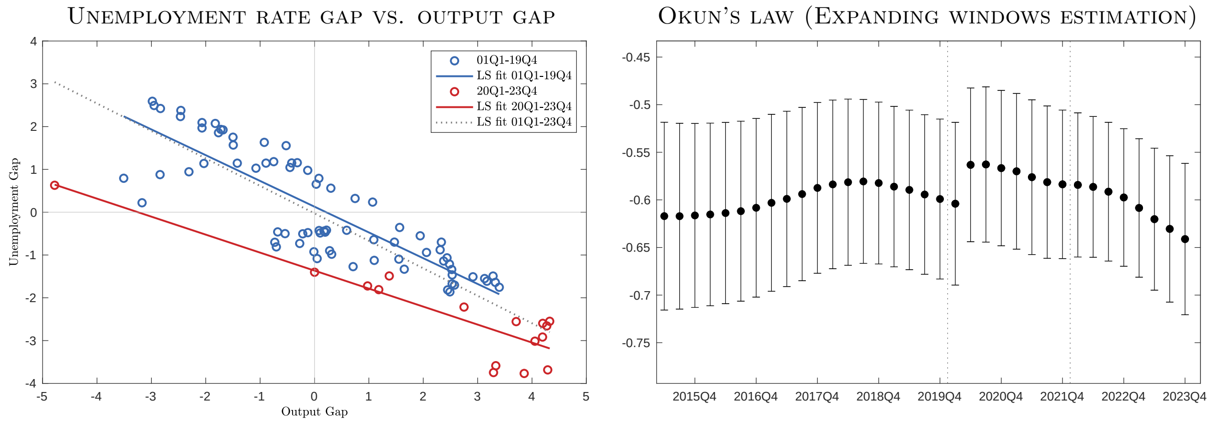
NOTES: The black solid line is our benchmark estimate and the grey shaded areas are the 68% and 84% confidence bands, the black dashed line is GDP YoY growth rate, the red lines are estimates obtained by excluding labor market indicators from the dataset.

To further corroborate the intuition that there is a tight relationship between our output gap estimate and labor market indicators, the left chart in Figure 5 shows that in our model, the unemployment rate gap and the output gap are negatively correlated; that is, our model captures the Okun's law relationship in the data. As shown by the slope of the least squares fit line in the chart, on average, for every percentage point increase in the output gap, the unemployment gap decreases 0.6 p.p. Moreover, as shown by the similar slope of the pre-Covid (blue) and post-Covid (red) least squares fit lines, this relationship has remained stable over time. We reach a similar conclusion even when we estimate an Okun's law regression of the output gap on the unemployment rate using an expanding window, as shown in the right chart.[6]

Lastly, Figure 6 shows the Generalized Impulse Response Functions (GIRFs) of the common component of the unemployment rate, GDP, potential output, and the output gap to a 1 p.p. shock

[6] We also estimated the relationship between the output gap and hours worked gap and find a positive relationship in line with the results of Morley et al. (2023). Moreover, we find this relation to be stable in time.

FIGURE 5: *Okun's Law*

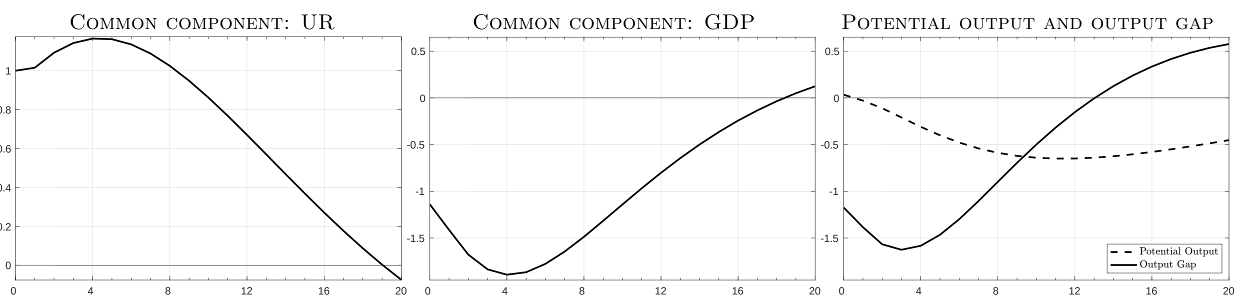


NOTES: The left chart shows the Okun's law relationship with the output gap on the horizontal axis and the unemployment rate gap on the vertical axis. Each circle correspond to an output gap - unemployment gap pair at time t. The blue circles refer to pre-Covid observations, the red circles refer to post-Covid observations. The dotted/blue/redlines are the least squares fit lines.

The right chart shows the least squares estimate, based on an expanding window starting from 2015:Q1, of the Okun's law slope β given by the regression $(\mathrm{UR}_t - D_{\mathrm{UR},t}) = \alpha + \beta \mathrm{OG}_t + \varepsilon_{\mathrm{UR},t}$, where $D_{\mathrm{UR},t}$ is the time-varying mean of the unemployment rate defined in (2). Each dot is an estimate of α while the whiskers are $\pm$ one HAC standard errors.

to the common component of the unemployment rate (Crump et al., 2021).[7] Results confirm that our model, on average, associates an unemployment rate increase with an output gap decrease. The common component of the unemployment rate remains 1 p.p. (or more) above the baseline for about a year and a half before decreasing and slowly returning to zero. In response, GDP decreases and keeps decreasing, reaching a through a year after the shock; then, it slowly returns to baseline. The model attributes most of the GDP response to movements in the output gap, while potential output slightly decreases only after a few quarters. The shock is fully absorbed in about 4 years.

FIGURE 6: *Generalized Impulse Response Functions to a shock to the unemployment rate*



NOTES: The black solid/dashed lines are the GIRFs to a 1 p.p. shock to the common component of the unemployment rate. The major ticks in the x-axis represent quarters after the shock.

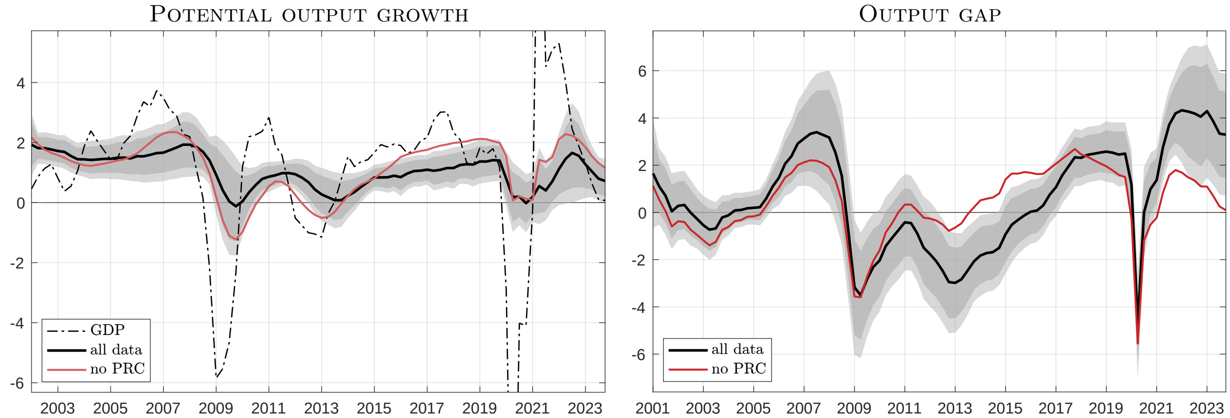
5.2 The Phillips curve

Figure 7 shows our estimated output gap after removing all 15 inflation indicators (i.e., price indexes, and natural gas and the oil prices) from the dataset. As can be seen, our output gap estimate would

[7] The lag-h GIRF of all variables is obtained by computing the differences between the h-step ahead forecast of their common component conditional on a shock to a given variable at time $T+1$ minus the h-step ahead unconditional forecast of the common component, i.e., when no shock is imposed. Both forecasts are computed conditional on all information available at time T (the last observation in our sample) by means of the Kalman filter (Bańbura et al., 2015).

have indicated a much tighter economy after the GFC and a less tight economy in 2022 and 2023 had we not included inflation indicators. In other words, our model interprets the low inflation after the GFC as a signal that there is slack in the economy (see also Jarociński and Lenza, 2018) and the high inflation in 2022 and 2023 as a signal that the economy is tight, even accounting for the surge in oil and natural gas prices.

FIGURE 7: *Output gap and potential output growth when excluding price indexes*



NOTES: The black solid line is our benchmark estimate and the grey shaded areas are the 68% and 84% confidence bands, the black dashed line is GDP YoY growth rate, the red lines are estimates obtained by excluding price indexes from the dataset.

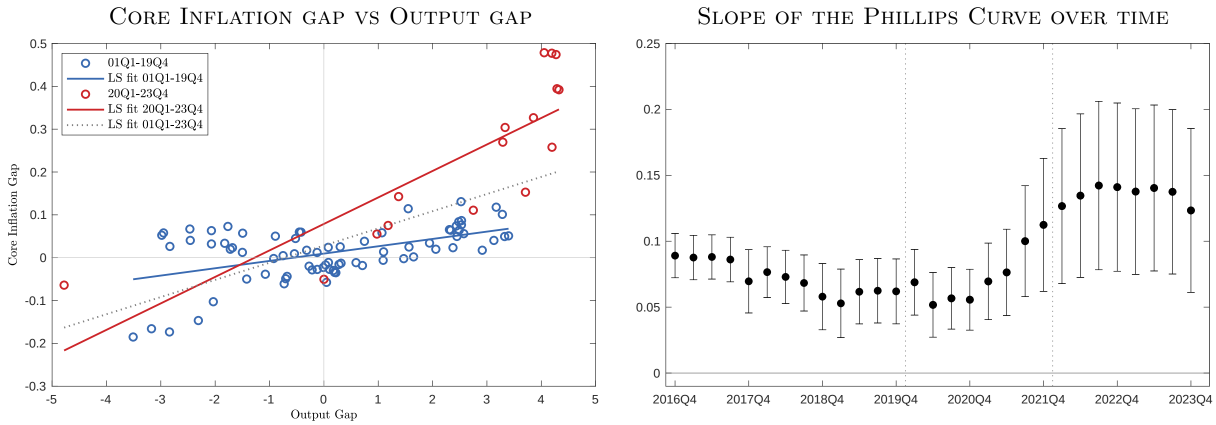
To further corroborate the intuition that there is a relationship between our output gap estimate and inflation indicators, the left plot in Figure 8 shows that in our model, the core inflation rate gap (i.e., the cyclical common component of core inflation) and the output gap are positively correlated; that is, there is Phillips correlation in the data and our model captures it. As shown by the slope of the least squares fit line in the chart, on average, for every percentage point increase in the output gap, the core inflation gap increases 4 basis points. Moreover, as shown by the difference between the slope of the pre-Covid (blue) and post-Covid (red) least squares fit lines, this correlation has increased significantly after Covid (the slope of the fit line increases from 0.017 to 0.062). The right plot in Figure 8, which shows expanding window estimates of the slope of the Phillips Curve, together with the respective standard errors, confirms that the relationship between inflation and the output gap has strengthened after Covid.

Lastly, Figure 9 shows the GIRFs of the common component of core inflation, GDP, potential output, and the output gap to a 0.5 p.p. shock to the common component of core inflation. Results confirm that our model, on average, associates an increase in inflation with an output gap increase. The GIRF of the common component of core inflation peaks one quarter after the shocks before decreasing and slowly returning to zero. In response, GDP initially increases, but then after about a year, it starts decreasing, reaching a trough about 2 years after the shock—the shock is fully absorbed in 5 years. The response of potential output is negative and persistent. The output gap initially increases, then decreases, and increases again before returning to zero.

6 The role of credit indicators

Borio et al. (2017) argue that incorporating financial indicators in the dataset is necessary to obtain meaningful estimates of the business cycle because credit expansions tend to overheat the economy,

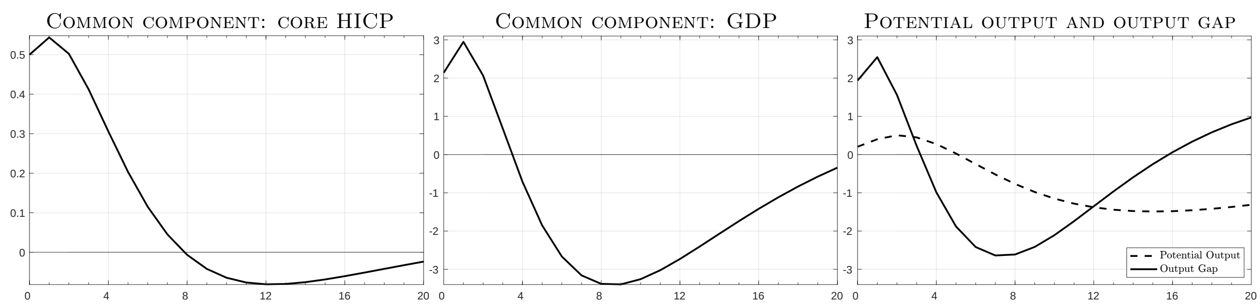
Figure 8: *Phillips Curve*



Notes: The left chart shows the Phillips Curve relationship with the output gap on the horizontal axis and the core inflation gap on the vertical axis. Each circle correspond to an output gap - core inflation gap pair at time t. The blue circles refer to pre-Covid observations, the red circles refer to post-Covid observations. The dotted/blue/red lines are the least squares fit lines.

The right chart shows the least squares estimate, based on an expanding window starting from 2015:Q1, of the slope of the Phillips Curve given by the following expectation-augmented specification (e.g., Conti, 2021): $\pi_t = c + \alpha \mathrm{OG}_t + \beta\pi_{t-1} + \gamma\mathbb{E}\pi_{t+k} + \varepsilon_{\pi,t}$, where π_t is core inflation and $\mathbb{E}\pi_{t+k}$ are the long-run (5-year ahead) inflation expectations in the Survey of Professional Forecasters. Each dot is an estimate of α while the whiskers are $\pm$ one HAC standard errors.

Figure 9: *Generalized Impulse Response Functions to a shock to core inflation*

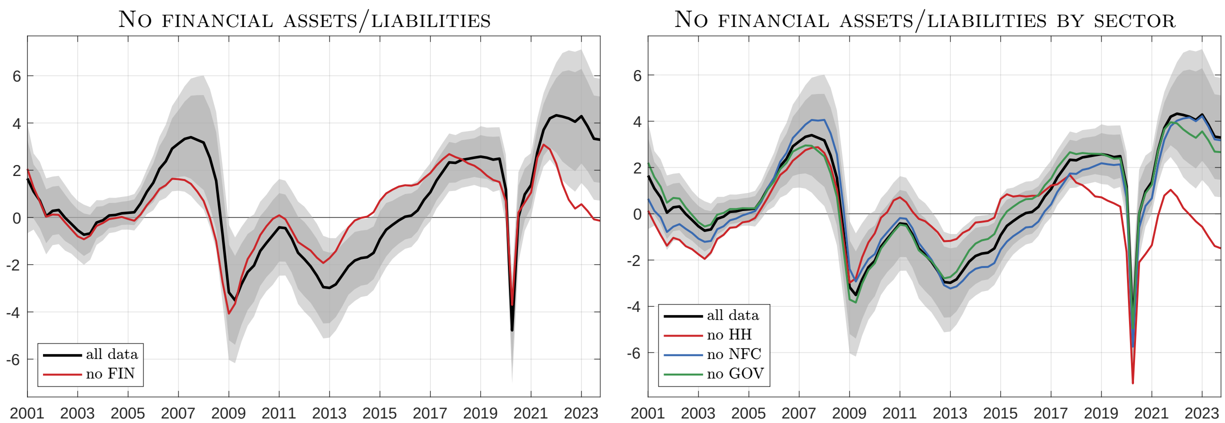


Notes: The black solid/dashed lines are the GIRF to a 0.5 p.p. shock to the common component of core inflation. The major ticks in the x-axis represent quarters after the shock.

particularly from the late 1990s onward. In support of Borio et al.'s results, Berger et al. (2022) find that a large share of the US economy overheating in the build-up of the financial crisis was due to financial imbalances in the credit and housing market. Moreover, Claessens et al. (2012), Rünstler and Vlekke (2018), and Winter et al. (2022) find that the business and financial cycles are correlated and co-move in the medium run, thus supporting the idea of a "medium-term" business cycle (Comin and Gertler, 2006). Our dataset includes various indicators of financial conditions, including credit indicators, monetary aggregates, and house prices, totaling 26 variables. What is the role of those financial variables (and in particular, credit indicators) in our model? Does our model support the view that the financial and business cycles are correlated? Out of its intrinsic interest, answering this question is very important because it may explain why our output gap estimate differs from those of the IMF and EC, which do not consider credit indicators.

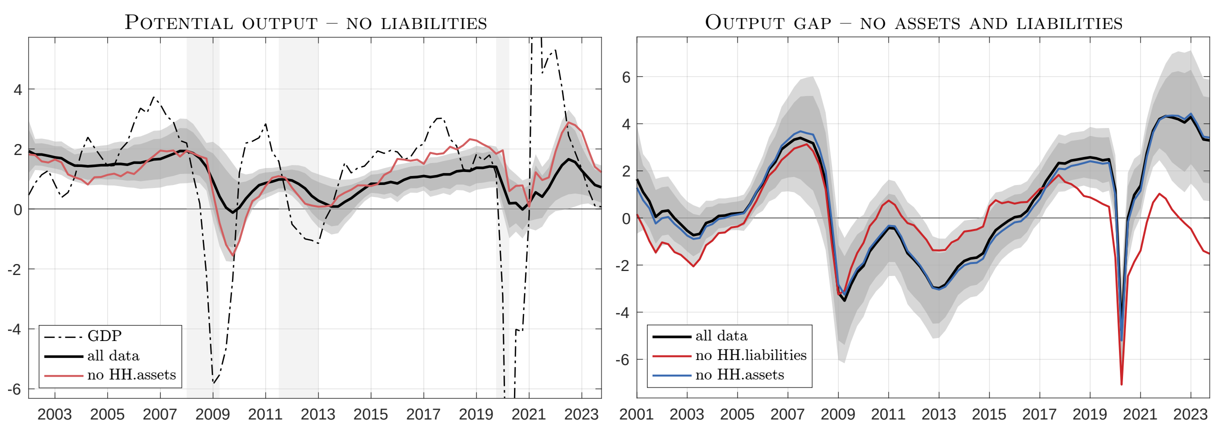
As shown in the left plot in Figure 10, removing all 26 financial variables from our dataset yields significant consequences. Is this outcome primarily due to removing many variables from the dataset? Or is there a specific sub-group of financial variables that carries more weight? The right plot in Figure 10 conclusively shows that household credit indicators are the primary driver of this result. Additionally, the left plot in Figure 11 clarifies that it is household liabilities that matter for the output gap estimation, while household assets do not have any significant effect. The results in Figures 10 and 11 provide further evidence that household leverage is an important driver of the business cycle, a trend that emerged in the early 2000s in many advanced economies (Mian et al., 2017). Indeed, while in the 1990s, non-financial corporations were the main driver of the financial cycle, households were behind both the pre-GFC excess leverage, which boosted households' demand, and the subsequent deleveraging, which curtailed household's demand (Mian and Sufi, 2018; Plagborg-Møller et al., 2020; Reichlin et al., 2020).

Figure 10: *Output gap excluding financial financial assets/liabilities or their sub-groups*



Notes: The black line is our estimate of the output gap in levels, the grey shaded areas are its 68% and 84% confidence bands, and the red/blue/green lines are the estimates obtained omitting all the information from: financial assets/liabilities (FIN); assets/liabilities of households (HH); assets/liabilities of Non-Financial Corporations (NFC); assets/liabilities of Government (GOV).

