# Estimating the Output Gap After COVID: How to Address Unprecedented Macroeconomic Variations *

Camilo Granados †
University of Texas at Dallas

Daniel Parra-Amado ‡
Banco de la República

October 18, 2023

**Abstract**

This study examines whether and how important it is to adjust output gap frameworks during the COVID-19 pandemic and similar unprecedentedly large-scale episodes. Our proposed modelling framework comprises a Bayesian Structural Vector Autoregressions with an identification setup based on a permanent-transitory decomposition that exploits the long-run relationship of consumption with output and whose residuals are scaled up around the COVID-19 period. Our results indicate that (i) a single structural error is usually sufficient to explain the permanent component of the gross domestic product (GDP); (ii) the adjusted method allows for the incorporation of the COVID-19 period without assuming sudden changes in the modelling setup after the pandemic; and (iii) the proposed adjustment generates apsproximation improvements relative to standard filters or similar models with no adjustments or alternative ones, but where the specific rare observations are not known. Importantly, abstracting from any adjustment may lead to over or underestimating the gap, to too-quick gap recoveries after downturns, or too-large volatility around the median potential output estimations.

*JEL Codes:* C11, C53, E3, E32, E37

*Key words:* Bayesian methods, business cycles, potential output, output gaps, structural estimation

---

*The authors thank the editor Sushanta K. Mallick, and the anonymous referees for their feedback, as well as Luca Benati for his guidance. Initial versions of this paper were developed during an academic visit to the Graduate Institute in Geneva and in the context of the Bilateral Assistance and Capacity Building for Central Banks (BCC) program. We also appreciate the comments and feedback of Cédric Tille, Jean-Paul Renne, Ivan Mendieta-Muñoz, Irina Panovska, Victor Valcarcel, Kirill Shakhnov, Benjamin Wong, as well as participants in the 7th International Workshop in Financial Markets and Nonlinear Dynamics (FMND) and the Society for Economic Measurement 8th International Conference (SEM). All remaining errors are ours.

†School of Economic, Political and Policy Sciences, University of Texas at Dallas. Richardson, TX 75080, USA. Email: camilo.granados@utdallas.edu

‡Corresponding author: Macroeconomic Modeling Department, Banco de la República, Colombia. Email: dparraam@banrep.gov.co, dparraa@unal.edu.co

# 1 Introduction

The COVID-19 pandemic unprecedentedly affected humanity, not only in terms of public health but also economically. Unlike other crises, economic deterioration has been globally synchronised. According to the International Monetary Fund, the world economy contracted by approximately $3.0\%$ in 2020, with both developed and emerging economies falling by $4.8\%$ and $1.9\%$, respectively. Ensuing recovery was relatively quick owing to the distribution of vaccines and the gradual opening during and after the lockdown. This allowed most economies to bounce back to positive growth rates in one to three quarters, which, given the magnitude of the initial downturn, led to the worldwide growth of $6.0\%$ in 2021. In summary, we witnessed extreme macroeconomic data, which also led to high levels of uncertainty.

The fluctuations observed naturally raise questions about the macroeconomic effects of the COVID-19 shock, particularly on variables such as the potential output and gap. On simple inspection, it is difficult to label a downturn of that magnitude as trivial for long-run variables. However, the pace of recovery makes it challenging to deem the shock as highly influential.

With this in mind, we aim to answer whether estimations of the output gap should be adjusted to account for the COVID-19 shock. We intend to determine a way to reconcile the magnitude of the shock with its transitory nature when approximating the potential output.

Our approach must, therefore, cover two fronts: first, how to obtain a good econometric framework for estimating the output gap, and second, how this can be adjusted in a manner that allows incorporating the COVID-19 shock information but prevents it from influencing the model as if it was representative of the data-generating process. For the first point, we rely on a permanent-transitory (PT) decomposition framework to identify the fluctuations of a set of macroeconomic variables (where the output is included) in a Bayesian Structural Vector Autoregression (BSVAR) setting; this is done by following Uhlig (2003, 2004). Based on the resulting model, we recover a path for the potential output that covers the COVID-19 period. For the second point, we adjust the model estimation with a scale factor around the rare shock date along the lines of Lenza and Primiceri (2022).

Primarily, we include an ample set of variables in our model that contrasts with the usual output gap estimation frameworks, such as univariate statistical filters or production function approaches.This enables us to include additional sources of information in our setup and account for the permanent income hypothesis through the relationship between consumption and long-run output, which, as mentioned by Cochrane (1994), facilitates identifying the permanent component of the output.

Nonetheless, the identification task in the context of SVAR models can be challenging, as these frameworks usually rely on imposing strong assumptions about the nature of shocks that can be too restrictive. For example, it is usual to impose that long-run output is driven only by supply shocks, while demand is only associated with transitory components (e.g., Barsky and Sims, 2011; Blinder and Rudd, 2013; Keating and Valcarcel, 2015; Chen and Gornicka, 2020). However, recent data has vindicated the potential long-term role of demand-driven phenomena; for example, in the Global Financial Crisis (GFC) and the protracted recovery that followed, a weak demand affected both the current output and its future expectations in such a persistent way that it shifted down the path of potential growth (Fontanari, Palumbo, and Salvatori, 2020).[1] The literature has followed suit and has recently pointed out that other shocks, such as demand (Furlanetto, Lepetit, Robstad, Rubio-Ramírez, and Ulvedal, 2021) and monetary (Jordà, Singh, and Taylor, 2020) shocks, can also have long-run effects.

The aforementioned consideration is even more valid in the context of the COVID-19 shock, which was considered a supply-driven shock, but eventually showed to involve demand-driven fluctuations.[2] We circumvent this issue of a separate identification of supply and demand shocks and their association with different terms by adopting an agnostic identification approach along the lines of Uhlig (2003, 2004), that is, based on the maximization of the explained fraction of long-horizon Forecast Error Variance (FEV) of the gross domestic product (GDP). Such an approach is particularly reasonable for gauging the potential output when we observe shocks such as the COVID-19 downturn that are perceived as a combination of supply- and demand-driven fluctuations (rather than either exclusively).

[1]This experience even led to revisiting the literature on hysteresis, such as Cerra, Fatás, and Saxena (forthcoming), Benati and Lubik (2021) and Aikman, Drehmann, Juselius, and Xing (2022).

[2]See Guerrieri, Lorenzoni, Straub, and Werning (2022) and Fornaro and Wolf (2020) for further discussion on demand to output spillovers and stagnation traps.

This identification scheme has been used by recent studies, such as Angeletos, Collard, and Dellas (2020) and Brignone and Mazzali (2022) and with the same objective of decomposing the permanent and transitory fluctuations of macroeconomic variables. We follow a similar approach while adjusting the econometric modelling along the lines of Lenza and Primiceri (2022), which allows us to incorporate the COVID-19 downturn in the sample but limits the impact of the rare event on the estimated parameters. The joint application of the identification setup and adjustment for high-magnitude shocks in the context of output gap estimations represents our contribution.

We apply our approach to the seven developed economies and find that a single structural shock is sufficient to characterise the long-run behaviour of GDP. By contrast, the remaining shocks tend to explain their transitory effects more significantly. This result aligns with the findings of Dieppe, Francis, and Kindberg-Hanlon (2021), Angeletos, Collard, and Dellas (2020) and Brignone and Mazzali (2022) for the US and European countries. Based on this result and the structural shocks, we approximate the GDP gap at each date as the weighted sum of the transitory shocks (and use the other shock to recover potential output) with weights based on the associated historical decomposition of the BSVAR model.

We address our main research question and compare the potential output (and gap) estimates with standard gap estimation methods and the BSVAR counterpart with no COVID-19 adjustment. We find that our proposal (PT identification with COVID-19 adjustment) prevents the potential output from falling too rapidly at the onset of the shock and does not induce a fast recovery in subsequent periods, a known drawback of the usual univariate filtering techniques.

Then, we compare our method with an alternative BSVAR with the same identification scheme but a stochastic volatility setup. This alternative is, in principle, also adjusting for the effect of the COVID-19 episode on the model. However, in contrast to our scaling method around the shock date, the adjustment is entirely endogenous because the variance is time-varying. This model leads to a stronger decrease in the potential output, but even by 2022 shows no sign of recovery, suggesting that the large magnitude of the shock could persistently affect the estimates. In light of this, our model represents a more appropriate alternative for a shock of large magnitude but small persistence that is less representative of the data-generating process of the sample.

Finally, we evaluate our method in a simulation setting to compare it with alternatives in a more general light, not only with the specific COVID-19 episode of reference in mind. The results indicate that the large-scale shock may induce standard filters to deliver negatively correlated estimators with the target (simulated "true" output gap). When that occurs, the structural identification (or the permanent component) of our model by itself corrects the issue, that is without any large-shock adjustment. However, only through the scale factor adjustment, we can obtain important gains in terms of the cross-correlations which at their peak can grow from less than 0.5 (with unadjusted methods) to beyond 0.8.

In summary, these results and exercises consistently indicate that the model's performance is not only associated with the structural identification of the BSVAR but also with the adjustment of the model to include the large shock. The benefits of adjusting the gap estimates in the presence of shocks of unprecedented magnitude are non-trivial. The performance gains terms are present even in models, successful at approximating the potential output. In addition, our setup prevents a strong decrease in the potential output after the outlying downturn and a quick recovery once its transitory nature is made evident; that is, it improves on the drawbacks of complex counterparts and standard filters.

**Related literature** Our paper is related to various strands of the literature. At large, this paper belongs to the large literature on the estimation of the output gap, and to a greater extent to those based on multivariate approaches.[3] More specifically, our paper is related to studies using PT decomposition-type of methods for estimating the output gap; among these papers, Angeletos, Collard, and Dellas (2020), Brignone and Mazzali (2022) and Dieppe, Francis, and Kindberg-Hanlon (2021) use the same approach of this paper, that is, based on explaining the highest possible share of the FEV of the output in the long-run, while studies such as Morley, Rodríguez-Palenzuela, Sun, and Wong (2023), Berger, Morley, and Wong (2023), and Berger and Ochsner (2022) use a Beveridge-Nelson (BN) type of decomposition based on the optimal forecast at long horizons. Our contribution relative to the first group of these studies is that we adjust our baseline model along the lines of Lenza and Primiceri (2022) to include the COVID-19 period in the sample. Simultaneously,

[3]For an overview of this literature see Álvarez and Gómez-Loscos (2018), Guisinger et al. (2018); and for a discussion about multivariate approaches see Cochrane (1990)

relative to the second group, rather than calling the optimal long-run forecast (obtained via BN decomposition) the potential GDP, we structurally identify a model where the share of the long-run variance is explained by a limited number of shocks.

This study also relates to the literature on the adjustment of econometric models to include COVID-19 periods. In particular, it closely follows the work of Lenza and Primiceri (2022) by scaling the model information around a researcher-specified date but allowing the scale factor parameters to be obtained in a Bayesian setting. Other studies proposing alternative adjustments in this direction are Hartwig (2022), Carriero, Clark, Marcellino, and Mertens (2022), and Ng (2021).

Having said this, it should be mentioned that this is not the only study leveraging in Lenza and Primiceri (2022) to adjust an output gap estimation method. This is also done in Morley, Rodríguez-Palenzuela, Sun, and Wong (2023). There, the authors adjust a VAR-X by the Covid-19 episode and afterward estimate the trend component of the output for the Eurozone using a BN decomposition. Our work is similar in that we apply jointly Lenza and Primiceri (2022) and an output gap estimation method. However, we differ in a number of relevant dimensions. First, we use a SVAR approach where we identify the structural errors driving the long-run component of GDP (from the forecast-error variance decomposition as in Uhlig, 2003, 2004), second, we estimate jointly the model and the scaling factors of Lenza and Primiceri (2022) in a Bayesian setting, and finally, we use a different notion of the potential output (and gap): we construct the potential output (gap) as the permanent (transitory) component resulting from the contribution of the structural shocks that explain the long-run (short-run) behavior of the GDP.

The remainder of this paper is organised as follows. We explain the methodology in Section 2. Section 3 describes our data and main results, including a comparison of the proposed estimates with those yielded by other methods. In Section 4, we evaluate the performance of the proposed method in a simulation exercise and conclude the paper.

## 2 Methodology

Our empirical strategy was divided into two stages. First, we fit a reduced-form Vector Autoregressive (VAR) model with a scale factor adjustment around the COVID-19 crisis as in Lenza and Primiceri (2022). This allows us to account for the increased variance in the macroeconomic variables around the shock date. Second, we recast our model into an SVAR form by identifying the main shock explaining the Colombian business cycle in the long run, which is done, along the lines of Uhlig (2003, 2004), that is, by maximizing the explained fraction of the total FEV of the GDP[4] at a long-run horizon (e.g. 15 or 25 years ahead).

In the first stage, following Lenza and Primiceri (2022), a scale factor $s_t$ is added to the VAR model's reduced-form residuals to capture the increased uncertainty during the COVID-19 crisis. $s_t$ is set to one in the sample period before the COVID-19 shock ($t^*$), $s_{t^*} = \bar{s_0}$, $s_{t^*+1} = \bar{s_1}$, $s_{t^*+2} = \bar{s_2}$ and $s_{t^*+j} = 1 + (\bar{s_2} - 1)\rho^{j-2}$ for $j \geq 3$.[5] The scaled (reduced-form) VAR model is given by:

$$Y_t = B_0 + B_1 Y_{t-1} + B_2 Y_{t-2} + \ldots + B_p Y_{t-p} + s_t u_t, \qquad u_t \sim N(0, \Sigma), \tag{1}$$

the COVID-19 outbreak dates back to the first quarter of 2020 ($t^* = 2020Q1$); therefore, $\bar{s}_0$ is estimated for that date, and $\bar{s}_1$, $\bar{s}_2$ for the next two quarters. Then, the scale factor decays at a rate $\rho$ for all future periods.[6] Thus, $\theta \equiv [\bar{s_0}, \bar{s_1}, \bar{s_2}, \rho]$ is the vector of additional parameters to be estimated jointly with those of the VAR $(B_0, B_1, \ldots, B_p, \Sigma)$. Equation (1) can be estimated as in Giannone, Lenza, and Primiceri (2015) by assuming the prior distributions of the coefficients to be conjugate Normal-Inverse Wishart and by including the scale factors into the posterior hyperparameters. They are jointly estimated using Bayesian techniques by drawing those parameters (that include $\theta$) from their posterior distributions within the context of a Metropolis-Hastings procedure. For the estimation, the priors of $\beta$ and $\Sigma$ used are,

$$\Sigma \sim IW(\Psi, d),$$

$$\beta | \Sigma \sim N(b, \Sigma \otimes \Omega),$$

[4]It is also possible on other variables, such as household consumption.