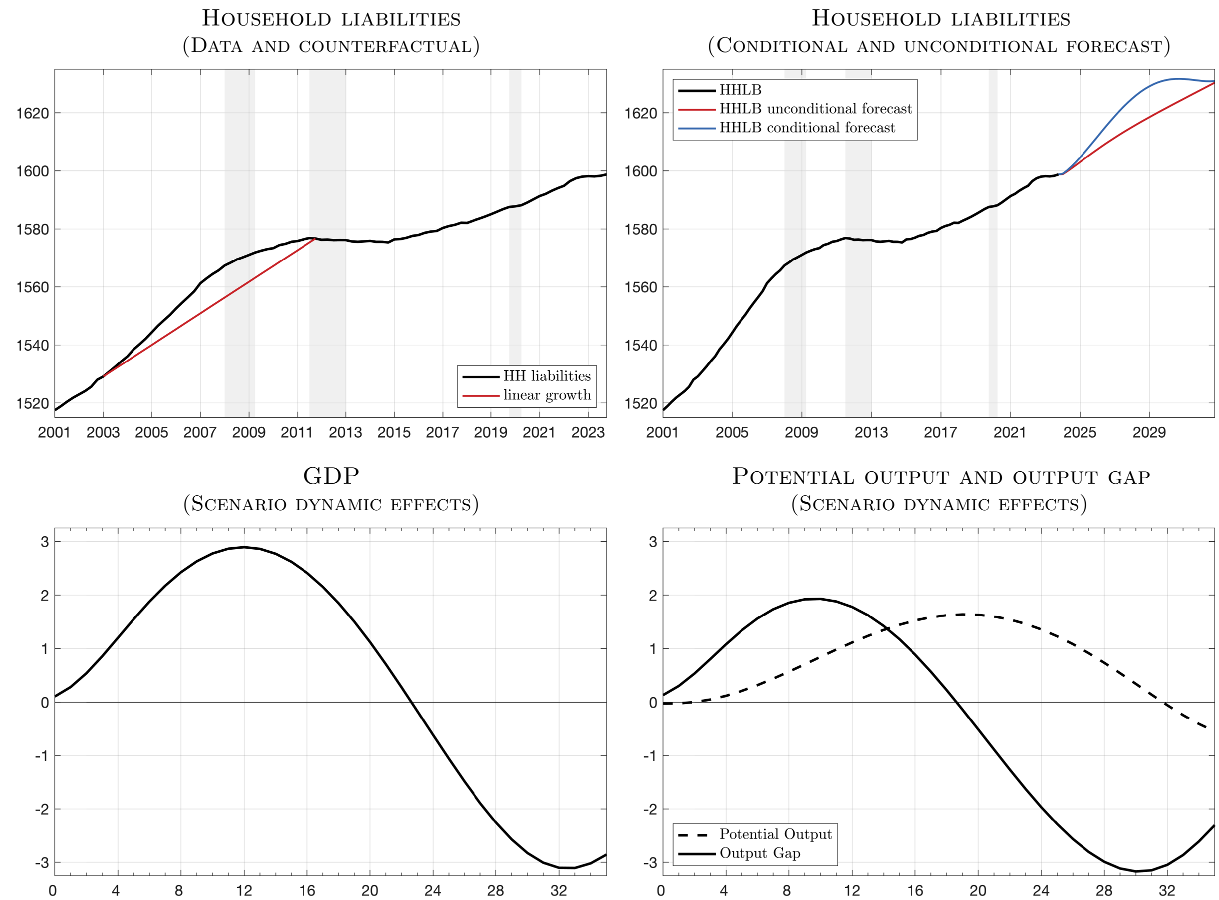
Figure 11: *Potential output and output gap excluding households assets/liabilities*



Notes: In the left chart, the black line is our estimate of the output gap in levels, the grey shaded areas are its 68% and 84% confidence bands, the red line is the estimate obtained omitting the information from all liabilities of households (HH.liabilities) and the blue line is the estimate obtained omitting the information from all assets of households (HH.assets). In the right chart, the black solid line is our estimate of year-on-year potential output growth, the grey shaded areas are the 68% and 84% confidence bands, the black dashed line is GDP YoY growth rate, and the red line is the estimate obtained omitting all the information from all liabilities of households (HH.liabilities).

In summary, credit indicators, notably household liabilities, have a crucial role when estimating the output gap—this is confirmed by the inflation forecasting exercise presented in Section 7. When we exclude these variables, in periods of debt build-up (e.g., before the GFC), we estimate a lower output gap, while in periods of deleveraging (e.g., from 2013 to 2016), we estimate a higher output gap. The scenario analysis in Figure 12 confirm this interpretation of the results in Figure 11.

FIGURE 12: *Scenario analysis*



NOTES: In the upper-left chart, the black line is the data (in 100×log-levels), and the red line is a linear path starting from the value of household liabilities in 2003:Q1 and ending in 2011:Q4. In the upper-right chart, the black line are the data,the blue line is the scenario we simulate, and the red line is the forecast of household liabilities when no alternative scenario is imposed. In the lower charts, the black solid/dashed lines are the dynamic effects of the simulated scenario.

Figure 12 shows the impact of a scenario in which household liabilities increase faster than in the baseline for about 3½ years—reaching a level about 6¼ p.p. higher than in the baseline—and then return to baseline after about 8 years. To calibrate this scenario, we looked at the difference between the actual times series of household liabilities between 2003:Q1 and 2011:Q4 against the counterfactual had household liabilities grew linearly in this period (upper-left chart of Figure 12),[8] and then smoothed it with a 5-th order degree polynomial. We then added this smoothed path to the unconditional forecast of household liabilities (the red line in the upper-right chart in Figure

[8] Between 2003:Q1 and 2011:Q4, household liabilities grew at an average annualized growth rate of 5.4%. This is the result of period of increased leverage between 2003 and 2008:Q1, when household liabilities grew at an average annualized growth rate of 7.6%, and a period of slower growth between 2008:Q1 and 2011:Q4, when household liabilities grew at an average annualized growth rate of 2.5%.

12), which gives us a path for household liabilities (blue line) conditional on which we can obtain forecasts for all the variables in the model.

The lower charts in Figure 12 show the dynamic effects of this scenario on the log level of GDP and on the output gap and potential output (Crump et al., 2021).[9] Specifically, GDP increases for a little over three years, reaching a peak at 3 p.p. above the baseline. Since then it starts declining and after six years turns negative. The response of the output gap mimics that of GDP, but it is a little faster. Potential output slowly increases for the first five years and then returns to the baseline. These results show that growth financed through household debt is not sustainable in the long run.

7 Has our output gap measure predictive power for inflation?

The output gap is often seen as an indicator of current/future inflationary pressure (or easing). Thus, predicting inflation is considered a *must-have* property for any output gap measure. This is the case, even though, as opposed to production-function-based model estimates of the output gap, our measure is not directly tied to inflation and was not designed to be an inflation gauge.

To assess the forecasting properties of our output gap estimate, we replicate the analysis conducted in Bańbura and Bobeica (2023), and we employ a simple autoregressive distributed lag model:

$$\pi_{t+4} = \alpha \pi_t + \beta OG_t + v_{t+4}, \tag{10}$$

where $\pi_t = 100 \log(P_t/P_{t-1})$ is the quarter-on-quarter inflation rate in quarter t, P_t is the harmonized consumer price index (either headline or core), $\pi_{t+4} = \sum_{i=1}^{4} \pi_{t+i}$ is year-on-year inflation in quarter $t+4$, and OG_t is the output gap. Bańbura and Bobeica (2023) labeled model (10) the "benchmark model," and they show that, despite being very simple, it delivers decent forecasts compared to more complex alternative models.

Our exercise compares the inflation forecast obtained using the benchmark model (10) with a forecast obtained by replacing our estimate of the output gap with different univariate and multivariate statistical models (see Appendix G for a description of the alternative models). This comparison is carried out through an expanding window exercise, where the first window is a 60-quarter window. We look at the forecasting performance of the different output gap measures over two distinct samples: a pre-Covid sample, 2015:Q4–2019:Q4, and a post-Covid sample, 2022:Q1–2023:Q4.

Table 1 compares the forecasting performance of the different output gap measures in terms of relative Root Mean Squared Error (RMSE)—numbers lower than one indicate a better forecasting performance when using our output gap estimate. As shown in rows (1)-(8), in the pre-Covid period, when inflation was low and stable, our output gap measure performed better than all the other alternatives in forecasting headline inflation and better than most alternatives when forecasting core inflation. However, in the post-Covid period, when inflation surged and then declined, our model outperformed all the other measures, sometimes substantially.

Row (9) in Table 1 compares the forecasting performance of two different estimates of the output

[9] The dynamic effects of the simulated scenario at a lag h on all variables are computed as the difference between the h-step ahead forecast of their common component conditional on the simulated path of household liabilities minus the h-step ahead unconditional forecast of the common component, i.e., when no alternative path is simulated. Both forecasts are computed conditional on all information available at time T (the last observation in our sample) by means of the Kalman filter (Bańbura et al., 2015).

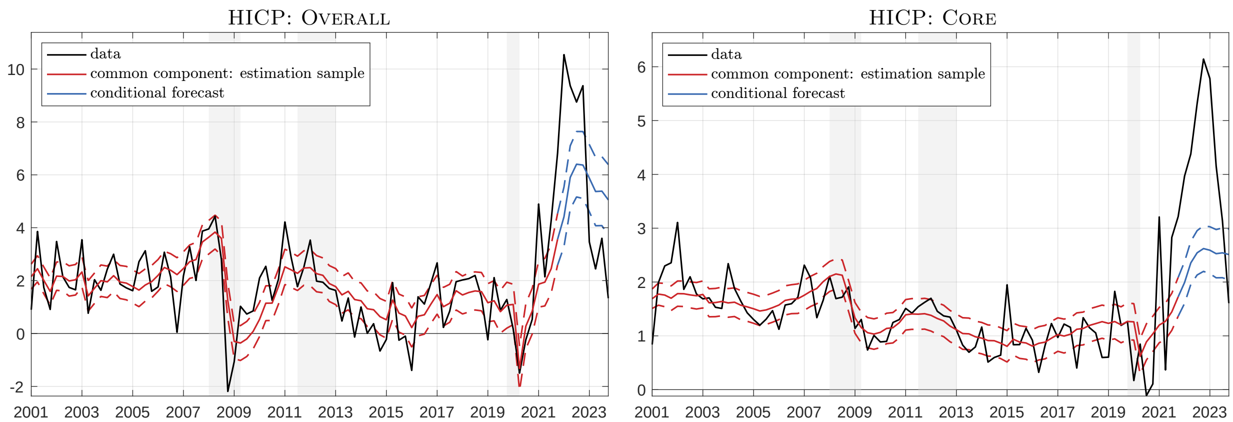
gap obtained with our model: the benchmark estimate and the one obtained excluding all credit indicators. The results in Table 1 add evidence to our claim that including credit variables in the dataset when estimating the output gap is crucial, as the model including credit indicators outperforms the model excluding credit indicators.

TABLE 1: *4-quarter ahead year-over-year inflation forecasting*
Relative Root Mean Squared Errors

	Output Gap Measure	2015:Q4-2019:Q4 Headline	2015:Q4-2019:Q4 Core	2022:Q1-2023:Q4 Headline	2022:Q1-2023:Q4 Core
(1)	HP Filter ($\lambda = 1600$)	0.91	1.02	0.81	0.96
(2)	HP Filter ($\lambda = 51200$)	0.97	0.97	0.90	0.97
(3)	Hamilton Filter	0.88	0.83	0.89	0.85
(4)	Boosted HP Filter ($\lambda = 1600$)	0.88	0.95	0.71	0.76
(5)	Boosted HP Filter ($\lambda = 51200$)	0.88	1.00	0.74	0.80
(6)	Christiano-Fitzgerald Filter	0.89	1.00	0.85	0.83
(7)	Butterworth Filter	0.93	0.95	0.83	0.78
(8)	Multivariate Beveridge-Nelson	0.90	0.97	-	-
(9)	No household	0.95	0.98	0.93	0.94

NOTES: The table shows the relative RMSE of forecasting year-on-year inflation using (10), where OG_t is either our output gap estimate, or an alternative output gap estimate. Our benchmark output gap estimate is always the numerator of the RMSE, thus numbers lower than 1 indicate a better forecasting performance when using our benchmark output gap estimate. Rows (1)–(8) compare our benchmark estimate with alternative models, while row (9) compares our benchmark estimate with the one obtained with our model but excluding all credit indicators. In row (8) forecasts are obtained with the output gap measure by Morley et al. (2023) which is available only until 2021:Q3. Therefore, we only present results for the pre-Covid forecasting exercise.

FIGURE 13: *Conditional forecasts of inflation*



NOTES: In both charts, the black line is the quarter-on-quarter annualized inflation, the red line is the common component estimated in-sample, and the blue line is the conditional forecast. The dashed lines are the 84% confidence intervals.

To conclude, having established that our model captures the Phillips correlation in the data and is capable of forecasting inflation, we use our model to understand how much of the inflation surge and deceleration is driven by macroeconomic co-movement. To this end, we produce Kalman filter based forecasts of the common component of inflation conditional on observing at each point in time all the variables in the model except price indicators (Bańbura et al., 2015).[10] The conditional forecasts in Figure 13 show that the macroeconomic co-movement can only account for some of the

[10] The choice of where to start the exercise requires some explanation. The first two consecutive quarter-on-quarter (qoq) at an annual rate (a.r.) core inflation reading above 2% occurred in the second half of 2021. After that, from

inflation acceleration and very little of the inflation deceleration. Thus, according to our model, the 2023 inflation deceleration was not due to the aggregate demand cooling down (as indicated by our almost flat output gap estimate in 2023) but to other factors, such as the expansion of aggregate supply or the normalization of inflation expectations.

8 Conclusions

This paper proposes a new measure of potential output and the output gap for the EA based on letting a large number of macroeconomic and financial indicators speak. To do so, we estimate a large-dimensional non-stationary dynamic factor model, which allows us to capture co-movements across series while incorporating relevant macroeconomic priors, such as the long-run decline in output growth. Our model is a modified version of the model Barigozzi and Luciani (2023) used to estimate the output gap in the US, which we upgraded so that it can also be estimated on short samples and can handle the anomalous dynamics induced by the Covid pandemic.

Our output gap estimate is in line with those published by the EC and the IMF in most of the sample. However, our estimate diverges significantly after the 2011–2012 sovereign debt recession when our output gap measure suggests that the EA economy was tighter than estimated by the EC and the IMF. This result suggests that the EA has a potential output issue, not a business cycle issue. Hence, if the goal is to achieve better economic conditions in the EA, its member countries should implement supply-side structural reforms that have long-run effects, while policies aiming at stimulating aggregate demand will have only short-term effects at best.

Moreover, we find that incorporating financial indicators in the dataset, particularly household liabilities, is necessary to pin down the output gap. Excluding these variables leads to estimating a lower output gap in periods of debt build-up (e.g., before the GFC) and a higher gap in periods of deleveraging (e.g., from 2013 to 2016), as growth financed through household debt is not sustainable in the long run.

References

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

Ahn, H. J. and M. Luciani (2024). Common and idiosyncratic inflation. FEDS 2020-024r1, Board of Governors of the Federal Reserve System.

Aruoba, S. B., F. X. Diebold, J. Nalewaik, F. Schorfheide, and D. Song (2016). Improving GDP measurement: A measurement-error perspective. *Journal of Econometrics 191*, 384–397.

Bai, J. and K. Li (2016). Maximum likelihood estimation and inference for approximate factor models of high dimension. *Review of Economics and Statistics 98*(2), 298–309.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*, 191–221.

2022:Q1, we had six consecutive qoq a.r. core inflation reading above 4%. As a result, we chose to start our exercise at the beginning of 2022. Qualitatively, the conclusion remains unchanged if we begin the exercise in 2022:Q3. However, quantitatively, there is a difference, as our peak forecast is about 1 p.p. lower. Nevertheless, this difference can almost entirely be attributed to the different starting points of the two forecasts..

Bai, J. and S. Ng (2004). A PANIC attack on unit roots and cointegration. *Econometrica 72*, 1127–1177.

Bańbura, M. and E. Bobeica (2023). Does the Phillips curve help to forecast euro area inflation? *International Journal of Forecasting 39*, 364–390.

Bańbura, M., D. Giannone, and M. Lenza (2015). Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections. *International Journal of Forecasting 31*(3), 739–756.

Barigozzi, M., M. Lippi, and M. Luciani (2021). Large-dimensional dynamic factor models: Estimation of impulse-response functions with $I(1)$ cointegrated factors. *Journal of Econometrics 221*, 455–482.

Barigozzi, M. and M. Luciani (2019). Quasi maximum likelihood estimation of Non-Stationary Large Approximate Dynamic Factor Model. arXiv 1910.09841.

Barigozzi, M. and M. Luciani (2023). Measuring the Output Gap using Large Datasets. *The Review of Economics and Statistics 105*, 1500–1514.

Barigozzi, M. and M. Luciani (2024). Quasi maximum likelihood estimation and inference of large approximate dynamic factor models via the EM algorithm. arXiv 1910.03821v5.

Berger, T., J. Richter, and B. Wong (2022). A unified approach for jointly estimating the business and financial cycle, and the role of financial factors. *Journal of Economic Dynamics and Control 136*, 104315.

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

Borio, C., P. Disyatat, and M. Juselius (2017). Rethinking potential output: embedding information about the financial cycle. *Oxford Economic Papers 69*, 655–677.

Buncic, D. and A. Pagan (2022). Discovering stars: Problems in recovering latent variables from models. SSRN working paper 4220302.

Burlon, L. and P. D'Imperio (2020). Reliable real-time estimates of the euro-area output gap. *Journal of Macroeconomics 64*, 103191.

Canova, F. (2022). FAQ: How do I estimate the output gap? mimeo.

Carriero, A., T. E. Clark, M. Marcellino, and E. Mertens (2022). Addressing COVID-19 outliers in BVARs with stochastic volatility. *The Review of Economics and Statistics*. available online.

Cette, G., J. Fernald, and B. Mojon (2016). The pre-Great Recession slowdown in productivity. *European Economic Review 88*, 3–20.

Claessens, S., M. A. Kose, and M. E. Terrones (2012). How do business and financial cycles interact? *Journal of International Economics 87*, 178–190.

Comin, D. and M. Gertler (2006). Medium-term business cycles. *The American Economic Review 96*.

Conti, A. M. (2021). Resurrecting the Phillips curve in low-inflation times. *Economic Modelling 96*, 172–195.

Crump, R. K., S. Eusepi, D. Giannone, E. Qian, and A. M. Sbordone (2021). A large bayesian var of the united states economy. *FRB of New York Staff Report* (976).

De Masi, P. (1997). IMF estimates of potential output: Theory and practice. IMF Working Paper 177.

Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2007). On the fit of new Keynesian models. *Journal of Business & Economic Statistics 25*, 123–143.

Doz, C., D. Giannone, and L. Reichlin (2012). A quasi-maximum likelihood approach for large, approximate dynamic factor models. *The Review of Economics and Statistics 94*, 1014–1024.

European Commission (2018). Staff working document on the review of the flexibility under the Stability and Growth Pact. COM(2018) 335.

Furlanetto, F., P. Gelain, and M. T. Sanjani (2021). Output gap, monetary policy trade-offs, and financial frictions. *Review of Economic Dynamics 41*, 52–70.

Gambetti, L. and A. Musso (2017). Loan supply shocks and the business cycle. *Journal of Applied Econometrics 32*, 764–782.

González-Astudillo, M. (2019). An output gap measure for the euro area: Exploiting country-level and cross-sectional data heterogeneity. *European Economic Review 120*, 103301.

Hallin, M. and R. Liška (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association 102*, 603–617.

Hamilton, J. D. (2018). Why you should never use the Hodrick-Prescott filter. *The Review of Economics and Statistics 100*, 831–843.

Hartl, T., R. Tschernig, and E. Weber (2022). Solving the unobserved components puzzle: A fractional approach to measuring the business cycle. mimeo.

Hasenzagl, T., F. Pellegrino, L. Reichlin, and G. Ricco (2022). A Model of the Fed's View on Inflation. *The Review of Economics and Statistics 104*, 686–704.

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

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

Justiniano, A., G. Primiceri, and A. Tambalotti (2013). Is there a trade-off between inflation and output stabilization? *American Economic Journal: Macroeconomics 5*, 1–31.

Kamber, G., J. Morley, and B. Wong (2018). Intuitive and reliable estimates of the output gap from a Beveridge-Nelson filter. *The Review of Economics and Statistics 100*, 550–566.

Kiley, M. T. (2013). Output gaps. *Journal of Macroeconomics 37*, 1–18.

Kim, C.-J. and J. Kim (2022). Trend-cycle decompositions of real GDP revisited: Classical and Bayesian perspectives on an unsolved puzzle. *Macroeconomic Dynamics 26*, 394–418.

Lenza, M. and G. E. Primiceri (2022). How to estimate a vector autoregression after March 2020. *Journal of Applied Econometrics 37*, 688–699.

Luciani, M. (2014). Forecasting with approximate dynamic factor models: The role of non-pervasive shocks. *International Journal of Forecasting 30*, 20–29.

Maroz, D., J. H. Stock, and M. W. Watson (2021). Comovement of economic activity during the Covid recession. mimeo.

McCracken, M. and S. Ng (2020). FRED-QD: A quarterly database for macroeconomic research. NBER working paper 26872.

McCracken, M. W. and S. Ng (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics 34*, 574–589.

Mian, A. and A. Sufi (2018). Finance and business cycles: The credit-driven household demand channel. *Journal of Economic Perspectives 32*, 31–58.

Mian, A., A. Sufi, and E. Verner (2017). Household debt and business cycles worldwide. *The Quarterly Journal of Economics 132*, 1755–1817.

Morley, J., C. R. Nelson, and E. Zivot (2003). Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different? *The Review of Economics and Statistics 85*, 235–243.

Morley, J., D. Rodríguez-Palenzuela, Y. Sun, and B. Wong (2023). Estimating the euro area output gap using multivariate information and addressing the COVID-19 pandemic. *European Economic Review 153*.

Morley, J., T. D. Tran, and B. Wong (2023). A simple correction for misspecification in trend-cycle decompositions with an application to estimating r. *Journal of Business & Economic Statistics*. available online.

Morley, J. and B. Wong (2020). Estimating and accounting for the output gap with large Bayesian vector autoregressions. *Journal of Applied Econometrics 35*, 1–18.

Ng, S. (2018). Comments on the cyclical sensitivity in estimates of potential output. *Brookings Papers on Economic Activity 49*, 412–423.

Ng, S. (2021). Modeling macroeconomic variations after COVID-19. NBER working paper 29060.

Onatski, A. and C. Wang (2021). Spurious factor analysis. *Econometrica 89*, 591–614.

Phillips, P. C. and S. Jin (2021). Business cycles, trend elimination, and the HP filter. *International Economic Review 62*, 469–520.

Phillips, P. C. and Z. Shi (2021). Boosting: Why you can use the HP filter. *International Economic Review 62*, 521–570.

Plagborg-Møller, M., L. Reichlin, G. Ricco, and T. Hasenzagl (2020). When is growth at risk? *Brookings Papers on Economic Activity Spring 2020*, 167–229.

Reichlin, L., G. Ricco, and T. Hasenzagl (2020). Financial variables as predictors of real growth vulnerability. Deutsche Bundesbank Discussion Papers 05/2020.

Rünstler, G. and M. Vlekke (2018). Business, housing, and credit cycles. *Journal of Applied Econometrics 33*.

Schmöller, M. E. and M. Spitzer (2021). Deep recessions, slowing productivity and missing (dis-) inflation in the euro area. *European Economic Review 134*, 103708.

Stock, J. H. and M. W. Watson (2016, January). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In J. B. Taylor and H. Uhlig (Eds.), *Handbook of Macroeconomics*, Volume 2, pp. 415–525. Elsevier.

Thum-Thysen, A., F. Blondeau, F. d'Auria, B. Döhring, A. Hristov, and K. Mc Morrow (2022). Potential output and output gaps against the backdrop of the COVID-19 pandemic. *Quarterly Report on the Euro Area 21*, 21–30.

Tóth, M. (2021). A multivariate unobserved components model to estimate potential output in the Euro Area: A production function based approach. ECB working paper 2523.

Winter, J. d., S. J. Koopman, and I. Hindrayanto (2022). Joint decomposition of business and financial cycles: Evidence from eight advanced economies. *Oxford Bulletin of Economics and Statistics 84*, 57–79.

Supplementary material for the paper:

Measuring the Euro Area Output Gap

Matteo Barigozzi
University of Bologna
matteo.barigozzi@unibo.it

Claudio Lissona
University of Bologna
claudio.lissona2@unibo.it

Matteo Luciani
Federal Reserve Board
matteo.luciani@frb.gov

Summary

In Appendix A, we provide full details on the dataset. In Appendix B, we formally state all assumptions of our model and provide motivation and comments for each of them. In Appendix D, we provide details of all estimation steps, and Appendix E explains how we compute the confidence bands to measure the uncertainty around our output gap estimate. Next, Appendix F compares our estimate of the output gap with the one obtained when identifying the common trend as suggested by Morley et al. (2023), and Appendix G compares it with alternative methodologies. In Appendix H, we show how our measures would change if we did not model the effect of Covid explicitly, and we also compare our measure of the output gap with the ones obtained by using a different estimate of the Covid factor or when modeling the Covid induced volatility as suggested by Lenza and Primiceri (2022). In Appendix I, we show the effect of not allowing some parameters to be time-varying. In Appendix J, we assess the reliability of our output gap estimate.

M. Barigozzi and C. Lissona gratefully acknowledges financial support from MIUR (PRIN2020, Grant 2020N9YFFE).

Disclaimer: the views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of the Board of Governors or the Federal Reserve System.

A Data Description

Table A2 provides a brief description for each of the 119 series in our dataset. Moreover, for each variable, Table A2 indicates the source, the unit of measure, the seasonal adjustment treatment, the transformation (if any), model for the deterministic component for the idiosyncratic component. Table A1 presents a glossary to proper understand the data description presented in Table A2.

All the series were retrieved starting from January 2000 up to June 2023. After dropping missing values and transforming the variables, the actual starting point for the analysis is 2001:Q1. Monthly series, which constitute around one-third of the dataset, are aggregated at the quarterly level by simple averages, hence, the sample used for the analysis is 2001:Q1-2023:Q4 ($T = 92$).

Most of the series in the dataset are available already seasonally adjusted from the source, while others, e.g., financial variables and producer price indexes, are only available not seasonality adjusted. In these cases, we deseasonalize the series using a simple dummy variable approach.

Table A1: *Glossary*

Source	**Unit**
EUR = Eurostat	CLV= Chain-linked volumes
OECD = Organization for Economic Co-operation and Development	1000-ppl= Thousands of persons
ECB = European Central Bank	1000-U = Thousands of Units
FRED = Federal Reserve Economic Data	CP = Current Prices

SA	**F**	**Trans**	**Trend**	**Idio**
NSA = No Seasonal Adjustment	Q = Quarterly	0 = No Transformation	0 = No Trend	0 = I(0)
SA = Seasonal Adjustment	M = Monthly	1 = First Differences	1 = Deterministic Trend	1 = I(1)
SCA = Seasonal and Calendar Adjustment		2 = Log-Transformations	2 = Time-Varying Trend	
MSA = Manual adjustment		3 = First-Differences in Logs	3 = Time-varying Mean	

Table A2: *Data Description: Euro Area*

N	ID	Series	Unit	SA	F	Source	Trans	Trend	Idio
		(1) **National Accounts**							
1	GDP	Real Gross Domestic Product	CLV(2015)	SCA	Q	EUR	2	2	0
2	EXPGS	Real Export Goods and services	CLV(2015)	SCA	Q	EUR	2	1	0
3	IMPGS	Real Import Goods and services	CLV(2015)	SCA	Q	EUR	2	1	0
4	GFCE	Real Government Final consumption expenditure	CLV(2015)	SCA	Q	EUR	2	1	1
5	HFCE	Real Households consumption expenditure	CLV(2015)	SCA	Q	EUR	2	1	1
6	CONSD	Real Households consumption expenditure: Durable Goods	CLV(2015)	SCA	Q	EUR	2	1	1
7	CONSND	Real Households consumption expenditure: Non-Durable Goods and Services	CLV(2015)	SCA	Q	EUR	2	1	1
8	GCF	Real Gross capital formation	CLV(2015)	SCA	Q	EUR	2	0	1
9	GCFC	Real Gross fixed capital formation	CLV(2015)	SCA	Q	EUR	2	0	1
10	GFACON	Real Gross Fixed Capital Formation: Construction	CLV(2015)	SCA	Q	EUR	2	1	1
11	GFAMG	Real Gross Fixed Capital Formation: Machinery and Equipment	CLV(2015)	SCA	Q	EUR	2	1	1
12	AHRDI	Adjusted Households Real Disposable Income	%change	SCA	Q	EUR	0	0	0
13	AHFCE	Actual Final Consumption Expenditure of Households	%change	SCA	Q	EUR	0	0	0
14	GNFCPS	Gross Profit Share of Non-Financial Corporations	Percent	SCA	Q	EUR	0	0	0
15	GNFCIR	Gross Investment Share of Non-Financial Corporations	Percent	SCA	Q	EUR	0	0	0
16	GHIR	Gross Investment Rate of Households	Percent	SCA	Q	EUR	0	0	0
17	GHSR	Gross Households Savings Rate	Percent	SCA	Q	EUR	0	0	0
		(2) **Labor Market Indicators**							
18	TEMP	Total Employment (domestic concept)	1000-ppl	SCA	Q	EUR	2	1	1
19	EMP	Employees (domestic concept)	1000-ppl	SCA	Q	EUR	2	1	1
20	SEMP	Self Employment (domestic concept)	1000-ppl	SCA	Q	EUR	2	0	1
21	THOURS	Hours Worked: Total	2015=100	SCA	Q	EUR	2	1	0
22	EMPAG	Quarterly Employment: Agriculture, Forestry, Fishing	1000-ppl	SCA	Q	EUR	2	1	0
23	EMPIN	Quarterly Employment: Industry	1000-ppl	SCA	Q	EUR	2	0	0
24	EMPMN	Quarterly Employment: Manufacturing	1000-ppl	SCA	Q	EUR	2	1	0
25	EMPCON	Quarterly Employment: Construction	1000-ppl	SCA	Q	EUR	2	0	1
26	EMPRT	Quarterly Employment: Wholesale/Retail trade, transport, food	1000-ppl	SCA	Q	EUR	2	1	1
27	EMPIT	Quarterly Employment: Information and Communication	1000-ppl	SCA	Q	EUR	2	1	1
28	EMPFC	Quarterly Employment: Financial and Insurance activities	1000-ppl	SCA	Q	EUR	2	0	1
29	EMPRE	Quarterly Employment: Real Estate	1000-ppl	SCA	Q	EUR	2	0	0
30	EMPPR	Quarterly Employment: Professional, Scientific, Technical activities	1000-ppl	SCA	Q	EUR	2	1	1
31	EMPPA	Quarterly Employment: PA, education, health ad social services	1000-ppl	SCA	Q	EUR	2	1	1
32	EMPENT	Quarterly Employment: Arts and recreational activities	1000-ppl	SCA	Q	EUR	2	1	0
33	UNETOT	Unemployment: Total	%active	SA	M	EUR	0	3	0
34	UNEO25	Unemployment: Over 25 years	%active	SA	M	EUR	0	0	1
35	UNEU25	Unemployment: Under 25 years	%active	SA	M	EUR	0	0	0
36	RPRP	Real Labour Productivity (person)	2015=100	SCA	Q	EUR	2	1	1
37	WS	Wages and salaries	CP	SCA	Q	EUR	2	1	0
38	ESC	Employers' Social Contributions	CP	SCA	Q	EUR	2	1	0
		(3) **Credit Aggregates**							
39	TAS.SDB	Total Economy - Assets: Short-Term Debt Securities	MLN€	MSA	Q	EUR	2	1	1
40	TAS.LDB	Total Economy - Assets: Long-Term Debt Securities	MLN€	MSA	Q	EUR	2	1	0
41	TAS.SLN	Total Economy - Assets: Short-Term Loans	MLN€	MSA	Q	EUR	2	1	0
42	TAS.LLN	Total Economy - Assets: Long-Term Loans	MLN€	MSA	Q	EUR	2	1	1
43	TLB.SDB	Total Economy - Liabilities: Short-Term Debt Securities	MLN€	MSA	Q	EUR	2	0	1
44	TLB.LDB	Total Economy - Liabilities: Long-Term Debt Securities	MLN€	MSA	Q	EUR	2	1	0
45	TLB.SLN	Total Economy - Liabilities: Short-Term Loans	MLN€	MSA	Q	EUR	2	1	1
46	TLB.LLN	Total Economy - Liabilities: Long-Term Loans	MLN€	MSA	Q	EUR	2	1	1
47	NFCAS	Non-Financial Corporations: Total Financial Assets	MLN€	MSA	Q	EUR	2	1	0
48	NFCAS.SLN	Non-Financial Corporations - Assets: Short-Term Loans	MLN€	MSA	Q	EUR	2	1	1
49	NFCAS.LLN	Non-Financial Corporations - Assets: Long-Term Loans	MLN€	MSA	Q	EUR	2	1	1
50	NFCLB	Non-Financial Corporations: Total Financial Liabilities	MLN€	MSA	Q	EUR	2	1	1
51	NFCLB.SLN	Non-Financial Corporations - Liabilities - Short-Term Loans	MLN€	MSA	Q	EUR	2	1	1
52	NFCLB.LLN	Non-Financial Corporations - Liabilities - Long-Term Loans	MLN€	MSA	Q	EUR	2	1	1
53	GGAS	General Government: Total Financial Assets	MLN€	MSA	Q	EUR	2	1	0
54	GGAS.SLN	General Government - Assets: Short-Term Loans	MLN€	MSA	Q	EUR	2	1	0
55	GGAS.LLN	General Government - Assets: Short-Term Loans	MLN€	MSA	Q	EUR	2	0	1
56	GGLB	General Government: Total Financial Liabilities	MLN€	MSA	Q	EUR	2	1	1
57	GGLB.SLN	General Government - Liabilities: Short-Term Loans	MLN€	MSA	Q	EUR	2	1	0
58	GGLB.LLN	General Government - Liabilities: Long-Term Loans	MLN€	MSA	Q	EUR	2	0	1
59	HHAS	Households: Total Financial Assets	MLN€	MSA	Q	EUR	2	1	1
60	HHAS.SLN	Households - Assets: Short-Term Loans	MLN€	MSA	Q	EUR	2	1	0
61	HHAS.LLN	Households - Assets: Long-Term Loans	MLN€	MSA	Q	EUR	2	1	0
62	HHLB	Households: Total Financial Liabilities	MLN€	MSA	Q	EUR	2	2	3
63	HHLB.SLN	Households - Liabilities: Short-Term Loans	MLN€	MSA	Q	EUR	2	0	1
64	HHLB.LLN	Households - Liabilities: Long-Term Loans	MLN€	MSA	Q	EUR	2	2	3

In absence of available data on durable and non-durable goods for the Euro Area, we follow Casalis and Krustev (2022) and build the aggregate series of durable consumption (CONSD) aggregating the data for the 20 individual Euro Area countries. Since data for services and non-durable goods are unavailable for many individual countries as well, we build an aggregate measure of non-durable goods (CONSND) which also includes semi-durable goods and services.

Table A2: *Data Description: Euro Area*

N	ID	Series	Unit	SA	F	Source	Trans	Trend	Idio
		(4) **Labor Costs**							
65	ULCIN	Nominal Unit Labor Costs: Industry	2016=100	SCA	Q	EUR	2	1	0
66	ULCMQ	Nominal Unit Labor Costs: Mining and Quarrying	2016=100	SCA	Q	EUR	2	1	1
67	ULCMN	Nominal Unit Labor Costs: Manufacturing	2016=100	SCA	Q	EUR	2	1	0
68	ULCCON	Nominal Unit Labor Costs: Construction	2016=100	SCA	Q	EUR	2	1	0
69	ULCRT	Nominal Unit Labor Costs: Wholesale/Retail Trade, Transport, Food, IT	2016=100	SCA	Q	EUR	2	1	0
70	ULCFC	Nominal Unit Labor Costs: Financial Activities	2016=100	SCA	Q	EUR	2	1	0
71	ULCRE	Nominal Unit Labor Costs: Real Estate	2016=100	SCA	Q	EUR	2	1	0
72	ULCPR	Nominal Unit Labor Costs: Professional, Scientific, Technical activities	2016=100	SCA	Q	EUR	2	1	0
		(5) **Exchange Rates**							
73	REER42	Real Exchange Rate (42 main industrial countries)	2010=100	NSA	M	EUR	2	0	0
74	ERUS	Exchange Rate (US dollar)	2010=100	NSA	M	EUR	2	0	0
		(6) **Interest Rates**							
75	IRT3M	3-Months Interest Rates	Percent	NSA	M	EUR	0	0	1
76	IRT6M	6-Months Interest Rates	Percent	NSA	M	EUR	0	0	1
77	LTIRT	Long-Term Interest Rates (EMU Criterion)	Percent	NSA	M	EUR	0	1	1
		(7) **Industrial Production and Turnover**							
78	IPMN	Industrial Production Index: Manufacturing	2021=100	SCA	M	EUR	2	0	1
79	IPCAG	Industrial Production Index: Capital Goods	2021=100	SCA	M	EUR	2	0	1
80	IPCOG	Industrial Production Index: Consumer Goods	2021=100	SCA	M	EUR	2	1	0
81	IPDCOG	Industrial Production Index: Durable Consumer Goods	2021=100	SCA	M	EUR	2	1	0
82	IPNDCOG	Industrial Production Index: Non Durable Consumer Goods	2021=100	SCA	M	EUR	2	1	0
83	IPING	Industrial Production Index: Intermediate Goods	2021=100	SCA	M	EUR	2	0	1
84	IPNRG	Industrial Production Index: Energy	2021=100	SCA	M	EUR	2	1	0
85	TRNMN	Turnover Index: Manufacturing	2021=100	SCA	M	EUR	2	1	1
86	TRNCAG	Turnover Index: Capital Goods	2021=100	SCA	M	EUR	2	1	1
87	TRNCOG	Turnover Index: Consumer Goods	2021=100	SCA	M	EUR	2	1	1
88	TRNDCOG	Turnover Index: Durable Consumer Goods	2021=100	SCA	M	EUR	2	0	0
89	TRNNDCOG	Turnover Index: Non Durable Consumer Goods	2021=100	SCA	M	EUR	2	1	0
90	TRNING	Turnover Index: Intermediate Goods	2021=100	SCA	M	EUR	2	0	0
91	TRNNRG	Turnover Index: Energy	2021=100	SCA	M	EUR	2	0	1
		(8) **Prices**							
92	PPICAG	Producer Price Index: Capital Goods	2021=100	MSA	M	EUR	3	0	0
93	PPIDCOG	Producer Price Index: Durable Consumer Goods	2021=100	MSA	M	EUR	3	0	0
94	PPINDCOG	Producer Price Index: Non Durable Consumer Goods	2021=100	MSA	M	EUR	3	0	0
95	PPIING	Producer Price Index: Intermediate Goods	2021=100	MSA	M	EUR	3	0	0
96	PPIFD	Producer Price Index: Food	2021=100	MSA	M	EUR	3	0	1
97	HICPOV	Harmonized Index of Consumer Prices: Overall Index	2010=100	SCA	M	ECB	3	0	1
98	HICPNEF	Harmonized Index of Consumer Prices: All Items: no Energy & Food	2010=100	SCA	M	ECB	3	0	0
99	HICPG	Harmonized Index of Consumer Prices: Goods	2010=100	SCA	M	ECB	3	0	1
100	HICPSV	Harmonized Index of Consumer Prices: Services	2010=100	SCA	M	ECB	3	0	0
101	HICPNG	Harmonized Index of Consumer Prices: Energy	2010=100	MSA	M	EUR	3	0	1
102	HICPFD	Harmonized Index of Consumer Prices: Food	2010=100	MSA	M	EUR	3	0	1
103	DFGDP	Real Gross Domestic Product Deflator	2015=100	SCA	Q	EUR	3	0	0
104	HPRC	Residential Property Prices (BIS)	MLN€	SCA	Q	FRED	3	0	0
105	POIL	Crude Oil Prices: Brent - Europe	€/barrel	MSA	Q	FRED	3	0	0
106	PNGAS	Global price of Natural gas, EU	€/MMbtu	MSA	Q	FRED	3	0	0
		(9) **Confidence Indicators**							
107	ICONFIX	Industrial Confidence Indicator	Index	SA	M	EUR	0	0	1
108	CCONFIX	Consumer Confidence Indicator	Index	SA	M	EUR	0	0	0
109	ESENTIX	Economic Sentiment Indicator	Index	SA	M	EUR	0	0	1
110	KCONFIX	Construction Confidence Indicator	Index	SA	M	EUR	0	0	1
111	RTCONFIX	Retail Confidence Indicator	Index	SA	M	EUR	0	0	0
112	SCONFIX	Services Confidence Indicator	Index	SA	M	EUR	0	0	1
113	BCI	Cyclically-Adjusted Business Confidence Index	2010=100	SA	M	OECD	0	0	1
114	CCI	Cyclically-Adjusted Consumer Confidence Index	2010=100	SA	M	OECD	0	0	0
		(10) **Monetary Aggregates**							
115	CURR	Money Stock: Currency	MLN€	SCA	M	ECB	2	1	1
116	M1	Money Stock: M1	MLN€	SCA	M	ECB	2	1	1
117	M2	Money Stock: M2	MLN€	SCA	M	ECB	2	1	1
		(11) **Others**							
118	SHIX	Share Prices	2010=100	SA	M	OECD	2	0	1
119	CAREG	Passenger's Cars Registrations	1000-U	SCA	M	ECB	2	0	1