[5]This setup allows the scale factor to take three different values in the first three periods after the outbreak and then decay at a rate $\rho$ in subsequent periods. The assummed dynamics behind this scale structure seems in line with empirical evidence for the year after the onset of the pandemic. Furthermore, notice the parameters are expected to be non-negative, which also aligns with the estimation results and the resulting parameters' distributions shown in Section 3.

[6]As mentioned, alternative adjustments to COVID-19 data for VAR models have emerged in the literature in both frequentist and Bayesian frameworks, several of which are based on the inclusion of additional pandemic-related variables as controls (dummies or indicators). See Ng (2021), Carriero, Clark, Marcellino, and Mertens (2022), and Hartwig (2022).

where $\beta \equiv vec([B_0, B_1, \ldots, B_p]')$ and $\gamma \equiv (\Psi, d, b$ and $\Omega)$ are the hyperparameter vectors. As mentioned by these authors, the Minnesota prior assumes that each variable included in the model follows a random walk process, and the prior for $\theta$ is defined analogously as a Normal distribution centered at 1.[7] The posterior of $\theta$ is used to capture the dynamics of $s_t$, which is jointly evaluated with the posterior of $\gamma$ as proposed by Lenza and Primiceri (2022):

$$p(\gamma, \theta|Y) \propto p(Y|\gamma, \theta) \cdot p(\gamma, \theta).$$

When (1) is estimated, we proceed with the second state, consisting of identifying structural shocks ($\varepsilon_t$) linked to the reduced-form errors by an impact matrix $A_0$ such that $u_t = A_0\varepsilon_t$ and $\Sigma = A_0 A_0'$.. It should be noted that there is not a unique $A_0$ that satisfies these relationships. For any candidate matrix $A_0$ an alternative matrix $\ddot{A}_0$ exists that can be derived using an orthonormal matrix $Q$ where $A_0 = \ddot{A}_0 Q$ and $QQ' = I$; in that sense, our approach also falls within the "set-identification" category.

In this context (which is common to most BSVAR identification setups) we apply our specific identification strategy; that is, the maximum fraction of the long-horizon FEV, along the lines of Uhlig (2003, 2004). This method seeks a target $q_1$ that satisfies:

$$q_1 = \operatorname{argmax} q_1' M q_1 \equiv q_1' \sum_{h=0}^{k} \ddot{A}_0{}' C_h' (e_j e_j') C_h \ddot{A}_0 q_1,$$

$$\text{subject to} \qquad q_1' q_1 = 1,$$

where $q_1$ is a column of $Q$ that explains the k-step-ahead forecast error of the $j$-th variable in $Y_t$ (in our case, the log of GDP), whose variance is given by $M$. Simultaneously, as shown in Uhlig (2003), $q_1$ is the eigenvector associated with the largest eigenvalue of the matrix $M$. $e_j$ is a selector vector with zeros everywhere and a 1 in the $j$-th position, and $C_h$ is a component of the long-run impact matrix of the VAR associated to the horizon $h$.[8] The constraint guarantees that $q_1$ is a unit-length

[7] For stationary series, we must adopt a prior mean of zero for all coefficients. However, as in our case in this article, when we are working with non-stationary series, it is better to set the prior mean to one to shrink towards a random walk. Further details see Giannone, Lenza, and Primiceri (2015).

[8] Note that $C(L) = I + C_1 L + C_2 L^2 + C_3 L^3 \cdots + C_h L^h + \ldots$ and the moving average representation of the model -in terms of the reduced form residuals- is given by $Y_t = \mu + C(L)u_t$ where $\mu$ is the unconditional mean of of $Y_t$ implied by the VAR model.

column vector that belongs to an orthonormal matrix.

Notably, the method recovers all eigenvalues and eigenvectors of $M$, which, given the decomposition method, are ordered from higher to lower fractions, explained by the FEV of the target variable. Thus, we can verify whether one or more shocks explain a larger component of the long-run FEV of the GDP. In other words, this approach identifies the shock that best explains the long-run component of the target variable.[9] This is verified in the following section.

## 2.1 Output Gap determination within the structural model

The BSVAR and identification setup above yields the structural errors, and more importantly for our objective, associates these to the long-run (permanent) or short-run (transitory) components of the variables. Based on such information we can reconstruct the estimated potential output and output gap as follows.

First, we compute the historical decomposition (HD) of the model that expresses the observed variables as a weighted sum of all structural shocks:

$$y_{1,t} = HD^{init}_{y_{1,t}} + HD^{\varepsilon^{(1)}}_{y_{1,t}} + HD^{\varepsilon^{(2)}}_{y_{1,t}} + \cdots + HD^{\varepsilon^{(k)}}_{y_{1,t}} \tag{2}$$

where, $y_{1,t}$ denotes the first variable in $Y_t$, and $HD^{\epsilon^k}_{y_{1,t}}$ represents the contribution of the $k$-th structural shock ($\varepsilon^{(k)}$) to its dynamics, and $HD^{init}_{y_{1,t}}$ is the sample analog to the unconditional mean of the variable $y_1$. Notice that analogous expressions hold for any variable in $Y_t$. For the sake of exposition, let us denote $y_{1,t}$ as the observed output.

The contribution of each structural error expressed in (2) will depend on the (finite) MA representation of the resulting VAR model,[10] and each component will indicate the impact of one of the structural shocks realized in each period $t-j$ on the output in period $t$, for example, the

[9] It is not necessarily capable of replicating its entire FEV, but it explains a large proportion of it. In other words, even more shocks could be used to increase the percentage of explained FEV if a second or third shock is also found to explain the permanent component of the target.

[10] Note that for practical applications this is carried out with finite-horizon analogs or approximations to the usual MA($\infty$) or Wold representation of the model, in fact the Historical Decomposition itself is usually defined just as a finite horizon version of this representation.

contribution of the $k$-th structural shock in the output at $t$ is:

$$HD_{y_{1,t}}^{\epsilon^{(k)}} = \sum_{j=0}^{t-1} \theta_{1k}^{(j)} \epsilon_{t-j}^{(k)},$$

with $\theta_{1k}^{(j)}$ denoting the coefficient in the row 1, column $k$ within the $j$-th matrix in the MA representation of the associated VAR.[11]

Now, the decomposition in (2) is appealing if we care about the dynamics explained by each structural shock but by itself does not disentangle the potential output from the cycle dynamics. However, our identification setup indicates which structural errors drive the permanent component and which the transitory, making it possible to split the observed dynamics into each. For example in the case that only the first structural shock drives the permanent component and the others (second to $k$-th shock) the transitory, each component of the output (potential GDP and gap) is obtained as,

$$y_{1,t} = \overbrace{HD_{y_{1,t}}^{init} + HD_{y_{1,t}}^{\varepsilon^{(1)}}}^{y_{1,t}^{pot}} + \overbrace{HD_{y_{1,t}}^{\varepsilon^{(2)}} + HD_{y_{1,t}}^{\varepsilon^{(3)}} + \cdots HD_{y_{1,t}}^{\varepsilon^{(k)}}}^{y_{1,t}^{gap}} = y_{1,t}^{pot} + y_{1,t}^{gap}, \tag{3}$$

where $y_{1,t}^{pot}$ would denote the permanent component of output, or the potential output, and $y_{1,t}^{gap}$ the transitory component, or output gap. It should also be noted that this method for obtaining the gap is analogous to the one used to compute these quantities by other VAR based methods such as Blanchard and Quah (1989), and Chen and Gornicka (2020), the latter based on the identification scheme of Forbes, Hjortsoe, and Nenova (2018)

## 2.2 Data

We assemble a dataset encompassing nine variables for seven developed economies:United States (USA), Canada (CAN), Australia (AUS), United Kingdom (GBR), Germany (GER), France (FRA) and Italy (ITA). Our model incorporates typical economic variables, see for example Morley et al. (2023), but we prioritize parsimony in dealing with the inherent dimensionality complexities of the multivariate setup. Although our main analytical focus lies at the end of the sample when the

[11]Such representation is the HD itself and can also be expressed as $Y_t = HD_y^{Init} + \sum_{j=0}^{t-1} \Theta^{(j)} A_0 \varepsilon_{t-j}$

COVID-19 shock happened, we extend the sample as much as possible for each country depending on the available information.[12] The variables included are the Gross Domestic Product (GDP), household consumption (CON), government consumption (GOV), and investment (INV), all in real terms and obtained from the respective national accounts data. We also include inflation (INF), real exchange rate (EXR), short-term interest rate (INT), long-term interest rate (LINT), and the real Brent oil price (OIL). Most of the data was obtained from the Organisation for Economic Co-operation and Development (OECD) website, central banks, and the statistical institutions of each country. Additional remaining variables are obtained from Datastream and Bloomberg.

For purposes of the estimation of the actual gap (reported in sections 2.3 and 3), we use the complete set of information (nine variables), and fit i) the model explained in Section 2 that corresponds to our baseline model (LP-Adjusted) based on a permanent-transitory (PT) identification (Uhlig, 2003, 2004) with a COVID-19 adjustment (as in Lenza and Primiceri, 2022), ii) a BSVAR model with PT identification but no COVID-19 adjustment (denoted PT-Decomp), and iii) a BSVAR model with a PT identification and a stochastic volatility setup, i.e, without depicting an explicit correction at the exact date of the COVID-19 shock (named as PT-Stocvol). On the other hand, for the evaluation stage (section 4), we use only three of the time series in each case: the real output, inflation, and interest rate, and then, we run a Monte Carlo simulation based on a New Keynesian model as in Benati (2008). Finally, in the estimation and evaluation sections we also consider standard univariate filters (HP and Christiano-Fitzgerald bandpass filter).

### 2.3 Empirical strategy

We set a nine-variable B-SVAR in levels for each economy with a lag length (in most cases of $p = 2$) choice given by the Bayesian and Hannan-Quinn Information criteria, and estimate the VAR in levels using a hierarchical modelling approach that allows us to make inferences about the informativeness of the prior distribution of the BSVAR, as proposed by Giannone, Lenza, and Primiceri (2015) which automatically determines a suitable measure of the shrinkage by considering a combination of conjugate priors such as a Minnesota prior and tighter priors when the model

[12] The sample sizes in each case are as follow: USA: 1964Q1:2023Q1, CAN: 1996Q1:2023Q1, AUS: 1994Q1:2023Q1, GBR: 1988Q1:2023Q1, GER:1991Q1:2023Q1, FRA:1980Q1:2023Q1, ITA:1994Q1:2023Q1.

includes many coefficients relative to the number of observations. As part of the procedure, we ran 20000 draws and kept half for estimation after burn-in. In addition, we explicitly modelled the COVID-19 extreme observations, as in Lenza and Primiceri (2022). From this first stage, we obtain a reduced-form VAR that has already been adjusted by the scale factor $(s_t)$ and incorporates the pandemic shock.

In the second stage, we identified the impact of the matrix of the SVAR by maximizing the explained share of the forecast variance error of the GDP for a 25-year horizon, as in Uhlig (2003, 2004). At the same time, we restrict that the share of the FEV one step ahead of consumption explained by the first structural error, or (the majority of the) permanent component, is larger than that of the output and for the latter to be larger than that of the investment. As explained by Cochrane (1994) and King, Plosser, Stock, and Watson (1987), this accounts for the fact that consumption is more closely aligned to the permanent component of GDP, while investment should reflect its most volatile and transitory components. After verifying these restrictions and keeping the draws that comply with them, we conducted PT decomposition and computed the permanent (and transitory) output component (as mentioned in section 2.1).[13]

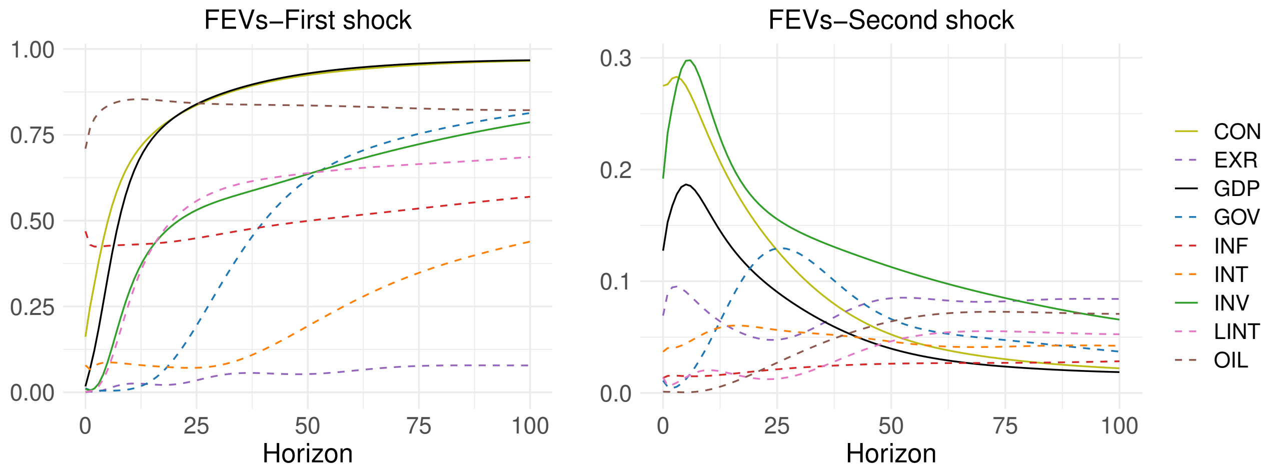
As aforementioned, the decomposition and resulting impact matrix already considers the ordering of the structural shocks according to their share of the explained variance of the target variable.[14] This can be verified in Figure 1, where we can see that for the US economy,[15] only the first structural error is necessary to account for approximately 95% of the long-run (permanent) component of the GDP. In contrast, the next most important shock in explaining the GDP's long-run FEV is instead better associated with the short-run or transitory component. The same result holds for the other economies in our sample where the first shock explain GDP in long-run between 59% and 96% (Figure 10 in Appendix A.2).