B Assumptions

In this section, we state all the formal assumptions underlying the model outlined in Section 3.1, and we provide both econometric and economic justifications for these assumptions. Throughout the text, we let t_{19Q4}, t_{20Q1} and t_{21Q4} denote 2019:Q4, 2020:Q1 and 2021:Q4, respectively.

Identifying assumptions for the space spanned by the factors

(1) The number of factors q is such that $q < n$ and is independent of n.

(2) The q-dimensional vector $\mathbf{f}_t$ is such that $\mathbb{E}[\Delta \mathbf{f}_t] = \mathbf{0}$ and $\mathbb{E}[\Delta \mathbf{f}_t \Delta \mathbf{f}_t'] = \mathbf{I}$.

(3) The $n \times q$ matrix $\boldsymbol{\Lambda} = (\boldsymbol{\lambda}_1 \cdots \boldsymbol{\lambda}_n)'$, with $\boldsymbol{\lambda}_i = (\lambda_{i1} \cdots \lambda_{iq})'$, $1 \leq i \leq n$, is such that $\lim_{n\to\infty} n^{-1}\boldsymbol{\Lambda}'\boldsymbol{\Lambda} = \mathbf{H}$ positive definite.

(4) The scalar g_t is such that $\mathbb{E}[\Delta g_t] = 0$ and $\mathbb{E}[(\Delta g_t)^2] = 1$.

(5) The n-dimensional vector $\boldsymbol{\gamma} = (\gamma_1 \cdots \gamma_n)'$, is such that $\lim_{n\to\infty} n^{-1}\boldsymbol{\gamma}'\boldsymbol{\gamma} > 0$.

(6) The q-dimensional vector $\mathbf{u}_t$ is such that $\mathbb{E}[\mathbf{u}_t] = \mathbf{0}$, $\mathbb{E}[\mathbf{u}_t\mathbf{u}_t'] = \boldsymbol{\Sigma}_u$ is positive definite, and $\mathbb{E}[\mathbf{u}_t\mathbf{u}_{t-k}'] = \mathbf{0}$ for $k \neq 0$. Moreover, s_t is deterministic with $s_t > 0$.

(7) The idiosyncratic innovations e_{it}, $1 \leq i \leq n$, are such that $\mathbb{E}[e_{it}] = 0$, $\mathbb{E}[e_{it}^2] = \sigma_{e_i}^2 > 0$ for all t. Moreover, there exist finite constants M_{ij} independent of t and $0 < \rho < 1$ independent of t, i, and j such that $|\mathbb{E}[e_{it}e_{j,t-k}]| \leq M_{ij}\rho^{|k|}$ for all $k \in \mathbb{Z}$, with $\sum_{i=1}^n M_{ij} \leq M$ and $\sum_{j=1}^n M_{ij} \leq M$ for some finite constant $M > 0$ independent of n.

(8) $\mathbb{E}[e_{it}\mathbf{u}_s] = \mathbf{0}$, for all i, t, s.

Assumptions (1)-(3) require the q factors $\mathbf{f}_t$ to be pervasive so that they have a non-negligible effect on the variables of interest (Bai and Ng, 2004; Barigozzi et al., 2021). Following Maroz et al. (2021), these assumptions are extended in parts (4) and (5) to the Covid factor, g_t, where pervasiveness stems from the common nature of the Covid shock affecting most of the series included in the dataset. Including both 2020 and 2021 in the Covid period is consistent with the evolution of the pandemic in Europe. Ng (2021) has proposed an alternative approach that uses external information from health statistics.

Assumption (6) assumes white noise innovations whose volatility changes over time after the Covid shock—time-varying volatility is not a prominent feature in the pre-2020 sample (Jarociński and Lenza, 2018). Accounting for the change in volatility due to Covid has proven to be fundamental both for estimation and forecasting, and here we adopt an approach similar to Lenza and Primiceri (2022) by introducing a scaling term s_t modeled independently for each period starting from 2020:Q1. Lenza and Primiceri (2022) analyze monthly US data and impose an exponential decay for s_t starting in June 2020. In contrast, we estimate one parameter for each period starting in 2020:Q1 because many series exhibit large variation even after the first half of 2020, which is not surprising given that (i) mobility restriction measures in the EA were much more restrictive than in the US, lasted for longer, and were also implemented in 2021, and (2) the Russia-Ukraine war had a much larger impact on Europe by pushing natural gas prices (and gasoline prices to a lesser extent) to the roof, and in creating a lot of macro-financial uncertainty. Moreover, as Morley et al. (2023) pointed out, quarterly data do not allow for a sharp identification of the decay parameter.

Assumption (7) allows the idiosyncratic innovations to be mildly cross-sectionally correlated and serially correlated with summable autocovariances, thus compatible with stationary ARMA dynamics (Bai and Ng, 2004; Barigozzi et al., 2021). Last, Assumption (8) requires the idiosyncratic and factor innovations to be uncorrelated at all leads and lags, a requirement consistent with the idea of global macroeconomic shocks being unrelated to local dynamics.

Assumptions on the dynamic specifications on the non-stationary idiosyncratic components and the secular components

(9) Let $\mathcal{I}_1$ be the set of indexes such that $\xi_{it} \sim I(1)$ if $i \in \mathcal{I}_1$, then $n_I = \#\{i : i \in \mathcal{I}_1\}$ is such that $0 < n_I < n$.

(10) Let $\mathcal{I}_b$ be the set of indexes such that $b_{it} \neq 0$ if $i \in \mathcal{I}_b$, then $n_B = \#\{i : \ i \in \mathcal{I}_b\}$ is such that $0 < n_B < n$. Moreover, $\mathrm{D}_{i0} = a_i \neq 0$, for all i.

(11) For $i =$ GDP, HHLB, HHLB.LLN, $\mathbb{E}[\eta_{it}] = 0$ and $\sigma^2_{\eta_i} = (400\widehat{\sigma}^2_{\Delta y_i})^{-1}$, while for $i =$ UNETOT HICPOV, HICPNEF, $\mathbb{E}[\epsilon_{it}] = 0$ and $\sigma^2_{\epsilon_i} = (400\widehat{\sigma}^2_{y_i})^{-1}$ where $\widehat{\sigma}^2_{\Delta y_i}$ and $\sigma^2_{y_i}$ are the sample variances of Δy_{it} and y_{it} respectively, computed for $1 \leq t \leq t_{\text{19Q4}}$ and $t_{\text{21Q4}} + 1 \leq t \leq T$.

Assumption (9) allows the idiosyncratic component to be $I(1)$ for some, but not all, of the series. This assumption is crucial when estimating the model on a large dataset. Imposing the assumption of all idiosyncratic components being $I(0)$ would be overly restrictive, as it implies cointegration for any q-dimensional vector of series (Barigozzi et al., 2021). While cointegration may hold for certain series, it is highly unlikely to hold for many others. To accommodate potential cointegration, we allow only a limited number n_I of variables to possess a non-stationary idiosyncratic component. Our dataset, where only $n_I = 57$ out of $n = 119$ series exhibit a non-stationary idiosyncratic component, supports this assumption.

Assumption (10) allows for a non-stationary secular component for some, but not all, of the variables in the dataset. This modeling choice is coherent with the properties of a standard macroeconomic dataset. Specifically, variables related to the real sector of the economy, such as consumption or investments, commonly display a distinct (upward) trend. Conversely, this may not hold for other variables, such as inflation rates or interest rates, for example. This intuition finds support in the empirical data, where only $n_B = 58$ out of $n = 119$ series exhibit a linear trend.

Specifically, letting $a_i = \mathrm{D}_{i0}$, Assumptions (10) and (11) imply that GDP, households' financial liabilities and long-term loans, have a secular component given by the local linear trend model

$$\mathrm{D}_{it} = a_i + b_{it}t, \qquad b_{it} = b_{i,t-1} + \eta_{it}, \qquad i = \text{GDP, LBHH, HHLB.LLN}. \tag{B1}$$

We introduce a local-linear trend for GDP to capture the gradual drift in the secular decline in long-run output growth documented both for the US and the EA (Cette et al., 2016; Antolin-Diaz et al., 2017; Gordon, 2018). The literature has identified several factors contributing to this slowdown, with particular emphasis on declining productivity growth. This decline has been more pronounced in the EA due to heterogeneity between core and peripheral countries, as peripheral countries are experiencing a larger misallocation of economic resources (Cette et al., 2016). Therefore, it is crucial to accurately account for these features to assess GDP's long-run dynamics. This assessment is essential for estimating potential output, as it avoids spuriously inflating the output gap with unexplained predictable variation (Ng, 2018).

We introduce a local-linear trend for household financial liabilities and long-term loans, which constitute about 85% of total household liabilities, to capture the slowdown in their average growth rates that occurred since the GFC.

Assumption (11) implies that the unemployment rate, overall and core inflation have a secular component given by the local level model

$$\mathrm{D}_{it} = a_{it}, \qquad a_{it} = a_{i,t-1} + \epsilon_{it}, \qquad i = \text{UNETOT, HICPOV, HICPNEF}. \tag{B2}$$

This specification captures relevant labor and demographic factors that may affect the unemployment rate secular trend, such as, for example, the aging of the population and the misallocation of resources in the labor market due to "soft budget constraints" or stringent labor market policies can lead to a mismatch between employers' needs and the skill-set of the unemployed (Cette et al., 2016). Similarly, this specification also allows us to account for the slowdown in inflation occurred after the GFC.

All the other variables in the dataset have either a deterministic linear trend or a constant mean, i.e, $\mathrm{D}_{it} = a_i + b_i t$ if $i \in \mathcal{I}_b$, or $\mathrm{D}_{it} = a_i$ otherwise. Although it is technically possible to model a time-varying component for all the variables in the dataset, such an approach would introduce complexities in the estimation framework, with the number of latent states increasing linearly with the number of series.

In Assumption (11) we fix the variances of the stochastic secular components following Del Negro et al. (2019) in order to effectively capture the gradual and persistent nature of the secular trends. This specification implies that the standard deviation of the secular trend, over a century, is approximately 1%, a choice consistent with the notion of a slow-moving secular component.

Assumptions on the dynamics of the factors, trend, and cycles

(12) The polynomial $\det(\mathbf{I} - \sum_{j=0}^{p-1} \mathbf{A}_j z^{j+1}) = 0$ has 1 root in $z = 1$ and the remaining $q-1$ roots in $|z| > 1$.

(13) The q-dimensional vector $\boldsymbol{\psi}$ is such that $\boldsymbol{\beta}'\boldsymbol{\psi} = \mathbf{0}$, where $\boldsymbol{\beta}$ is the $q \times (q-1)$ matrix having as columns the cointegrating vectors of $\mathbf{f}_t$, i.e., such that $\boldsymbol{\beta}'\mathbf{f}_t$ is weakly stationary.

(14) The q-dimensional vector $\boldsymbol{\omega}_t$ is weakly stationary and such that $\mathbb{E}[\boldsymbol{\omega}_t] = \mathbf{0}$ and $\mathbb{E}[\boldsymbol{\omega}_t\boldsymbol{\omega}_t'] = \boldsymbol{\Sigma}_\omega$ is positive definite.

(15) The scalar ν_t is such that $\mathbb{E}[\nu_t] = 0$ and $\mathbb{E}[\nu_t^2] = \sigma_\nu^2 > 0$.

(16) $\mathbb{E}[\nu_t \boldsymbol{\omega}_t] = \mathbf{0}$ for all t.

Assumption (12) imposes that 1 common trend drives the non-stationarity in the common factors, hence, that the factors are cointegrated with $q-1$ cointegrating relations—our data provide strong support for the presence of just one common trend. This is a standard assumption in the literature, which often assumes that common productivity trend is the sole driver of long-run economic growth (see, e.g., Del Negro et al., 2007).

Assumptions (13) and (14) imply that $\boldsymbol{\omega}_t$, defined in (6), belongs to the cointegration space of the common factors. This view is consistent with theoretical models assuming that the output gap represents deviations from long-run equilibria determined by a common productivity trend (Del Negro et al., 2007; Tóth, 2021).

Assumption (15) assumes that ν_t is a stochastic process—hence τ_t is a common stochastic trend—but it does not constraint ν_t to be a white noise—hence τ_t to be a random walk. Indeed, our estimates suggest that ν_t is autocorrelated, in line with the theoretical arguments by Lippi and Reichlin (1994).

Finally, Assumption (16) implies contemporaneous orthogonality between potential output and the output gap, which is also assumed in the non-parametric approaches used by Barigozzi and Luciani (2023).

C Number of Common Factors

In this section, we present the results of various information criteria employed to select the number of common factors. Specifically, we look at the log-information criteria IC of Bai and Ng (2002) (BN), Hallin and Liška (2007) (HL), and Alessi et al. (2010) (ABC), as well at the eigenvalue-ratio criterion by Ahn and Horenstein (2013) (AH), and the test proposed by Onatski (2009) (ON). In order to avoid spurious effects from the Covid period, we employ standardized and de-meaned first-differenced data up to 2019:Q4.

TABLE C1: *Number of common factors*

	ABC	AH	BN	HL	ON
q	4	1	4	4	1

NOTES: ABC = Alessi et al. (2010) criteria; AH = Ahn and Horenstein (2013) test; BN = Bai and Ng (2002) criteria, HL = Hallin and Liška (2007) criteria, ON = Onatski (2009).

As shown in Table C1, the ABC, BN, and HL criteria suggest $q = 4$, while the AH criteria and the ON test suggest $q = 1$. We picked $q = 4$ because (i) the criteria of Ahn and Horenstein (2013) may underperform if there is a significant difference in the explanatory power of the different factors, which is the case in our dataset where the first factor has a much larger explanatory power than the others; and (ii) the Onatski (2009) may not perform well when T is small compared to n.

D Estimation in detail

In this section, we provide details on the estimation procedure described in Section 3.2. Then, the extended state-space form of the model is given by:

$$y_{it} = \mathrm{D}_{it} + \boldsymbol{\lambda}_i'\mathbf{f}_t + \gamma_i g_t \mathbb{I}_{t_{20\mathrm{Q1}} \le t \le t_{21\mathrm{Q4}}} + \zeta_{it} + z_{i,t}, \qquad z_{it} \overset{i.i.d.}{\sim} (0, R_i), \quad 1 \le i \le n,\ 1 \le t \le T, \tag{D1a}$$

$$\mathbf{f}_t = \sum_{j=0}^{p-1} \mathbf{A}_j \mathbf{f}_{t-j} + \{s_t \mathbb{I}_{t \ge t_{20\mathrm{Q1}}} + (1 - \mathbb{I}_{t \ge t_{20\mathrm{Q1}}})\}\mathbf{u}_t, \qquad \mathbf{u}_t \overset{i.i.d.}{\sim} (\mathbf{0}, \boldsymbol{\Sigma}_u), \tag{D1b}$$

$$\mathrm{D}_{it} = \begin{cases} a_i & \text{if } i \in \mathcal{I}_a, \\ a_i + b_i t & \text{if } i \in \mathcal{I}_b, \\ \mathrm{D}_{it-1} + b_{it-1},\ b_{it} = b_{it-1} + \eta_{it} & \text{if } i = \textsc{gdp}, \textsc{hhlb}, \textsc{hhlb.lln}, \qquad \eta_{it} \overset{i.i.d.}{\sim} (0, \sigma^2_{\eta_i}), \\ \mathrm{D}_{it-1} + \epsilon_{i,t} & \text{if } i = \textsc{unetot}, \textsc{hicpov}, \textsc{hicpnef}, \quad \epsilon_{i,t} \overset{i.i.d.}{\sim} (0, \sigma^2_{\epsilon_i}), \end{cases} \tag{D1c}$$

$$\zeta_{it} = \begin{cases} \zeta_{it} + e_{it} & \text{if } i \in \mathcal{I}_1, \quad e_{it} \overset{i.i.d.}{\sim} (0, \sigma^2_{e_i}) \\ 0 & \text{if } i \notin \mathcal{I}_1, \end{cases} \tag{D1d}$$

$$R_i = \begin{cases} \sigma^2_z & \text{if } i \in \mathcal{I}_1, \\ \sigma^2_{e_i} & \text{if } i \notin \mathcal{I}_1, \end{cases} \tag{D1e}$$

where $\mathcal{I}_a$ denotes the set of series with only a constant intercept, while $\mathcal{I}_b$ denotes the set of series with a deterministic linear trend. Similarly, $\mathcal{I}_1$ denotes the set of series with a $I(1)$ idiosyncratic component.