[13] As a check, we increased the number of draws to 100.000 and obtained similar results.

[14] This implies that the first structural shock reported by the method is the one with the largest share of explained variance of the long-run GDP, followed by the shock with the second largest share, and so on, until the last shock that explains the smallest share.

[15] In some cases we show the results only for the US for the sake of a clearer and simpler exposition. However, the analogous figures and results for the other economies are included in the appendix.

**Figure 1:** Shocks explaining the highest share of long-run GDP FEV (USA)




Note: The Figure shows the Forecast Error Variance (FEV) explained by the two structural shocks with the largest explained share for the GDP in the long-run. Given a single shock explains almost 100% of GDP for large horizons (left-panel) it is associated as the main driver of the permanent component of output, while the second highest (right panel) and remaining shocks are instead considered as driving the transitory component of output.

In light of these results, we compute the output gap based on the historical decomposition components attributed to the second to ninth structural shocks – i.e., those explaining only the transitory component of GDP – and use only the first one to recover the potential GDP.[16] In other words, we compute the output gap by rerunning the baseline BSVAR model (LP-Adjusted) and shutting off the first structural shock. Thus, the transitory component is estimated by using the other eight shocks.[17] In on a related point, it should also be noted that the first structural error will explain the majority of the long-run FEV of the GDP (target variable), but not necessarily the largest share of the FEV for other variables. The relative importance of the shocks to the other variables can be seen in the FEV decomposition per variable, as shown in Figure 13 for the US economy and for the other six economies from Figure 14 to 19 in the Appendix A.5.[18]

[16]Analogously, the potential GDP can be obtained as the original series minus the transitory component.

[17]The methodology is not limited to the selection of a single shock to identify the permanent component of GDP in the long run. For example, in the case of Italy, it could consider employing the first two shocks (84%) as the first shock explains 59% of the long-term GDP variations, while the second accounts for nearly 25%. In this case, the transitory component (output gap) is explained by using seven shocks (from the third to ninth structural shocks).

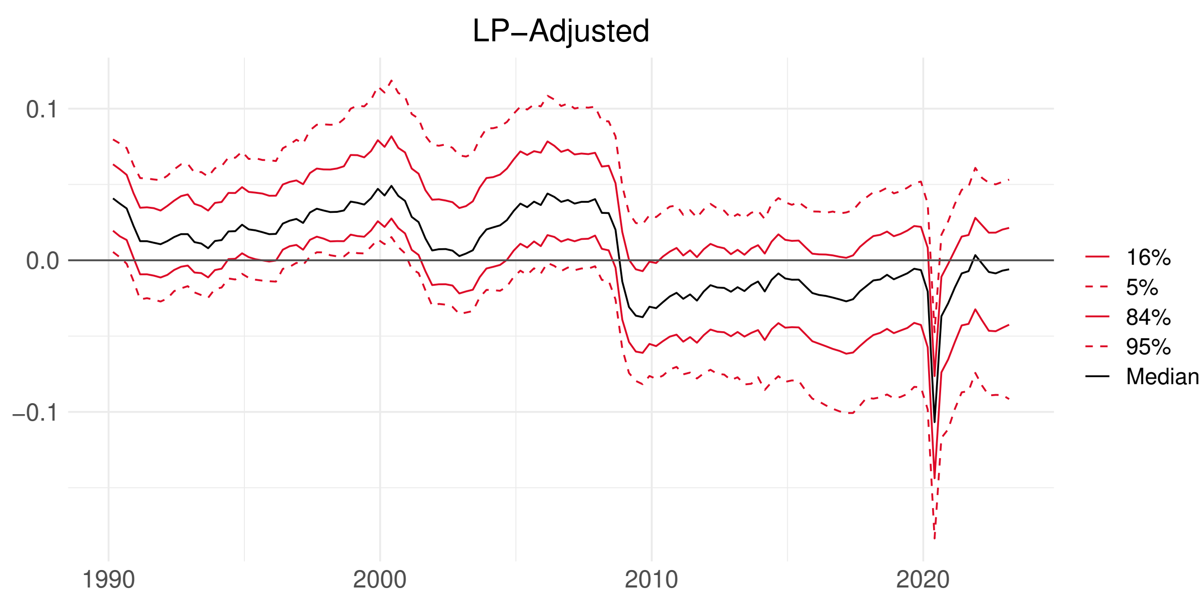
[18]We leave additional results that are related to other variables and shocks for the appendix, as we are only concerned with approximating the target variable, but also because the trade-off of this method is that you compromise a structural interpretation of the shocks as separable types of drivers (e.g. monetary, financial, global, local, supply, or demand, among others), as by construction, the method only gauges the overall importance of shocks at different horizons.

# 3 Results

## 3.1 Baseline Results

Figure 2 shows the output gap for the US economy obtained from our proposed baseline BSVAR model. This approach incorporates both a PT decomposition and a scale factor adjustment which deals with the observations during the COVID-19 period (labeled as LP-Adjusted). Before COVID-19, our estimated output gap reflected the recessions of the early 1990s and 2000s and the global financial crisis (GFC) of 2008.[19] Another notable feature of our method is its ability to illustrate the uncertainty associated with the estimate. As observed, there is a significant increase in volatility during the COVID-19 period in comparison to previous periods.

**Figure 2:** Baseline results: Output gap for US economy




Notes: The Figure shows the results of the baseline BSVAR model with Permanent-Transitory decomposition and Covid-19 adjustment (LP-Adjusted). The solid black line represents the median estimates. The solid and dotted red lines represent the percentiles of 5%, 16%, 84% and 95%, respectively.

During the COVID-19 pandemic, the US gap underwent a steep decline ($-10.6\%$) in the second quarter of 2020; however, unlike in the 2008 recession, the downturn was not persistent. Instead, it bounced back in the following quarters. As in most economies, the decrease is largely explained by

[19]Those crises correspond to the shaded areas in Figure 3 associated to the turning points determined by the National Bureau of Economic Research (NBER) for recession. The NBER recession data are available at http://www.nber.org/cycles/cyclesmain.html

lockdown measures, while the recovery is induced by the gradual reopening of the economy. It is worth highlighting that, unlike the aftermath of the 2008 financial crisis that led to a negative gap persisting for more than a decade, a partial closure of the US GDP gap was observed following the COVID-19 pandemic. A similar pattern is witnessed in the case of the other six economies, although many of them show a positive output gap in 2023Q1 (see Figure 11 in Appendix A.3). The output gap fell to -13.6%, on average for the seven countries, during the second quarter of 2020. The largest drop was observed in the United Kingdom (fell to $-24\%$), while the least affected was Australia ($-7.8\%$).