Initialization

In order to apply the Kalman Filter and Smoother, we need initial estimates of all the quantities described in Equations (D1a)-(D1e). We denote with the superscript "19" all quantities computed with data up to 2019. Let $\breve{y}^{19}_{it} = (y^{19}_t - \breve{a}^{19}_i - \breve{b}^{19}_i \cdot t)/\sigma^2_{\Delta y^{19}_i}$, where $\sigma^2_{\Delta y^{19}_i}$ is the sample variance of Δy^{19}_{it}, and $\breve{a}^{19}_i$ and $\breve{b}^{19}_i$ are estimated by regressing a constant and a time trend on y^{19}_{it}, whenever $i \in \mathcal{I}_b$ or $i =$ GDP, HHLB, HHLB.LLN. If $i \in \mathcal{I}_a$ or $i =$ UNETOT, HICPOV, HICPNEF we let $\breve{y}_{it} = y^{19}_{it} - \breve{a}^{19}_i$, where $\breve{a}^{19}_i$ is the sample average of y_{it}. The standardized slopes are denoted as $\widehat{b}^{19}_i = \breve{b}^{19}_i/\sigma^2_{\Delta y^{19}_i}$. We initialize the loadings using the estimator of Barigozzi et al. (2021): the $n \times q$ matrix of estimated loadings $\widehat{\mathbf{\Lambda}}^{19} = (\widehat{\boldsymbol{\lambda}}^{19}_1, \dots, \widehat{\boldsymbol{\lambda}}^{19}_q)'$ is obtained by principal components on the standardized first differences of the data, i.e. $(\Delta y^{19}_{it} - \overline{\Delta y^{19}_i})/\sigma^2_{\Delta y^{19}_i}$, where $\overline{\Delta y^{19}_i}$ is the sample mean of Δy^{19}_{it}. Given the loadings, we also obtain a first estimate of the q common factors, $\widehat{\mathbf{f}}^{19}_t = n^{-1}\widehat{\mathbf{\Lambda}}^{19\prime}\breve{\mathbf{y}}^{19}_t$ and of the idiosyncratic components, $\widehat{\xi}^{19}_{it} = \breve{y}^{19}_{it} - \widehat{\boldsymbol{\lambda}}^{19\prime}_i\widehat{\mathbf{f}}^{19}_t$. Furthermore, we obtain $\widehat{\mathbf{A}}^{19}_j$, $j = 1, \dots, p$, by fitting a VAR(p) on $\widehat{\mathbf{f}}^{19}_t$. Given the residuals $\widehat{\mathbf{u}}^{19}_t = \widehat{\mathbf{f}}^{19}_t - \widehat{\mathbf{J}}^{19}\widehat{\mathbf{f}}^{19}_{t-1}$, where $\widehat{\mathbf{J}}^{19}$ is the companion form representation of the autoregressive matrices $\widehat{\mathbf{A}}^{19}_1, \dots, \widehat{\mathbf{A}}^{19}_p$, an estimate of the latent variance is given by $\widehat{\mathbf{\Sigma}}^{19}_u = \widehat{\mathrm{Cov}}(\widehat{\mathbf{u}}^{19}_t)$, where $\widehat{\mathrm{Cov}}$ is the sample covariance matrix.

We are left with the initialization of the variances of the time-varying parameters of the model and of the idiosyncratic components. We calibrate the variances of the time-varying parameters following Del Negro et al. (2019). In this way, we assume these quantities to slowly drift over time, with a standard deviation of approximately 1% over a century. This results in $\widehat{\sigma}^2_{\eta_i} \approx 4 \times 10^{-3}$ for $i =$ GDP, $\widehat{\sigma}^2_{\eta_i} \approx 10^{-3}$ for $i =$ HHLB, HHLB.LLN and $\widehat{\sigma}^2_{\epsilon} \approx 10^{-2}$ for $i =$ UNETOT, HICPOV, HICPNEF. For the idiosyncratic variances, if $i \notin \mathcal{I}_1$ we set $\widehat{R}^{19}_i = \widehat{\mathrm{Var}}(\widehat{\xi}^{19}_{it})$. On the other hand, if $i \in \mathcal{I}_1$, we set $\widehat{R}^{19}_i = 10^{-2}$, a value sufficiently high to guarantee more accurate estimates (Opschoor and van Dijk, 2023).

Table D1 provides an overview of the initial values of the states for the Kalman Filter. We denote as $\mathbf{f}^{19}_{t|t}, \mathbf{D}^{19}_{t|t}, \mathbf{b}^{19}_{t|t}$ and $\boldsymbol{\zeta}^{19}_{t|t}$ the Kalman Filter estimates (up to 2019:Q4) of the states at time t, estimated using all the information up to time t. Similarly, we denote as $\mathbf{f}^{19}_{t|T}, \mathbf{D}^{19}_{t|T}, \mathbf{b}^{19}_{t|t}$ and $\boldsymbol{\zeta}^{19}_{t|T}$ the Kalman Smoother estimates (up to 2019:Q4) of the states at time t, estimated using all the

Table D1: *Initialization of states for the Kalman Filter*

$\mathbf{f}^{19}_{0\|0} = \widehat{\mathbf{f}}^{19}_0$	
$\text{vec}\left(\mathbf{P}^{19}_{0\|0}\right) = \left(\mathbf{I}_{pq^2} - \widehat{\mathbf{A}}^{19} \otimes \widehat{\mathbf{A}}^{19}\right)^{-1} \text{vec}\left(\widehat{\boldsymbol{\Sigma}}^{19}_u\right)$	
$\text{D}^{19}_{i,0\|0} = 0$	if $i \in \mathcal{I}_a$
$\text{D}^{19}_{i,0\|0} = \widehat{b}^{19}_i$	if $i \in \mathcal{I}_b$, $i =$ GDP, HHLB, HHLB.LLN
$\text{D}^{19}_{i,0\|0} = 0$	if $i =$ UNETOT, HICPOV, HICPNEF
$b^{19}_{i,0\|0} = \widehat{b}^{19}_i$	if $i \in \mathcal{I}_b$, $i =$ GDP, HHLB, HHLB.LLN
$P^{D(19)}_{i,0\|0} = 0$	if $i \in \{\mathcal{I}_a, \mathcal{I}_b\}$
$P^{D(19)}_{i,0\|0} = \frac{1}{(1-0.99)^2}\widehat{\sigma}^2_{\eta_i}$	if $i =$ GDP, HHLB, HHLB.LLN
$P^{D(19)}_{i,0\|0} = \widehat{\sigma}^2_{\epsilon_i}$	if $i =$ UNETOT, HICPOV, HICPNEF
$P^{b(19)}_{i,0\|0} = \frac{1}{(1-0.99)^2}\widehat{\sigma}^2_{\eta_i}$	if $i =$ GDP, HHLB, HHLB.LLN
$\zeta^{19}_{i,0\|0} = \widehat{\xi}^{\zeta(19)}_{i1}$	if $i \in \mathcal{I}_1$
$P^{\zeta(19)}_{i,0\|0} = \frac{1}{(1-0.99)^2}\widehat{\text{Var}}\left(\Delta\widehat{\xi}_{it}\right)$	if $i \in \mathcal{I}_1$

information up to time T.

Estimation is carried on using a standardized version of the data in levels, that is:

$$\tilde{y}_{it} = \begin{cases} \frac{y_{it}-\breve{a}_i}{\sigma^2_{\Delta y_i}} & \text{if} \quad i \in \{\mathcal{I}_b, \text{GDP, HHLB, HHLB.LLN}\} \\ \frac{y_{it}-\bar{y}_i}{\sigma^2_{\Delta y_i}} & \text{if} \quad i \in \{\mathcal{I}_a, \text{UNETOT,HICPOV,HICPNEF}\} \end{cases}$$

with $\tilde{y}^{19}_{it}$ being the standardized data up to 2019. When computing $\sigma^2_{\Delta y_i}$, $\breve{a}_i$ and $\breve{b}_i$ $\widehat{b}_i$) for the entire sample, we treat Covid outliers as missing values.

Step 1: Estimate the model up to 2019:Q4 (pre-Covid step)

Given the initial values of the parameters and the states, we run the Kalman Filter and Smoother using standardized data up to 2019, $\tilde{\mathbf{y}}^{19}_t = (\tilde{y}_{1,t}, \dots, \tilde{y}_{n,t})'$, to obtain a new estimate of the states, namely the factors $\mathbf{f}^{19}_{t|T}$, the time-varying secular components $\mathbf{D}^{19}_{t|T}$ and slopes $\mathbf{b}^{19}_{t|t}$, and the non-stationary idiosyncratic components $\boldsymbol{\zeta}^{19}_{t|T}$, along with the corresponding conditional covariances.

Given the smoothed states, we estimate all the parameters as follows:

- Factor loadings:

$$\widehat{\boldsymbol{\lambda}}^{19'}_i = \left(\sum_{t=1}^{T}\left(\tilde{y}^{19}_{it} - \text{D}^{19}_{i,t|T} - \zeta^{19}_{i,t|T}\right)\mathbf{f}^{19'}_{t|T}\right)\left(\sum_{t=1}^{T}\mathbf{f}^{19}_{t|T}\mathbf{f}^{19'}_{t|T} + \mathbf{P}^{19}_{t|T}\right)^{-1}$$

- PARAMETERS OF THE LAW OF MOTION OF THE COMMON FACTORS:

$$\widehat{\mathbf{A}}^{19} = \left(\sum_{t=2}^{T} \mathbf{f}_{t|T}^{19}\mathbf{f}_{t-1|T}^{19\prime} + \mathbf{P}_{t,t-1\,|\,T}^{19}\right)\left(\sum_{t=2}^{T} \mathbf{f}_{t-1|T}^{19}\mathbf{f}_{t-1|T}^{19\prime} + \mathbf{P}_{t-1\,|\,T}^{19}\right)$$

$$\widehat{\mathbf{\Sigma}}_{u}^{19} = \frac{1}{T}\left(\sum_{t=2}^{T}\left(\mathbf{f}_{t|T}^{19}\mathbf{f}_{t|T}^{19\prime} + \mathbf{P}_{t\,|\,T}^{19}\right) - \widehat{\mathbf{A}}^{19}\sum_{t=2}^{T}\left(\mathbf{f}_{t|T}^{19}\mathbf{f}_{t-1|T}^{19\prime} + \mathbf{P}_{t,t-1\,|\,T}^{19}\right)\right)$$

- SLOPES OF SECULAR TREND:

$$\widehat{b}_{i}^{19} = \left(\sum_{t=1}^{T}\left(\tilde{y}_{it}^{19} - \widehat{\boldsymbol{\lambda}}_{i}^{19\prime}\mathbf{f}_{t|T}^{19} - \zeta_{i,t|T}^{19}\right)t\right)\left(\sum_{t=1}^{T} t^{2}\right)^{-1}$$

- VARIANCE OF $I(1)$ IDIOSYNCRATIC COMPONENTS:

$$\begin{aligned}\widehat{\sigma}_{e_i}^{2,19} &= \frac{1}{T}\sum_{t=2}^{T}\left(\zeta_{i,t|T}^{19}\zeta_{i,t|T}^{19\prime} + P_{i,t\,|\,T}^{\zeta(19)}\right) + \frac{1}{T}\sum_{t=2}^{T}\left(\zeta_{it-1|T}^{19}\zeta_{it-1|T}^{19\prime} + P_{i,t-1\,|\,T}^{\zeta(19)}\right) - \\ &- \frac{2}{T}\sum_{t=2}^{T}\left(\zeta_{it|T}^{19}\zeta_{it-1|T}^{19\prime} + P_{i,t,t-1\,|\,T}^{\zeta(19)}\right)\end{aligned}$$

- COVARIANCE PREDICTION ERROR:

$$\begin{aligned}\widehat{R}_{i}^{19} &= \frac{1}{T}\sum_{t=1}^{T}\Big\{\left(\tilde{y}_{i,t}^{19} - \widehat{\boldsymbol{\lambda}}_{i}^{19\prime}\mathbf{f}_{t|T}^{19} - \mathbb{I}_{i\in\mathcal{I}_b}\mathrm{D}_{i,t|T}^{19} - \mathbb{I}_{i\in\mathcal{I}_1}\zeta_{i,t|T}^{19}\right)^{2} + \widehat{\boldsymbol{\lambda}}_{i}^{19\prime}\mathbf{P}_{t,T\,|\,T}^{19}\widehat{\boldsymbol{\lambda}}_{i}^{19} + \\ &+ \mathbb{I}_{i\in\mathcal{I}_b}P_{i,t\,|\,T}^{\mathrm{D}(19)} + \mathbb{I}_{i\in\mathcal{I}_1}P_{i,t\,|\,T}^{\zeta(19)}\Big\}\end{aligned}$$

Given the estimated factors, $\mathbf{f}_{t|T}^{19}$ we obtain an estimate of the trend and cyclical component for the pre-Covid period by means of the EM algorithm. To run the algorithm, we need an initial estimate of the parameters $\boldsymbol{\psi}$, $\mathbf{\Sigma}_{\omega}$, and σ_{ν}^{2}.

(a) We compute an initial estimate of $\boldsymbol{\psi}^{19}$, denoted as $\widehat{\boldsymbol{\psi}}^{(0),19}$ by PCA on the long-run variance-covariance matrix of the factors (Zhang et al., 2019).

(b) We compute an initial estimate of the common trend, denoted as $\widehat{\tau}_{t}^{(0),19}$, by projecting $\mathbf{f}_{t\,|\,T}^{19}$ onto $\widehat{\boldsymbol{\psi}}^{(0),19}$. This also yields an initial estimate of the cyclical component $\widehat{\boldsymbol{\omega}}_{t}^{(0)} = \mathbf{f}_{t\,|\,T}^{19} - \widehat{\boldsymbol{\psi}}^{(0),19}\widehat{\tau}_{t}^{(0),19}$.

(c) We initialize $\sigma_{\nu}^{2,19}$ following Del Negro et al. (2017, 2019), so that $\widehat{\sigma}_{\nu}^{2(0),19} = (400\widehat{\sigma}_{\Delta\tau^{19}}^{2})^{-1}$, where $\widehat{\sigma}_{\Delta\tau^{19}}^{2}$ is the sample variance of $\Delta\widehat{\tau}_{t}^{(0),19}$. We chose a very small value for the variance of $\widehat{\sigma}_{\nu}^{2(0),19}$ to incorporate our prior assumption of a slow-moving trend.

(d) Lastly, since by construction $\widehat{\boldsymbol{\omega}}_{t}^{(0),19}$ has a sample covariance matrix of reduced rank $(q-1)$, in order to run the EM algorithm we initialize this covariance as $\widehat{\mathbf{\Sigma}}_{\omega}^{(0),19} = T^{-1}\sum_{t=1}^{T}\widehat{\boldsymbol{\omega}}_{t}^{(0),19}\widehat{\boldsymbol{\omega}}_{t}^{(0),19\prime} + \kappa\mathbf{I}_{q}$, where we set $\kappa = 10^{-2}$. This choice is consistent with the recommendations by Opschoor and van Dijk (2023) who show that smaller values of κ might be detrimental for the performance of the algorithm.

Once the initial estimates for the algorithm have been computed, in the E-step we run the Kalman Filter and Smoother to obtain a new estimate of the trend, namely $\tau_{t\,|\,T}^{(1),19}$, along with an estimate of its conditional variance and covariance, $P_{t\,|\,T}^{\tau(1),19}$ and $P_{t,t-1\,|\,T}^{\tau(1),19}$, respectively. The smoothed trend is then used to estimate the parameters in the M-step, until convergence of the algorithm is achieved. For a generic iteration k of the algorithm, the parameters are estimated as follow:

- Trend loadings:

$$\widehat{\boldsymbol{\psi}}^{(k),19} = \left(\sum_{t=1}^{T} \mathbf{f}_{t|T}^{19}\tau_{t|T}^{(k),19}\right)\left(\sum_{t=1}^{T}\tau_{t|T}^{2(k),19} + P_{t|T}^{\tau(k),19}\right)^{-1}$$

- Variance of common trend:

$$\widehat{\sigma}_{\nu}^{2(k),19} = \frac{1}{T}\sum_{t=2}^{T}\left(\tau_{t|T}^{2(k),19} + P_{t|T}^{\tau(k),19}\right) + \frac{1}{T}\sum_{t=2}^{T}\left(\tau_{t-1|T}^{2(k),19} + P_{t-1|T}^{\tau(k),19}\right) - \frac{2}{T}\sum_{t=2}^{T}\left(\tau_{t|T}^{(k),19}\tau_{t-1|T}^{(k),19} + P_{t,t-1|T}^{\tau(k),19}\right)$$

- Covariance of transitory component

$$\widehat{\boldsymbol{\Sigma}}_{\omega}^{(k),19} = \frac{1}{T}\sum_{t=1}^{T}\left\{\left(\mathbf{f}_{t|T}^{19} - \widehat{\boldsymbol{\psi}}^{(k),19}\tau_{t|T}^{(k),19}\right)\left(\mathbf{f}_{t|T}^{19} - \widehat{\boldsymbol{\psi}}^{(k),19}\tau_{t|T}^{(k),19}\right)' + \widehat{\boldsymbol{\psi}}^{(k),19}P_{t|T}^{\tau(k),19}\widehat{\boldsymbol{\psi}}^{(k),19\prime}\right\}$$

The algorithm is stopped using the likelihood-based criterion of Doz et al. (2012), with a threshold of 10^{-3}. At convergence, we obtain an estimate of the trend and transitory component up to 2019:Q4, $\tau_{t|T}^{19}$ and $\boldsymbol{\omega}_{t|T}^{19}$, respectively, along with the estimated parameters $\widehat{\boldsymbol{\psi}}^{19}$, $\widehat{\sigma}_{\nu}^{2,19}$ and $\widehat{\boldsymbol{\Sigma}}_{\omega}^{19}$.

Step 2: Estimate the Covid factor and volatility (Covid step)

Given the estimated parameters up to 2019:Q4, we run the Kalman Filter and Smoother using all the (standardized) data, $\tilde{\mathbf{y}}_t = (\tilde{y}_{1,t}, \ldots, \tilde{y}_{n,t})'$, to obtain the estimated states given the pre-Covid parameters. In doing so, we truncate the Kalman smoother in correspondence of 2020:Q1, to avoid spurious backward effects from the presence of Covid outliers. The estimated states are denoted as $\mathbf{f}_{t|T}^{(0)}$, $\mathbf{D}_{t|T}^{(0)}$, $\mathbf{b}_{t|T}^{(0)}$ and $\boldsymbol{\zeta}_{t|T}^{(0)}$.

Given the smoothed states, let:

$$\widehat{\boldsymbol{\xi}}_t = \tilde{\mathbf{y}}_t - \widehat{\boldsymbol{\Lambda}}^{19}\mathbf{f}_{t|T}^{(0)} - \mathbf{D}_{t|T}^{(0)}$$

and denote as $\widehat{\boldsymbol{\Xi}} = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_t)'$ the $T \times n$ matrix of idiosyncratic components. Then we estimate the Covid factor by estimating the first principal component using the $n \times n$ variance-covariance matrix of the estimated idiosyncratic components from 2020:Q1 to 2021:Q4, denoted as $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\Xi}^C}$. This is the procedure proposed by Maroz et al. (2021) that we modify to account for non-stationarity in the idiosyncratic component. This done by partitioning the matrix of idiosyncratic components during the Covid period as $\widehat{\boldsymbol{\Xi}}^C = (\widehat{\boldsymbol{\Xi}}^{C,1}|\widehat{\boldsymbol{\Xi}}^{C,0})$, where $\widehat{\boldsymbol{\Xi}}^{C,1}$ and $\widehat{\boldsymbol{\Xi}}^{C,1}$ are the matrices of estimated idiosyncratic components in the period 2020:Q1 to 2021:Q4 for $i \in \mathcal{I}_1$ and $i \in \mathcal{I}_0$, respectively. Then, we estimate $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\Xi}^C}$ (Hamilton, 2020, Chapter 17; Bai, 2004):

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\Xi}^C} = \begin{bmatrix} \frac{1}{(T^C)^2}\widehat{\boldsymbol{\Xi}}^{C,1\prime}\widehat{\boldsymbol{\Xi}}^{C,1} & \frac{1}{(T^C)^{3/2}}\widehat{\boldsymbol{\Xi}}^{C,1\prime}\widehat{\boldsymbol{\Xi}}^{C,0} \\ \frac{1}{(T^C)^{3/2}}\widehat{\boldsymbol{\Xi}}^{C,0\prime}\widehat{\boldsymbol{\Xi}}^{C,1} & \frac{1}{T^C}\widehat{\boldsymbol{\Xi}}^{C,0\prime}\widehat{\boldsymbol{\Xi}}^{C,0} \end{bmatrix}$$

where $T^c = 8$ denotes the time-periods between 2020:Q1 and 2021:Q4. Given $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\Xi}^C}$, we obtain the Covid factor and the corresponding loadings as:

$$\widehat{\boldsymbol{\gamma}} \;=\; \sqrt{n}\cdot\widehat{\mathbf{V}}_{\bar{\boldsymbol{\Xi}}^C}$$

$$\widehat{\mathbf{g}} \;=\; \frac{1}{\sqrt{n}}\cdot(\widehat{\boldsymbol{\Xi}}^C\widehat{\mathbf{V}}_{\bar{\boldsymbol{\Xi}}^C})$$

where $\widehat{\mathbf{g}}$ is the $T^C\times 1$ vector with entries $\widehat{g}_t$ and $\widehat{\mathbf{V}}_{\boldsymbol{\Xi}^C}$ is the $n\times 1$ eigenvector corresponding to the largest eigenvalue of $\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\Xi}}^C}$. Given $\widehat{\mathbf{g}}$, the associated loadings are $\widehat{\boldsymbol{\gamma}} = (\widehat{\gamma}_1,\dots,\widehat{\gamma}_n)'$.

Next, give the estimate of states and the Covid factor, we account for the presence of changes in the volatility after the Covid shock by modifying the Lenza and Primiceri (2022) procedure to accommodate for quarterly data. Let $s_t^* = s_t\mathbb{I}_{t\geq t_{20Q1}} + (1-\mathbb{I}_{t\geq t_{20Q1}})$, the likelihood writes as:

$$\begin{aligned}\mathcal{L}(\mathbf{f}^{(0)}\,|\,\mathbf{A},\boldsymbol{\Sigma}_u,s_1^*,\dots,s_T^*) \;&\propto\; \prod_{t=2}^{T}\left|(s_t^*)^2\boldsymbol{\Sigma}_u\right|^{-\frac{1}{2}}\cdot\exp\left\{-\frac{1}{2}\sum_{t=2}^{T}\left(\mathbf{f}_{t|T}^{(0)}-\mathbf{A}\mathbf{f}_{t-1|T}^{(0)}\right)'\left(s_t^2\boldsymbol{\Sigma}_u\right)^{-1}\left(\mathbf{f}_{t|T}^{(0)}-\mathbf{A}\mathbf{f}_{t-1|T}^{(0)}\right)\right\}\\ &\propto\; \left(\prod_{t=2}^{T}(s_t^*)^{-n}\right)|\boldsymbol{\Sigma}_u|^{-\frac{T-1}{2}}\cdot\exp\left\{-\frac{1}{2}\sum_{t=2}^{T}\left(\mathbf{f}_{t|T}^{*(0)}-\mathbf{A}\mathbf{f}_{t-1|T}^{*(0)}\right)'(\boldsymbol{\Sigma}_u)^{-1}\left(\mathbf{f}_{t|T}^{*(0)}-\mathbf{A}\mathbf{f}_{t-1|T}^{*(0)}\right)\right\}\end{aligned}$$

where $\mathbf{f}_{t|T}^{*(0)} = \mathbf{f}_{t|T}^{(0)}/s_t^*$, with corresponding variance-covariance matrix $\mathbf{P}_{t|T}^{*(0)}$.

The maximum-likelihood estimators of $\mathbf{A}$ and $\boldsymbol{\Sigma}_u$ given the factors are:

$$\check{\mathbf{A}} \;=\; \left(\sum_{t=2}^{T}\mathbf{f}_{t|T}^{*(0)}\mathbf{f}_{t-1|T}^{*(0)'}+\mathbf{P}_{t,t-1\,|\,T}^{*(0)}\right)^{-1}\left(\sum_{t=1}^{T}\mathbf{f}_{t-1|T}^{*(0)}\mathbf{f}_{t-1|T}^{*(0)'}+\mathbf{P}_{t-1\,|\,T}^{*(0)}\right)$$

$$\check{\boldsymbol{\Sigma}}_u \;=\; \frac{1}{T}\left(\sum_{t=2}^{T}\left(\mathbf{f}_{t|T}^{*(0)}\mathbf{f}_{t|T}^{*(0)'}+\mathbf{P}_{t\,|\,T}^{*(0)}\right)-\check{\mathbf{A}}\sum_{t=2}^{T}\left(\mathbf{f}_{t|T}^{*(0)}\mathbf{f}_{t-1|T}^{*(0)'}+\mathbf{P}_{t,t-1\,|\,T}^{*(0)}\right)\right)$$

Substituting $\check{\mathbf{A}}$ and $\check{\boldsymbol{\Sigma}}_u$ in the likelihood, we obtain the concentrated likelihood:

$$\mathcal{L}(\mathbf{f}^{(0)}\,|\,\check{\mathbf{A}},\check{\boldsymbol{\Sigma}}_u,s_1^*,\dots,s_T^*) \;=\; \prod_{t=2}^{T}(s_t^*)^{-n}\cdot\left|\check{\boldsymbol{\Sigma}}_u\right|^{-\frac{T}{2}}$$

By numerically maximizing the concentrated likelihood we obtain the volatility parameters $\widehat{s}_t$.

Step 3: Full sample estimation

Given the Covid factor estimated in Step 2, we obtain all the other parameters in the model:

- Factor loadings:

$$\widehat{\boldsymbol{\lambda}}_i' \;=\; \left(\sum_{t=1}^{T}\left(\tilde{y}_{it}-\mathrm{D}_{i,t\,|\,T}^{(0)}-\zeta_{i,t\,|\,T}^{(0)}\right)\mathbf{f}_{t\,|\,T}^{(0)'}-\widehat{\gamma}_i\widehat{g}_t\right)\left(\sum_{t=1}^{T}\mathbf{f}_{t\,|\,T}^{(0)}\mathbf{f}_{t\,|\,T}^{(0)'}+\mathbf{P}_{t\,|\,T}^{(0)}\right)^{-1}$$

- Slopes of secular trend:

$$\widehat{b}_i \;=\; \left(\sum_{t=1}^{T}\left(\tilde{y}_{it}-\widehat{\boldsymbol{\lambda}}_i'\mathbf{f}_{t|T}^{(0)}-\zeta_{i,t|T}^{(0)}-\widehat{\gamma}_i\widehat{g}_t\right)t\right)\left(\sum_{t=1}^{T}t^2\right)^{-1}$$

- Variance of $I(1)$ idiosyncratic components:

$$\begin{aligned}\widehat{\sigma}^2_{e_i} &= \frac{1}{T}\sum_{t=2}^{T}\left(\zeta^{(0)}_{i,t|T}\zeta^{(0)\prime}_{i,t|T} + P^{\zeta(0)}_{i,t\,|\,T}\right) + \frac{1}{T}\sum_{t=2}^{T}\left(\zeta^{(0)}_{it-1|T}\zeta^{(0)\prime}_{it-1|T} + P^{\zeta(0)}_{i,t-1\,|\,T}\right) - \\ &\quad - \frac{2}{T}\sum_{t=2}^{T}\left(\zeta^{(0)}_{it|T}\zeta^{(0)\prime}_{it-1|T} + P^{\zeta(0)}_{i,t,t-1\,|\,T}\right)\end{aligned}$$

- Covariance prediction error:

$$\begin{aligned}\widehat{R}_i &= \frac{1}{T}\left\{\sum_{t=1}^{T}\left(\tilde{y}^{(0)}_{i,t} - \widehat{\boldsymbol{\lambda}}^{(0)\prime}_i\mathbf{f}^{(0)}_{t|T} - \mathbb{I}_{i\in\mathcal{I}_b}\mathrm{D}^{(0)}_{i,t|T} - \mathbb{I}_{i\in\mathcal{I}_1}\zeta^{(0)}_{i,t|T} - \widehat{\gamma}_i\widehat{g}_t\right)^2 + \widehat{\boldsymbol{\lambda}}_i'\mathbf{P}^{(0)}_{t,T\,|\,T}\widehat{\boldsymbol{\lambda}}_i + \right.\\ &\quad \left. + \mathbb{I}_{i\in\mathcal{I}_b}P^{\mathrm{D}(0)}_{i,t\,|\,T} + \mathbb{I}_{i\in\mathcal{I}_1}P^{\zeta(0)}_{i,t\,|\,T}\right\}\end{aligned}$$

Given the estimated parameters, using data net of the Covid component, i.e. $\tilde{\mathbf{y}}_t - \widehat{\boldsymbol{\gamma}}\widehat{g}_t$, we obtain a final estimates of all the states, $\mathbf{f}_{t\,|\,T}$, $\mathbf{D}_{t\,|\,T}$, $\mathbf{b}_{t\,|\,T}$, $\boldsymbol{\zeta}_{t\,|\,T}$ and their conditional covariances with the Kalman Filter and Smoother. Note that, for this final run, the Kalman Filter and Smoother do not need to be truncated in 2019:Q4 because we have already controlled for the Covid pandemic.

With the final estimates of the factors, $\mathbf{f}_{t\,|\,T}$ we obtain a final estimate of the trend and transitory component by means of the EM algorithm. We initialize the EM algorithm with the estimated trend loadings up to 2019:Q4, normalized so that $\widehat{\boldsymbol{\psi}}^{19\prime}\widehat{\boldsymbol{\psi}}^{19} = 1$. We denote this initial estimate as $\widehat{\boldsymbol{\psi}}^{(0)}$. An initial estimate of the trend component is obtained by projecting $\mathbf{f}_{t\,|\,T}$ onto $\widehat{\boldsymbol{\psi}}^{(0)}$, i.e. $\tau^{(0)}_t = (\widehat{\boldsymbol{\psi}}^{(0)\prime}\widehat{\boldsymbol{\psi}}^{(0)})^{-1}\widehat{\boldsymbol{\psi}}^{(0)\prime}\mathbf{f}_{t\,|\,T}$. This also yields an initial estimate of the transitory component $\boldsymbol{\omega}^{(0)}_t = \mathbf{f}_{t\,|\,T} - \widehat{\boldsymbol{\psi}}^{(0)}\tau^{(0)}_t$. Finally, σ^2_ν is initialized as $(400\widehat{\sigma}^2_{\Delta\tau})^{-1}$, where $\widehat{\sigma}^2_{\Delta\tau}$ is the sample variance of $\widehat{\tau}^{(0)}_t$.

After having computed the initial estimates for the algorithm, we estimate the trend via the usual EM algorithm.

For a generic iteration k of the algorithm, we estimate the parameters as follow:

- Trend loadings:

$$\widehat{\boldsymbol{\psi}}^{(k)} = \left(\sum_{t=1}^{T}\mathbf{f}_{t\,|\,T}\tau^{(k)}_{t|T}\right)\left(\sum_{t=1}^{T}\tau^{2(k)}_{t|T} + P^{\tau(k)}_{t|T}\right)^{-1}$$

- Variance of common trend:

$$\begin{aligned}\widehat{\sigma}^{2(k)}_{\nu} &= \frac{1}{T}\sum_{t=2}^{T}\left(\tau^{2(k)}_{t|T} + P^{\tau(k)}_{t|T}\right) + \frac{1}{T}\sum_{t=2}^{T}\left(\tau^{2(k)}_{t-1|T} + P^{\tau(k)}_{t-1\,|\,T}\right) - \\ &\quad - \frac{2}{T}\sum_{t=2}^{T}\left(\tau^{(k)}_{t|T}\tau^{(k)}_{t-1|T} + P^{\tau(k)}_{t,t-1\,|\,T}\right)\end{aligned}$$

- Covariance of transitory component

$$\widehat{\boldsymbol{\Sigma}}^{(k)}_{\omega} = \frac{1}{T}\sum_{t=1}^{T}\left\{\left(\mathbf{f}_{t\,|\,T} - \widehat{\boldsymbol{\psi}}^{(k)}\tau^{(k)}_{t|T}\right)\left(\mathbf{f}_{t\,|\,T} - \widehat{\boldsymbol{\psi}}^{(k)}\tau^{(k)}_{t|T}\right)' + \widehat{\boldsymbol{\psi}}^{(k)}P^{\tau(k)}_{t|T}\widehat{\boldsymbol{\psi}}^{(k)\prime}\right\}$$

The algorithm is stopped using the likelihood-based criterion of Doz et al. (2012), with a threshold of 10^{-3}. At convergence, we obtain an estimate of the trend and transitory components, $\tau_{t|T}$ and $\boldsymbol{\omega}_{t|T}$, respectively, along with the estimated parameters $\widehat{\boldsymbol{\psi}}$, $\widehat{\sigma}_{\nu}^2$ and $\widehat{\boldsymbol{\Sigma}}_{\omega}$.

Ultimately, given the estimated trend and transitory components, the estimated output gap and potential output are defined as:

$$
\begin{aligned}
\widehat{\mathrm{PO}}_t &= \mathrm{D}_{\mathrm{GDP},t|T} + \widehat{\boldsymbol{\lambda}}'_{\mathrm{GDP}} \widehat{\boldsymbol{\psi}} \tau_{t|T}, \\
\widehat{\mathrm{OG}}_t &= \widehat{\boldsymbol{\lambda}}'_{\mathrm{GDP}} \boldsymbol{\omega}_{t|T}
\end{aligned}
$$

E Confidence bands

To obtain confidence bands for our quantities of interest, we follow the procedure outlined in Barigozzi and Luciani (2023). In particular, we simulate all the states in the model using the simulation smoother of Durbin and Koopman (2002), and we generate all the stationary residuals of the model using a stationary block bootstrap procedure (Politis and Romano, 1994). In practice, we have an estimate of all the states, namely $\mathbf{f}_{t|T}$, $\mathbf{D}_{t|T}$, $\mathbf{b}_{t|T}$ and $\boldsymbol{\zeta}_{t|T}$, an estimate of the Covid factor $\widehat{g}_t$ and the estimated volatility parameters $\widehat{s}_t$, $t \geq t_{\text{20Q1}}$. Then, the algorithm is structured as follows:

1. Simulate the states by the simulation smoother (Durbin and Koopman, 2002)

   (a) Common factors:

      i. simulate $\tilde{\mathbf{f}}_1^{(b)} \sim N\left(\mathbf{f}_{1|T}, \mathbf{P}_{1|T}\right)$;

      ii. simulate $\tilde{\mathbf{u}}_t^{(b)} \sim N\left(\mathbf{0}_q, \check{\boldsymbol{\Sigma}}_u\right)$;

      iii. for $t = 2, \ldots, T$ generate $\tilde{\mathbf{f}}_t^{(b)} = \sum_{k=1}^p \check{\mathbf{A}}_k \tilde{\mathbf{f}}_{t-k}^{(b)} + \{s_t \mathbb{I}_{t \geq t_{\text{20Q1}}} + (1 - \mathbb{I}_{t \geq t_{\text{20Q1}}})\} \tilde{\mathbf{u}}_t^{(b)}$.

   (b) $I(1)$ idiosyncratic components. For each $i \in \mathcal{I}_1$:

      i. simulate $\tilde{\zeta}_{i1}^{(b)} \sim N\left(\zeta_{i,1|T}, P_{i,1|T}^{\zeta}\right)$;

      ii. simulate $\tilde{e}_{it}^{(b)} \sim N\left(0, \widehat{\sigma}_{e_i}^2\right)$;

      iii. for $t = 2, \ldots, T$ generate $\tilde{\zeta}_{it}^{(b)} = \tilde{\zeta}_{it-1}^{(b)} + \tilde{e}_{it}^{(b)}$.

   (c) Time-varying secular components:

      - $i =$ GDP, HHLB, HHLB.LLN:

        i. simulate $\tilde{b}_{i,1}^{(b)} \sim N\left(b_{i,1|T}, P_{i,1|T}^{b}\right)$, and set $\tilde{\mathrm{D}}_{i,1}^{(b)} = \tilde{b}_{i,1}^{(b)}$;

        ii. simulate $\tilde{\eta}_t^{(b)} \sim N\left(0, \widehat{\sigma}_{\eta_i}^2\right)$;

        iii. for $t = 2, \ldots, T$ generate $\tilde{b}_{i,t}^{(b)} = \tilde{b}_{i,t-1}^{(b)} + \tilde{\eta}_t^{(b)}$;

        iv. for $t = 2, \ldots, T$ generate $\tilde{\mathrm{D}}_{i,t}^{(b)} = \tilde{\mathrm{D}}_{i,t-1}^{(b)} + \tilde{b}_{i,t}^{(b)}$.

      - $i =$ UNETOT, HICPOV, HICPNEF:

        i. simulate $\tilde{\mathrm{D}}_{i,1}^{(b)} \sim N\left(b_{i,1|T}, P_{i,1|T}^{\mathrm{D}}\right)$

        ii. simulate $\tilde{\eta}_t^{(b)} \sim N\left(0, \widehat{\sigma}_{\epsilon_i}^2\right)$;

        iii. for $t = 2, \ldots, T$ generate $\tilde{\mathrm{D}}_{i,t}^{(b)} = \tilde{\mathrm{D}}_{i,t-1} + \tilde{\eta}_t^{(b)}$.

2. Simulate the stationary residuals of the model, $\mathbf{z}_t = (z_{1,t}, \dots, z_{n,t})'$, using a stationary block-bootstrap (Politis and Romano, 1994) with an average block length of four quarters. Denote the resulting simulated residuals as $\tilde{\mathbf{z}}_t^{(b)} = (\tilde{z}_{1,t}^{(b)}, \dots, \tilde{z}_{n,t}^{(b)})'$. This step, along with step 2(b), gives us the simulated idiosyncratic component, i.e. $\tilde{\xi}_{i,t}^{(b)} = \tilde{\zeta}_{i,t}^{(b)} + \tilde{z}_{i,t}^{(b)}$ if $i \in \mathcal{I}_1$ and $\tilde{\xi}_{i,t}^{(b)} = \tilde{z}_{i,t}^{(b)}$ if $i \in \mathcal{I}_0$.

3. Generate the data. For $t = 1, \dots, T$, generate:
   (a) $\tilde{y}_{it}^{(b)} = \tilde{\mathrm{D}}_{it}^{(b)} + \widehat{\boldsymbol{\lambda}}_i' \tilde{\mathbf{f}}_t^{(b)} + \widehat{\gamma}_i \widehat{g}_t + \tilde{\xi}_{it}^{(b)}$, for $i =$ GDP, HHLB, HHLB.LLN, UNETOT, HICPOV, HICPNEF
   (b) $\tilde{y}_{it}^{(b)} = \mathrm{D}_{i,t|T} + \widehat{\boldsymbol{\lambda}}_i' \tilde{\mathbf{f}}_t^{(b)} + \widehat{\gamma}_i \widehat{g}_t + \tilde{\xi}_{it}^{(b)}$, for all other variables.