### 3.1.1 Comparison with alternative estimation methods

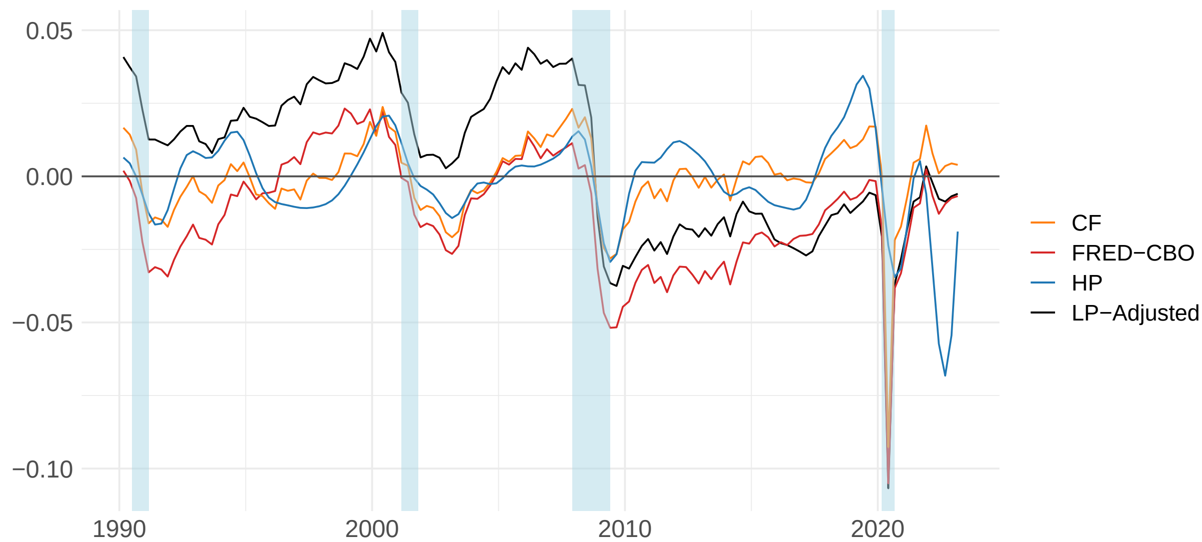
We also compare our estimations with those generated by usual filtering techniques, namely the Hodrick-Prescott (HP) and Christiano-Fitzgerald (CF) filters, as well as to an estimation computed using the real potential GDP estimated by the U.S. Congressional Budget Office (FRED-CBO).[20] The output gap estimates for the compared methods and our proposal are shown in Figure 3. We can see that the univariate filters (HP, CF) tend to deliver a large gap right before COVID-19 and a rapid and sizeable subsequent recovery, which sends that gap onto positive territory (and at or beyond $2\%$) in a few quarters. These features may indicate an overestimation of the gap, specifically when we see that the other estimates, including our proposal, do not display such behaviour, and instead suggest a dynamic yet more moderate recovery. Notably, when tying these results to the associated potential output dynamics, these results indicate that our proposal does not lower the potential output significantly during the period, which is related to adjusting the model to incorporate the COVID-19 observations in the estimation sample without assuming drastic changes in its data-generating process.

Concerning the official CBO estimate, it is important to note that our model consistently generates a negative gap after the 2008 financial crisis, in accordance with the economic consensus which is related to recent studies on hysteresis and the scarring effects of protracted recessions (e.g., Cerra, Fatás, and Saxena, 2023; Aikman, Drehmann, Juselius, and Xing, 2022). Similarly, using the CBO estimate as a benchmark, the correlation with our model's estimated is notably higher than

[20] We take the data from the FRED webpage (https://fred.stlouisfed.org/graph/?g=f1cZ)

the filters (86% vs 30% and 75%). Similar findings are obtained for the other six economies (see Figure 11 in Appendix A.3).

**Figure 3:** Comparison methodologies for output gap estimation (USA)



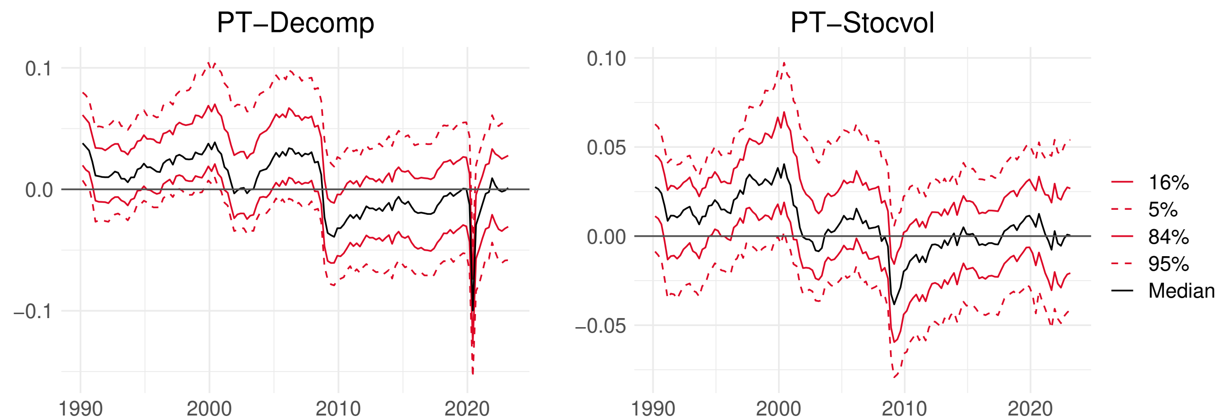

Notes: The black line represents the median of the baseline model (LP-Adjusted). Tthe orange and blue represents the Hodrick-Prescott (HP) and Christiano-Fitzgerald (CF) univariate filters. The red one is the output gap using the real potential GDP estimated by the U.S. Congressional Budget Office (CBO). Shaded areas indicate U.S. recessions and the COVID-19 period.

Finally, we compare the proposed BSVAR model (LP-Adjusted) with two models using the same type of identification setup (PT decomposition as in Uhlig, 2003). The first alternative is a BSVAR without a scaling adjustment for the COVID-19 episode (PT-Decomp), and the second is a model where, instead of using a scale factor at a known date, we allow for a stochastic volatility of the errors (PT-Stocvol).

The associated output gaps of the two BSVAR alternatives are shown in Figure 4. A large contrast between the baseline and the alternatives emerges at first sight: both the PT-Decomp and the PT-Stocvol generate a less negative gap during the COVID-19 outbreak, which implies that the potential output is affected more drastically relative to our baseline model. In that sense, as with some of the simpler filters, the alternatives tend to overestimate the impact of the shock on the long-run output.

Regarding the volatility around the estimates, the PT-Decomp displays the largest uncertainty, as reflected by wider percentile ranges than in the baseline. Nevertheless, the PT-Stocvol successfully mitigates volatility (yielding at a similar range as the baseline); however, it is the method where the estimated potential output is affected the most during the downturn. Therefore, we do not obtain an adequate estimate of the output gap by using the stochastic volatility setup. Comparable results are obtained for the remaining six economies.