4. Using $\tilde{\mathbf{y}}_t^{(b)} = (\tilde{y}_{1,t}^{(b)}, \dots, \tilde{y}_{n,t}^{(b)})'$, estimate the model as described in Appendix D to get a new estimate of the loadings, $\widehat{\boldsymbol{\Lambda}}^{(b)} = (\widehat{\boldsymbol{\lambda}}_1^{(b)'}, \dots, \widehat{\boldsymbol{\lambda}}_N^{(b)'})'$, and all the other parameters in the model, as well as a new estimate of the states $\mathbf{f}_{t|T}^{(b)}, \mathbf{D}_{t|T}^{(b)}, \mathbf{b}_{t|T}^{(b)}$ and $\boldsymbol{\zeta}_{t|T}^{(b)}$.

5. Center the estimated states: $\bar{\mathbf{f}}_{t|T}^{(b)} = \mathbf{f}_{t|T} - \tilde{\mathbf{f}}_t^{(b)}$, $\bar{\mathbf{D}}_{t|T}^{(b)} = \mathbf{D}_{t|T} - \tilde{\mathbf{D}}_t^{(b)} + \mathbf{D}_{t|T}^{(b)}$, $\bar{\mathbf{b}}_{t|T}^{(b)} = \mathbf{b}_{t|T} - \tilde{\mathbf{b}}_t^{(b)} + \mathbf{b}_{t|T}^{(b)}$ and $\bar{\boldsymbol{\zeta}}_t^{(b)} = \boldsymbol{\zeta}_{t|T} - \tilde{\boldsymbol{\zeta}}_t^{(b)} + \boldsymbol{\zeta}_{t|T}^{(b)}$.

6. Run the trend cycle decomposition on the estimated factors $\bar{\mathbf{f}}_t^{(b)}$ to get a new estimate of the common trend $\tau_{t|T}^{(b)}$, the transitory component $\boldsymbol{\omega}_{t|T}^{(b)}$, and the parameter $\widehat{\boldsymbol{\psi}}^{(b)}$.

7. Estimate potential output as $\widehat{\mathrm{PO}}_t^{(b)} = \bar{\mathrm{D}}_{\mathrm{GDP},t|T}^{(b)} + \widehat{\boldsymbol{\lambda}}_{\mathrm{GDP}}^{(b)'} \widehat{\boldsymbol{\psi}}^{(b)} \tau_{t|T}^{(b)}$, and the output gap as $\widehat{\mathrm{OG}}_t^{(b)} = \widehat{\boldsymbol{\lambda}}_{\mathrm{GDP}}^{(b)'} \boldsymbol{\omega}_{t|T}^{(b)}$.

Repeating this procedure B times, we obtain a distribution of the output gap: $\{\widehat{\mathrm{OG}}_t^{(b)}, b = 1, \dots, B\}$. Then, we construct the $(1-\alpha)$ confidence interval as $[\widehat{\mathrm{OG}}_t + z_{\alpha/2} \widehat{\sigma}_t^{\mathrm{OG}}, \widehat{\mathrm{OG}}_t + z_{1-\alpha/2} \widehat{\sigma}_t^{\mathrm{OG}}]$, where $\widehat{\sigma}_t^{\mathrm{OG}}$ is the sample standard deviation of $\{\widehat{\mathrm{OG}}_t^{(b)} - \widehat{\mathrm{OG}}_t\}$ and $z_{\alpha/2} = -z_{1-\alpha/2}$ is the $\alpha/2$-th quantile of a standard normal distribution.

F Identification à la Morley et al. (2023)

Morley et al. (2023), henceforth MTW, propose an alternative trend smoothing approach to correct the estimated trend whenever it displays some serial correlation in first differences despite being assumed to be a random walk. We can apply the MTW approach in our setting by estimating an ARMA(1,1) model on the first difference of the estimated common trend $\Delta \widehat{\tau}_t^{(0)}$. The trend estimate corrected as in MTW is given by

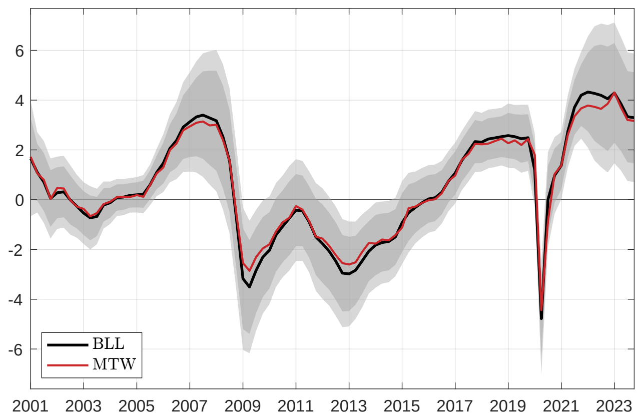
$$\Delta \tilde{\tau}_t = \left(\frac{1 + \widehat{\theta}}{1 - \widehat{\phi}} \right) \widehat{\varepsilon}_t, \tag{F1}$$

where $\widehat{\phi}$ and $\widehat{\theta}$ are the estimated ARMA parameters, and $\widehat{\varepsilon}_t$ are the ARMA residuals. By cumulating $\Delta \tilde{\tau}_t$, we obtain the corrected estimate of the common trend, $\tilde{\tau}_t$.

In practice, smoothing as in (F1) seems to be less efficient than our proposal of using the Kalman smoother due to the presence of some residual Covid volatility in the estimated factors that affects the ARMA estimation. In particular, $\widehat{\mathrm{Var}}(\Delta \widehat{\tau}_t) = \widehat{\sigma}_\nu^2 = 0.092$ (computed by excluding observation in 2020 and 2021 from the calculation) when estimating the model with our method, and $\widehat{\mathrm{Var}}(\Delta\widetilde{\tau}_t) = (1+\widehat{\theta})(1-\widehat{\phi})^{-1}\widehat{\sigma}_{\varepsilon}^2 = 0.17$, when estimating the model as in (F1), using MTW's method, where $\widehat{\sigma}_{\varepsilon}^2$ is the sample variance of $\widehat{\varepsilon}_t$. Although these values are quite similar, they become very different if we compute them using the whole sample; indeed, they increase to 0.095 and 0.52, respectively. This difference is essentially due to anomalous fluctuations in the ARMA residuals $\widehat{\varepsilon}_t$ during 2020-2021.

Figure F1 compares our output gap estimate with the one obtained by applying the correction proposed by Morley et al. (2023). As expected, the two approaches yield very similar estimates, but in 2021 and 2022, when the MTW corrections reduce the estimate of the output gap from +4% to +3%.

FIGURE F1: *Output gap estimate with identification à la Morley et al. (2023)*



NOTES: The black bold line is our estimate of the output gap. The grey shaded areas are the 68% and 84% confidence bands. The light red line is the estimated output gap obtained applying the correction of the preliminary estimated trend proposed by Morley et al. (2023) (MTW).

G Comparison with alternative measures

In this section, we compare our estimates of potential output and the output gap with those obtained using four different univariate filters and one multivariate approach:

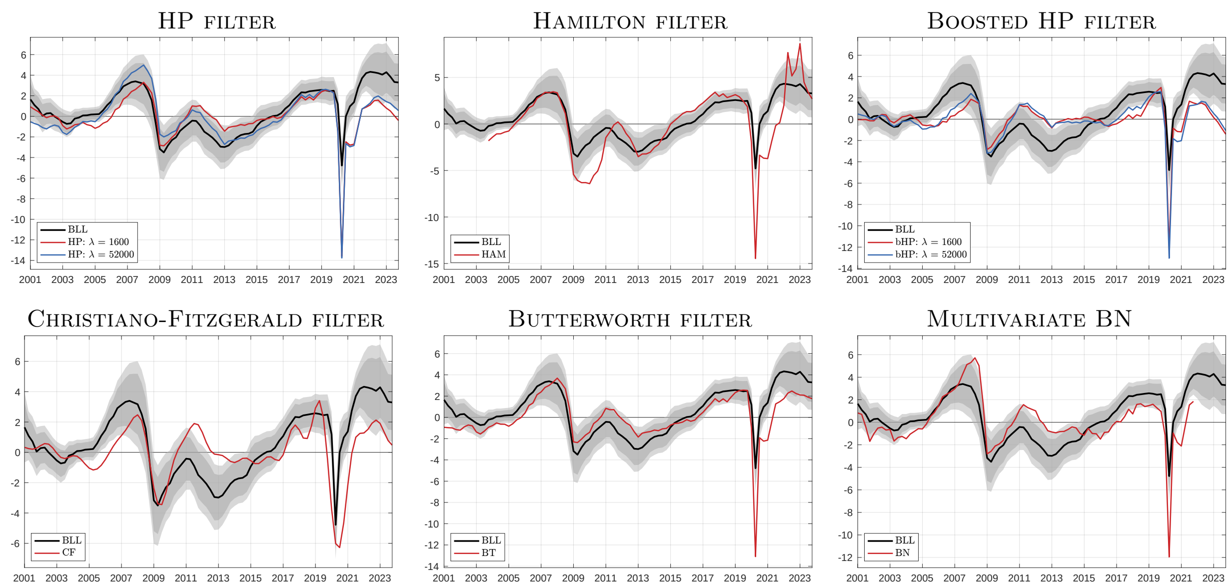
1) The Hodrick and Prescott (1997) filter (HP), with two different values for the smoothing parameter λ: (i) $\lambda = 1600$, commonly used for quarterly data, (ii) $\lambda = 51200$, as proposed by Borio (2014) to capture variability at lower frequencies.
2) The Hamilton (2018) filter (Ham), where the trend is the 8−step ahead forecast of quarterly GDP growth, obtained using the 4 most recent values of quarterly GDP for each time t, and the cycle is the residual obtained from this regression.
3) The boosted HP filter (bHP) of Phillips and Shi (2021), which improves on the standard HP filter by applying the filter recursively on the residuals extracted from previous iterations. The number of iterations m is a tuning parameter that controls the intensity of the updating, and it is chosen to minimize the information criterion proposed by the authors.
4) The Christiano and Fitzgerald (2003) filter (CF), with cutoff frequencies for the transitory component between 8 and 32 quarters.
5) A Butterworth filter (BT) for the transitory component, as proposed by Canova (2022). This

filter can be cast in state-space form, and its squared-gain function defines the frequencies attributed to the cycles. Here, we employ a first-order polynomial $n = 1$, with a cutoff point for the frequency set at $\omega = 0.04$ and scale $G_0 = 1$.

6) The multivariate Beveridge-Nelson (BN) decomposition based on a large Bayesian VAR, as proposed by Morley et al. (2023). The authors estimate the Euro Area output gap from 1999:Q1 to 2021:Q3. Here, due to the lack of availability of their data, we keep their original estimates, truncating the figure in correspondence with our starting point, i.e. 2001:Q1.

Figure G1 presents the results of this exercise. Overall, the output gap obtained with our methodology aligns with those estimated with univariate models in terms of peaks and troughs. However, there are several differences in terms of shape and amplitude.

FIGURE G1: *Output gap estimates with alternative methods*



NOTES: The black line is our benchmark estimate and the grey shaded areas are the 68% and 84% confidence bands, the red and blue lines are alternative estimates.

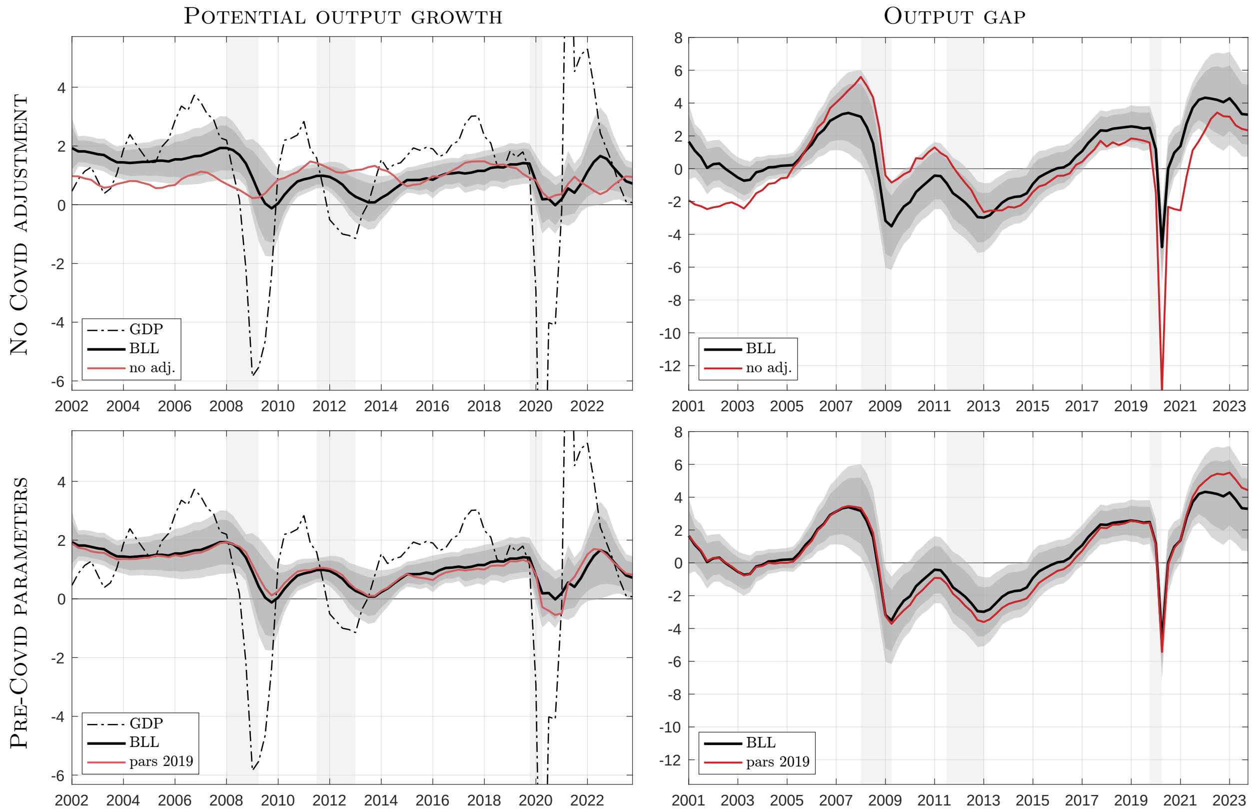
H Accounting for Covid

H.1 Why adjusting for the Covid shock matters

Section 3.2 explained how we accounted for the Covid shock when estimating the model. The upper plots in Figure H1 compare our benchmark estimates with the one we would have obtained if we had estimated the model over the full sample without applying any adjustment for the Covid shock ("no adj."). As can be seen, ignoring the Covid shock affects the estimates of potential output and the output gap throughout the sample, which is undesirable; moreover, ignoring the Covid shock distorts the estimate of common dynamics in 2020 and 2021.

Having shown that accounting for the Covid shock is necessary, the question is whether a strategy different from the one we adopted would have been desirable. For example, what if we had estimated all parameters up to 2019:Q4 and then extracted the states by simply truncating the Kalman Smoother? Despite being effective, this strategy is sub-optimal because estimating the parameters

Figure H1: *Output gap when using pre-Covid parameters or without Covid adjustment*



Notes: The black solid line is our benchmark estimate and the grey shaded areas are the 68% and 84% confidence bands, the black dashed line is GDP YoY growth rate, the red lines are the estimates obtained with two alternative estimation strategies: 1. (left) fixing parameters estimated up to 2019:Q4 (pars 2019); 2. (right) estimating the model with no adjustment for Covid (no adj.).

up to 2019 becomes less and less justifiable as new data come in. Moreover, if the increase in volatility induced by the Covid shock turns out to be very persistent, confidence intervals would be underestimated because they only account for the pre-Covid volatility regime. By accounting for the Covid shock, we avoid both these issues.

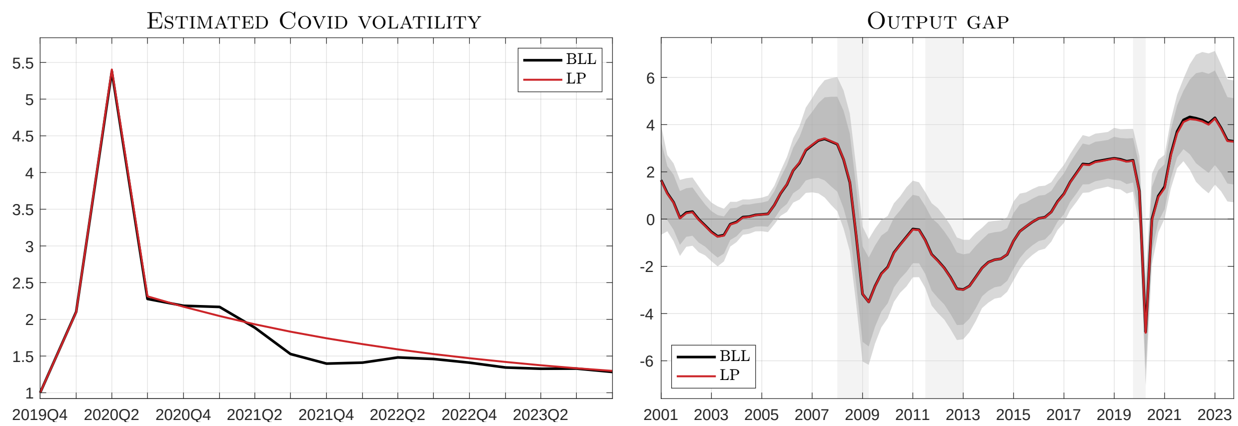
The lower plots in Figure H1 compare our benchmark estimate with the one obtained by estimating the parameters up to the last quarter of 2019 ("pars 2019"). The two estimates are very similar up to the pandemic, after which the estimate using the up-to-2019 parameters points towards much larger fluctuations in potential output growth and a much larger output gap, which, if taken at face value, signals a very tight economy.

H.2 Covid volatiliy

Section 3.2 explained how we accounted for the effect of the Covid shock on the volatility of the common factors. Specifically, we follow Lenza and Primiceri (2022) and introduce a factor s_t that scales the volatility of the common factors from the beginning of the pandemic onward. Lenza and Primiceri (2022) analyzed monthly US data and impose an exponential decay for s_t starting in June 2020. In contrast, we estimate one parameter for each period starting in 2020:Q1. This choice is motivated by both the different impact and policy response of the Covid pandemic in Europe and by the fact that, as Morley et al. (2023) pointed out, quarterly data do not allow for a sharp identification of the decay parameter.

The left plot in Figure H2 shows the estimated scaling factor s_t obtained under both our parametrization (black line) and under the exponential decay parametrization proposed by Lenza and Primiceri (2022) (red line). Our estimate of the volatility closely tracks the evolution of the pandemic, as it spikes in the first two quarters of 2020 when mobility restrictions were most stringent in Europe, and pick-up again in 2021:Q1 when the spread of the Delta variant reached its peak. Moreover, we estimate that the volatility is very persistent—the decay we get is very close to 1 (≈ 0.98)—much more persistent than estimated by Lenza and Primiceri (2022), who estimated their model on monthly US data. This difference is likely due to the evolution of the pandemic in Europe, where mobility restriction measures were much more restrictive than in the US, lasted for longer, and were also implemented in 2021. Moreover, the Russia-Ukraine war had a much larger impact on Europe by pushing natural gas (and gasoline prices to a lesser extent) to the roof and creating a lot of macro-financial uncertainty. This result motivates the need to allow for time variation in the factor volatility until the end of the sample.

Figure H2: *Output gap estimate with alternative Covid volatility*



Notes: In the left plot, the black line is our estimate of Covid volatility (s_t), the red line is the estimate obtained by assuming the exponential decay parametrization of the Covid volatility as in Lenza and Primiceri (2022). In the right plot, the black line is our benchmark estimate and the grey shaded areas are the 68% and 84% confidence bands, the red line is the estimate obtained with the exponential decay parametrization of the Covid volatility.

As shown in the right plot in Figure H2, the two parametrizations lead to virtually identical results.