**Figure 4:** US Output gap - two alternative BSVAR models



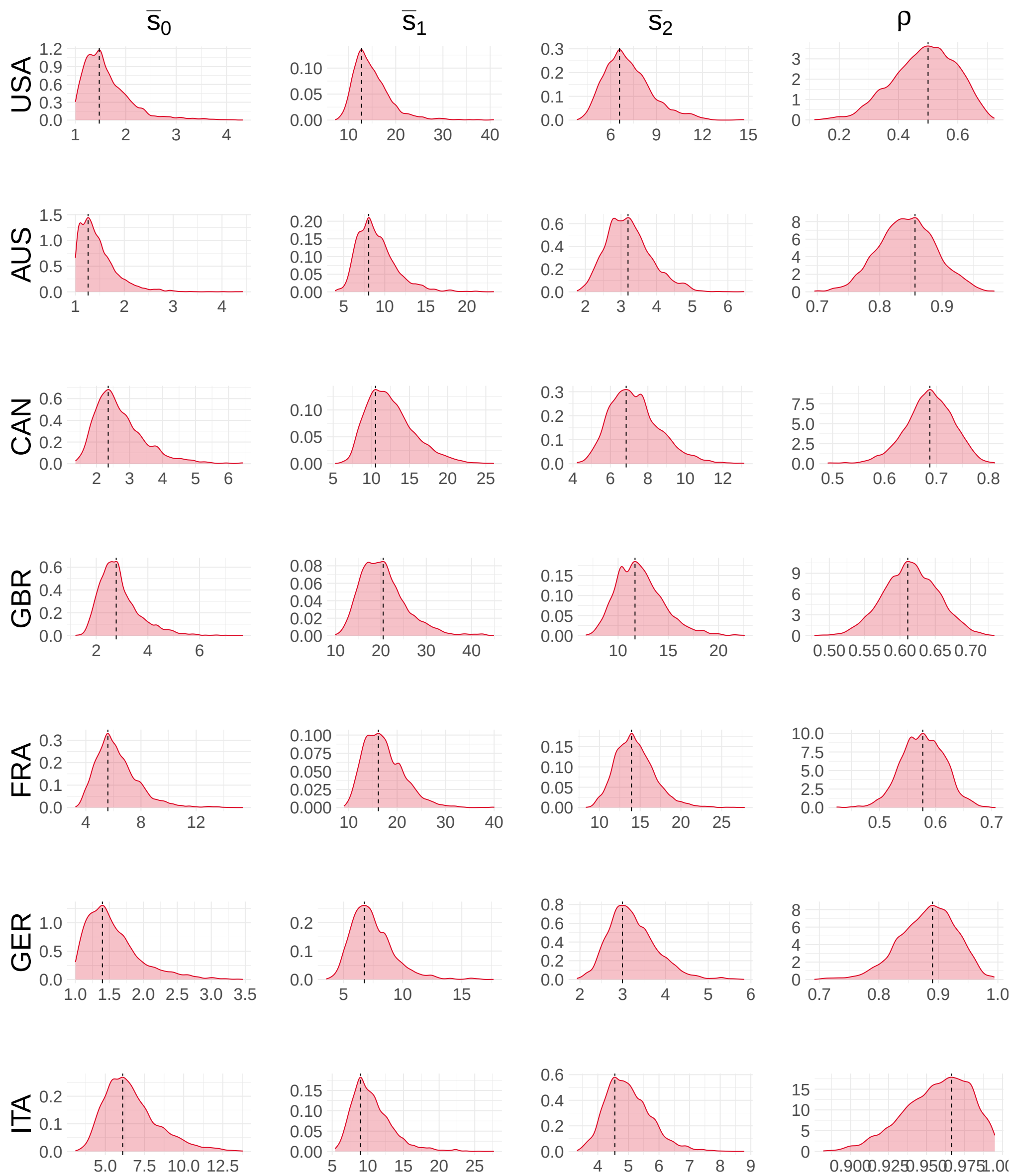

Notes: Left panel shows the results of the BSVAR model with Permanent-Transitory decomposition and no Covid-19 adjustment (PT-Decomp). Rigth panel shows the results of the Stochastic Volatility BSVAR with Permanent-Transitory decomposition and no Covid-19 adjustment (PT-Stocvol). The solid black line represents the median estimates. The solid and dotted red lines represent the percentiles of 5%, 16%, 84% and 95%, respectively.

### 3.1.2 Outlier observations around the COVID pandemic

Given that our main concern is to study the adjustment of potential output estimates to drastic magnitude shocks, such as those observed in the COVID-19 outbreak, verifying the estimates of the scale factors generated by our baseline estimates can be insightful. Principally, if scaling is irrelevant, the posterior estimates should suggest $\bar{s}_0 = \bar{s}_1 = \bar{s}_2 = 1$; otherwise, they should be sizeable. We estimate these parameters, as in Lenza and Primiceri (2022), and present our estimate of scale factors in Figure 5.

The parameters posteriors are drawn based on a Metropolis Hastings algorithm with a Minnesota Prior. Thus, we estimated the scaling factors together with other hyperparameters in a hierarchical structure. The resulting posteriors for $\bar{s}_0, \bar{s}_1, \bar{s}_2$ peak around 1.5, 13, and 7, respectively, indicating that, in effect, it is relevant for this sample to scale up the errors around the COVID-19 observations to account for the steep increase in volatility of that period, but that may not characterise its data-

**Figure 5:** Posterior distribution of the overall volatility scaling factors



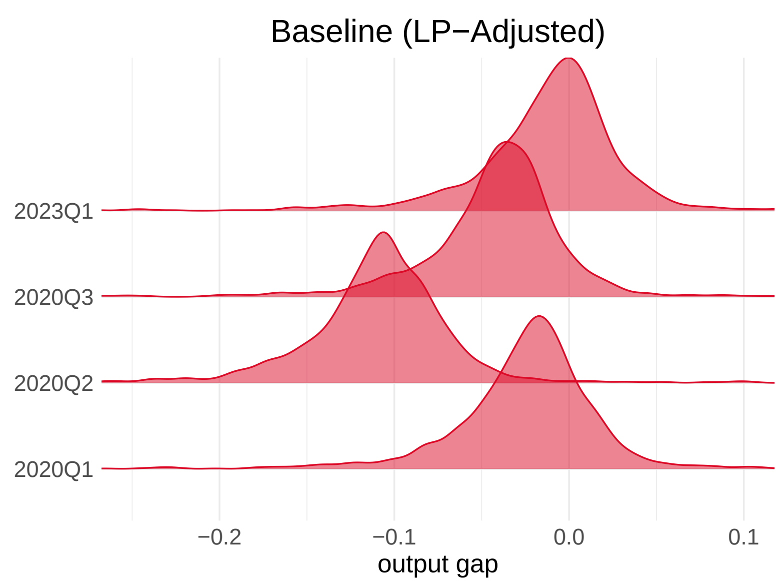

Notes: Each column represents the scale parameters $\theta \equiv [\bar{s_0}, \bar{s_1}, \bar{s_2}, \rho]$. Each row represents the results for each country. The vertical dotted black line marks the the mode of the distribution.

generating process, nor should it drastically influence the BVAR estimates. A country-by-country analysis shows that Italy and France were the most adversely impacted in 2020Q1, exhibiting an estimated parameter $\bar{s}_0$ nearly twice as high as that of the rest of the economies. We must keep in mind that, at the pandemic's outset, these nations were the most severely affected by the virus. In a similar vein, the UK demonstrates an increased volatility during the second and third quarters of

2020 in comparison to the other countries. Regarding the posterior of the decay coefficient ($\rho$), in the US case, it peaks around 0.5, which, together with $\bar{s}_2$, implies that the volatility scale factor falls by a half after 2020Q3 and afterward non-linearly towards one. The coefficient $\rho$ estimated for Italy is the highest, which implies that the COVID-19 shock had a greater persistence in that country compared to other economies.