H.3 Alternative estimator for the Covid factor

The approach we described in Appendix D to estimate the Covid factor has the benefit of retaining all the information in $\widehat{\boldsymbol{\xi}}_t$, but it has the problem of relying only on eight data points. That said, since we are only interested in the first eigenvector of $\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\Xi}}^C}$, and given the extent to which the series co-moved during the Covid period, even a few data points should be informative. However, the estimates could be imprecise, thereby motivating an alternative estimation strategy.

As an alternative approach, we estimate the Covid factor by estimating the first principal component using the $T^C \times T^C$ variance-covariance matrix of the estimated idiosyncratic components from 2020:Q1 to 2021:Q4, denoted as $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Xi}^C}$. In order to estimate $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Xi}^C}$, we consider $\widehat{\xi}_{it}$ if $i \in \mathcal{I}_0$ and

$\Delta\widehat{\xi}_{it}$ if $i \in \mathcal{I}_1$, i.e. we take first-differences of all non-stationary idiosyncratic components.[xi] We obtain the Covid factor and the corresponding loadings as:

$$\tilde{\mathbf{g}} \;=\; \sqrt{T^C}\cdot \tilde{\mathbf{V}}_{\bar{\boldsymbol{\Xi}}^C}$$

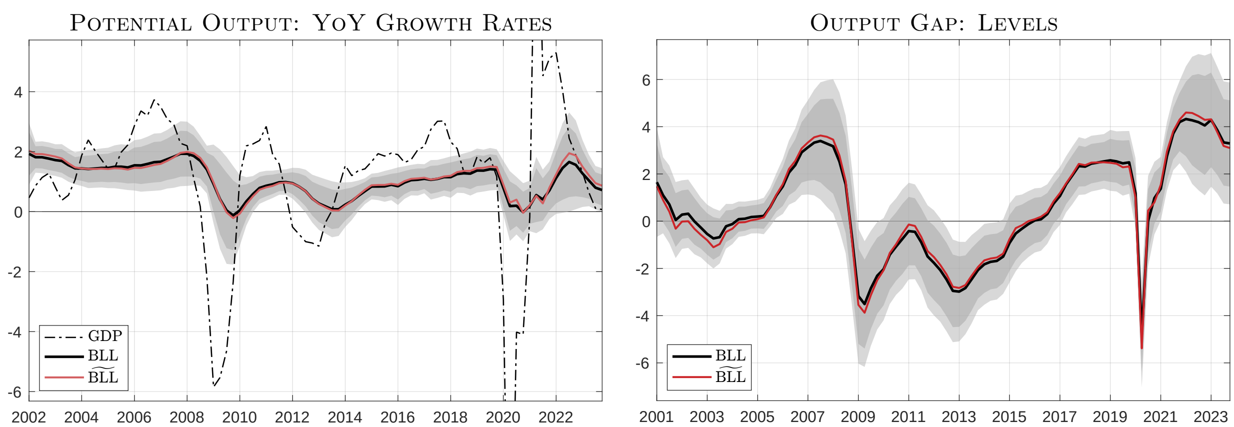
$$\tilde{\boldsymbol{\gamma}} \;=\; \frac{1}{\sqrt{T^C}}\cdot (\widehat{\boldsymbol{\Xi}}^C \tilde{\mathbf{V}}_{\bar{\boldsymbol{\Xi}}^C})$$

where $\tilde{\mathbf{g}}$ is the $T^C \times 1$ vector with entries $\tilde{g}_t$ and $\tilde{\mathbf{V}}_{\bar{\boldsymbol{\Xi}}^C}$ is the $T^C \times 1$ eigenvector corresponding to the largest eigenvalue of $\tilde{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\Xi}}^C}$. Given $\tilde{\mathbf{g}}$, the associated loadings are $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_n)'$.

This second strategy allows us to estimate $\tilde{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\Xi}}^C}$ with N data points, thereby yielding a more precise estimate. However, this comes at the cost of missing important information due to differencing of the non-stationary idiosyncratic component. Which one of the two approaches is better?

As a robustness exercise, in this Appendix we look at what would have been the output gap estimate, had we adopted the second strategy to estimate the Covid factor that we just laid out. As shown in Figure H3, the results obtained with the alternative Covid factor are almost identical to those obtained in the benchmark specification. This is not surprising, since the co-movements observed in most of the series during the Covid period are so large to be easily identified even with a limited range of observations.

Figure H3: *Output gap estimate with the alternative Covid factor*



Notes: The black solid line is our benchmark estimate and the grey shaded areas are the 68% and 84% confidence bands. The red solid lines are the estimates obtained with the alternative Covid factor ($\widetilde{\text{BLL}}$). The level of the output gap is the percentage deviation from potential.

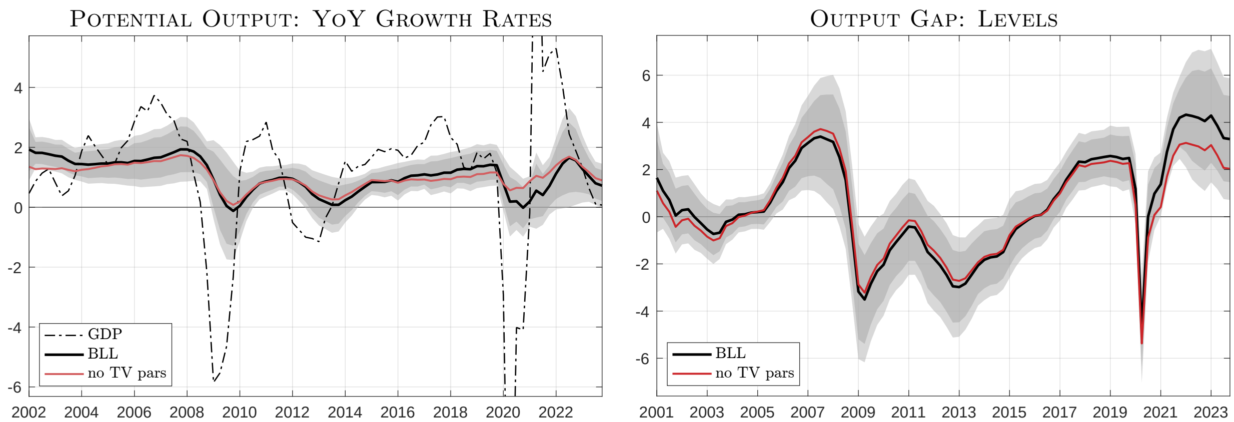
I No time-varying parameters

Figure I1 compares the benchmark estimates of the output gap (right plot) and YoY potential output growth (left plot) with those obtained without allowing for a local linear trend for GDP, household liabilities, and long-term loans and no time-varying mean for the unemployment rate. Removing the time variation in the secular trends leads to a flatter estimate of potential output growth in the post-pandemic periods. This result shows that allowing for a time-varying trend

[xi] In this case $\widehat{\boldsymbol{\xi}}_i$ will be a $T-1$ vector denoting at time t the level of the idiosyncratic component for $i \in \mathcal{I}_0$ and the growth rate of the idiosyncratic component for $i \in \mathcal{I}_1$.

for GDP is crucial to properly capture the slowdown in potential output in the latter part of the sample.

Figure I1: *Output gap estimate when imposing no time-varying parameters*



Notes: The black solid line is our benchmark estimate, the grey shaded areas are the 68% and 84% confidence bands, the black dashed line is GDP YoY growth rate, the red line is the estimate obtained without time-varying parameters (no TV pars). The level of the output gap is the percentage deviation from potential.

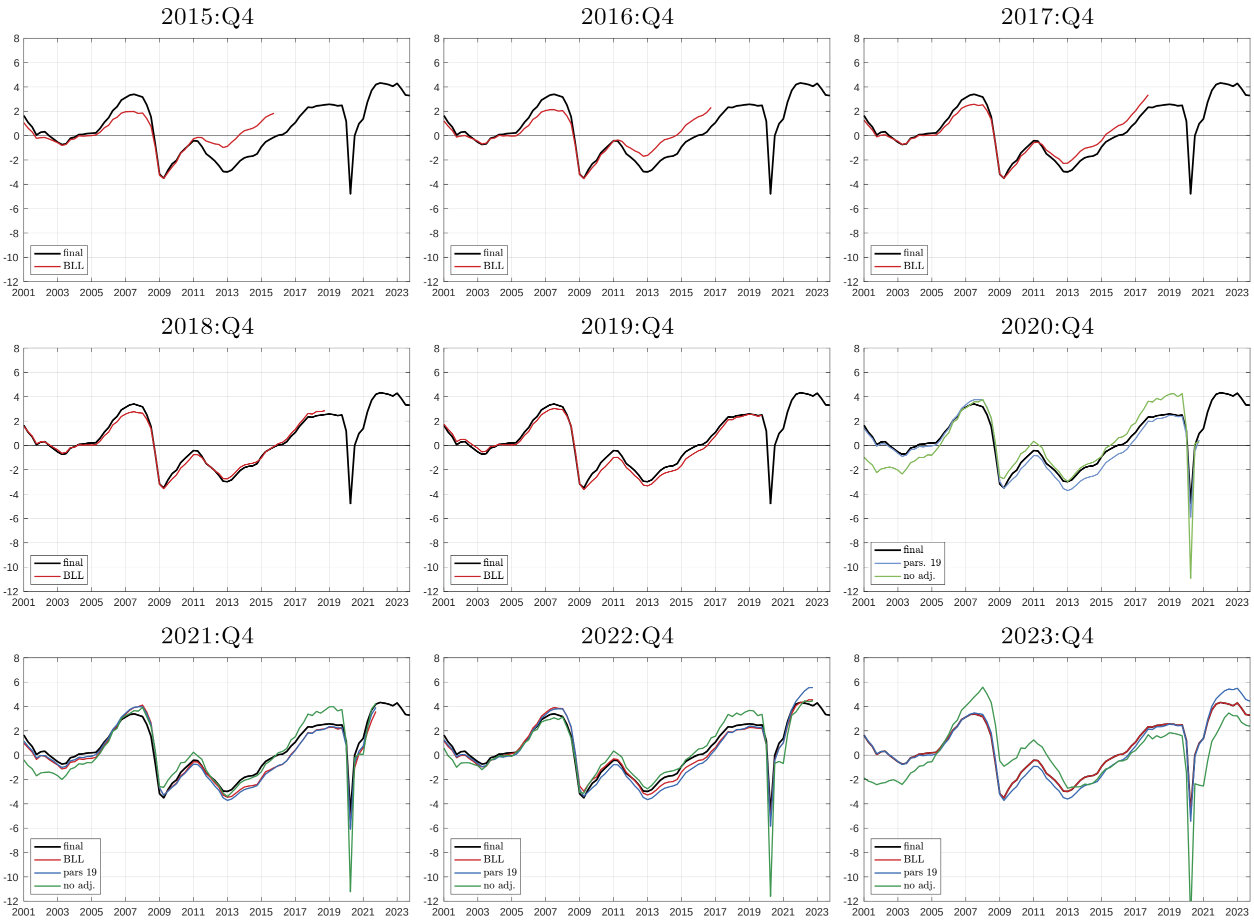
J Real-time reliability

There is skepticism in the literature on the reliability of model-based output gap estimates in real time because of the size of end-of-sample revisions (Orphanides and van Norden, 2002). Model-based estimates of the output gap are subject to revisions in real time because new information leads to changes in both the estimates of the model parameters and the latent states. In this section, we assess the reliability of our output gap estimate through a *quasi*-real-time exercise on expanding windows, where the first window begins in 2000:Q1 ends in 2015:Q1 ($T = 57$).

Looking at the upper charts in Figure J1, it is clear that the model is slow in recognizing how deep the 2012 Sovereign debt crisis is, only doing so once data for 2016 becomes available. Inasmuch as this result is disappointing, we cannot help but notice that 60 observations are probably too few to pin down the output gap accurately. As more information becomes available, the model's estimate of the output gap stabilizes, so much so that from 2017 onward, the *quasi*-real-time estimate is very close to the final estimate.

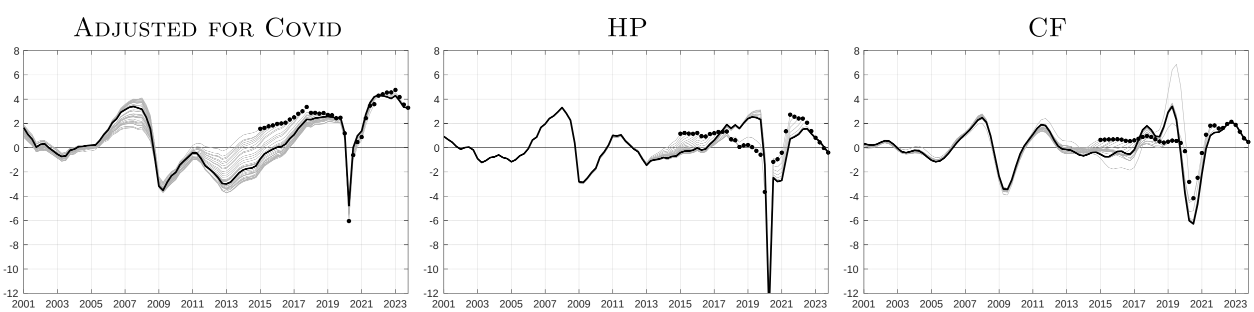
Moving to the Covid pandemic and its aftermath, it is evident that our strategy for adapting to the COVID shock was viable only starting from the second half of 2020 or even 2021. Moreover, it is reasonable to assume that anybody would have understood that doing nothing and allowing the Covid shock to affect the estimates was a huge mistake. Thus, we consider the *quasi*-real-time performance of the simple strategy of freezing the parameters to pre-Covid data. As shown by the green line, if we had followed this approach, we would have had very reliable output gap estimates, as opposed to extreme estimates, had we chosen to do nothing. Finally, by comparing the red and the green lines, we can appreciate the impact of the adjustment we implemented to account for the Covid shock.

Figure J1: *Output Gap: quasi-real time, expanding window*



Notes: The black line is the estimate of the output gap obtained using the final sample of data, up to 2023:Q4. The red line in Figure J1 displays the evolution of our *quasi*-real-time estimate of the output gap obtained using the procedure outlined in Section 3. The blue line shows the estimate we would have obtained without any adjustment for the Covid shock. Lastly, the green line is the estimate we would have obtained if we had frozen the parameter estimate at the value in 2019:Q4.

Figure J2: *Output gap: quasi-real time results*



Notes: the black line is the estimate of the output gap obtained over the full sample, and the thin grey lines are the estimate obtained on all the other subsamples. Each black dots represent the estimate of the output gap for quarter Q and year Y obtained on the sample ending at quarter Q and year Y.

Figure J2 compares the *quasi*-real-time estimate of the output gap from our model with those obtained from an HP filter and a Christiano-Fitzgerald band-pass filter. As mentioned earlier, our estimate's weakness is that the one obtained on the sample ending in 2015 is quite different from the final estimate. However, its strength is that from 2017 onward, the *quasi*-real-time estimates converge to the final and are very robust. In contrast, the HP filter and the Christiano-Fitzgerald filter yield *quasi*-real-time estimates that are never to far from the final estimate. However, they converge to the final estimate very late, making them unreliable for real-time analysis.

References

Ahn, S. C. and A. R. Horenstein (2013). Eigenvalue ratio test for the number of factors. *Econometrica 81*, 1203–1227.

Alessi, L., M. Barigozzi, and M. Capasso (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics & Probability Letters 80*(23-24), 1806–1813.

Antolin-Diaz, J., T. Drechsel, and I. Petrella (2017). Tracking the slowdown in long-run GDP growth. *The Review of Economics and Statistics 99*, 343–356.

Bai, J. (2004). Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics 122*, 137–183.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*, 191–221.

Bai, J. and S. Ng (2004). A PANIC attack on unit roots and cointegration. *Econometrica 72*, 1127–1177.

Barigozzi, M., M. Lippi, and M. Luciani (2021). Large-dimensional dynamic factor models: Estimation of impulse-response functions with $I(1)$ cointegrated factors. *Journal of Econometrics 221*, 455–482.

Barigozzi, M. and M. Luciani (2023). Measuring the Output Gap using Large Datasets. *The Review of Economics and Statistics 105*, 1500–1514.

Borio, C. (2014). The financial cycle and macroeconomics: What have we learnt? *Journal of Banking & Finance 45*, 182–198.

Canova, F. (2022). FAQ: How do I estimate the output gap? mimeo.

Casalis, A. and G. Krustev (2022). Cyclical drivers of euro area consumption: What can we learn from durable goods? *Journal of International Money and Finance 120*, 102241.

Cette, G., J. Fernald, and B. Mojon (2016). The pre-Great Recession slowdown in productivity. *European Economic Review 88*, 3–20.

Christiano, L. J. and T. J. Fitzgerald (2003). The bandpass filter. *International Economic Review 44*, 435–465.

Del Negro, M., D. Giannone, M. P. Giannoni, and A. Tambalotti (2017). Safety, liquidity, and the natural rate of interest. *Brookings Papers on Economic Activity Spring 2017*, 235–316.

Del Negro, M., D. Giannone, M. P. Giannoni, and A. Tambalotti (2019). Global trends in interest rates. *Journal of International Economics 118*, 248–262.

Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2007). On the fit of new Keynesian models. *Journal of Business & Economic Statistics 25*, 123–143.

Doz, C., D. Giannone, and L. Reichlin (2012). A quasi-maximum likelihood approach for large, approximate dynamic factor models. *The Review of Economics and Statistics 94*, 1014–1024.

Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*, 603–616.

Gordon, R. J. (2018). Declining American economic growth despite ongoing innovation. *Explorations in Economic History 69*, 1–12.

Hallin, M. and R. Liška (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association 102*, 603–617.

Hamilton, J. D. (2018). Why you should never use the Hodrick-Prescott filter. *The Review of Economics and Statistics 100*, 831–843.

Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.

Hodrick, R. J. and E. C. Prescott (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, 1–16.

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

Lenza, M. and G. E. Primiceri (2022). How to estimate a vector autoregression after March 2020. *Journal of Applied Econometrics 37*, 688–699.

Lippi, M. and L. Reichlin (1994). Diffusion of technical change and the decomposition of output into trend and cycle. *The Review of Economic Studies 61*, 19–30.

Maroz, D., J. H. Stock, and M. W. Watson (2021). Comovement of economic activity during the Covid recession. mimeo.

Morley, J., D. Rodríguez-Palenzuela, Y. Sun, and B. Wong (2023). Estimating the euro area output gap using multivariate information and addressing the COVID-19 pandemic. *European Economic Review 153*.

Morley, J., T. D. Tran, and B. Wong (2023). A simple correction for misspecification in trend-cycle decompositions with an application to estimating r. *Journal of Business & Economic Statistics*. available online.

Ng, S. (2018). Comments on the cyclical sensitivity in estimates of potential output. *Brookings Papers on Economic Activity 49*, 412–423.

Ng, S. (2021). Modeling macroeconomic variations after COVID-19. NBER working paper 29060.

Onatski, A. (2009). Testing hypotheses about the number of factors in large factor models. *Econometrica 77*, 1447–1479.

Opschoor, D. and D. J. van Dijk (2023). Slow Expectation-Maximization convergence in low-noise dynamic factor models. SSRN working paper 4408065.

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

Phillips, P. C. and Z. Shi (2021). Boosting: Why you can use the HP filter. *International Economic Review 62*, 521–570.

Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association 89*, 1303–1313.

Tóth, M. (2021). A multivariate unobserved components model to estimate potential output in the Euro Area: A production function based approach. ECB working paper 2523.

Zhang, R., P. Robinson, and Q. Yao (2019). Identifying cointegration by eigenanalysis. *Journal of the American Statistical Association 114*, 916–927.