To further illustrate the impact of the COVID-19 shock on the output gap, we can depict the distributions of the draw estimates for dates around the episode, as shown in Figure 6. We reveal the quarter of the shock (2020Q1), the subsequent two quarters, and the first quarter of 2023 as a reference for a date when the potential output dynamics are, in principle, back to normal (here implicitly recognise the transitory nature of the pandemic shock).

**Figure 6:** Distribution of the US output gap estimation during COVID-19 shock and 2023Q1.




Notes: The green line represents the median of the baseline model (LP-Adjusted). Tthe orange and blue represents the Hodrick-Prescott (HP) and Christiano-Fitzgerald (CF) univariate filters. The red one is the output gap using the real potential GDP estimated by the U.S. Congressional Budget Office (CBO). Shaded areas indicate U.S. recessions.

As we can see in the figure, the distribution of the gap has a large shift to the left, implying that the potential GDP was not largely affected by the downturn (and instead, the gap lowered in line with the observed GDP). In addition, the distribution spread increased, reflecting an increase in uncertainty around the estimate during the pandemic. Afterwards, we observe the distribution shifts back to pre-COVID-19 levels, although it still reflects increased volatility. In summary, we can see that the impact on the mean gap was transitory, although a somewhat larger uncertainty remains. Nonetheless, the larger uncertainty is approximately one percentage point higher than before, rather than orders of magnitude larger, as may be induced by a model without a scale factor adjustment for the COVID-19 downturn.

## 4 Evaluation of the method

Evaluating the relative performance of our estimates is a hard task, given that our target, the potential GDP, is an unobserved variable. There is no well-defined target against which to perform a "horse race" using a set of competing methods.

However, an assessment of these methods is in order, and alternative evaluation methods can be proposed. These usually imply assuming knowledge of relevant features of the actual potential output that can be tested. One route taken by the literature (e.g., Chen and Gornicka (2020)) comprises setting up a Phillips Curve with the output gap on the right-hand side of the equation; subsequently, an estimation method of the output gap is assessed according to its capacity to forecast inflation in the context of the Phillips Curve. Here, we assume that the output gap is a relevant variable for determining inflation and that the relationship captured in the Phillips curve is stable over time; that is, the curve setup is an appropriate device for testing the relationship between the output gap and inflation.

Although that is a feasible venue, it also opens discussions about how stable the Phillips Curve in each country is considered, whether such a relationship exists (e.g., McLeay and Tenreyro, 2020) or if its slope has flattened over time (Hazell, Herreño, Nakamura, and Steinsson, 2022). These discussions are relevant more in recent times when the trade-off between output stabilisation and inflation is strongly felt worldwide. However, such debates are beyond the scope of our study and may divert attention from what we aim for in this study, approximating the potential output.

Alternatively, and along similar lines as Canova (2020), we take a more direct approach and assume to count with a real measure of the potential output and then approximate it with a set of methods whose estimates are assessed based on their co-movement with the actual output gap. We do this in the context of a Monte Carlo simulation, where the set of economic variables and potential output is simulated based on an economic model taken as given.

### 4.1 The model used to simulate the output gap

We consider a standard three-equation New Keynesian DSGE model along the lines of Benati (2008), but where the output is assumed to have a unit root component that behaves as a random walk with a drift:

$$y_t^P = \delta + y_{t-1}^P + v_t, \qquad v_t \sim WN(0, \sigma_v^2), \tag{4}$$

The associated log-linearized model is given by:

$$\pi_t = \frac{\beta}{1+\alpha\beta}\pi_{t+1|t} + \frac{\alpha}{1+\alpha\beta}\pi_{t-1} - \kappa\hat{y}_t + u_t, \qquad u_t \sim WN(0, \sigma_u^2), \tag{5}$$

$$\hat{y}_t = \gamma\hat{y}_{t+1|t} + (1-\gamma)\hat{y}_{t-1} - \sigma^{-1}(R_t - \pi_{t+1|t}) - (1-\gamma)\Delta y_t^P, \tag{6}$$

$$R_t = \rho R_{t-1} + (1-\rho)\left[\phi_\pi \pi_t + \phi_y \hat{y}_t\right] + \epsilon_{R,t}, \qquad \epsilon_{R,t} \sim WN(0, \sigma_R^2). \tag{7}$$

The first two equations, the hybrid Phillips curve, and dynamic IS feature both backward- and forward-looking components, whereas the monetary policy rule, is given by a Taylor rule with smoothing. $\pi_t$ is inflation, $R_t$ is the nominal rate, and the real GDP is $Y_t$ which in the model is rescaled by its unit root component ($Y^P$) as $\hat{y}_t = \ln\left(Y_t/Y_t^P\right)$ to achieve stationarity. The latter implies that $\hat{y}_t$ is the output gap or the output as a deviation of the potential GDP given by its stochastic trend. The other variables were set as log deviations of their non-stochastic steady-state values.

The parameters of the model, $\Theta = \{\sigma_R^2, \sigma_u^2, \sigma_v^2, \kappa, \sigma, \alpha, \gamma, \rho, \phi_\pi, \phi_y\}$, were estimated for each considered economy using Bayesian methods. The posterior mode is found via simulated annealing, as in Benati (2008), and the posterior distribution of $\Theta$ is characterized by implementing a Random-Walk Metropolis-Hastings algorithm, as in An and Schorfheide (2007). Both simulated annealing and Metropolis simulations require the evaluation of the likelihood (and posterior) of the model based on its Sims canonical form and associated state-space representation.

Table 1 shows median -across country-specific estimates- the parameters' priors, posterior modes, and percentiles obtained in our estimations. An additional step in the simulations is the scale factor adjustment of the variances, which is revised every 10% of the iterations and adjusted depending on the fraction of accepted draws in the subset draws. With that, the average acceptance ratio across -country-specific- simulations is 0.219.

**Table 1:** Prior, Posterior modes and standard deviations for the parameters (median values)

| | | Prior | | Posterior | |
|---|---|---|---|---|---|
| Parameter | Prior Density | Mode | Standard Deviation | Mode | 90% coverage percentiles |
| $\sigma_R^2$ | Inverse Gamma | 0.01 | 0.01 | 0.0008 | [0.0007, 0.00011] |
| $\sigma_u^2$ | Inverse Gamma | 0.01 | 0.01 | 0.0010 | [0.0008, 0.0012] |
| $\sigma_v^2$ | Inverse Gamma | 0.01 | 0.01 | 0.0012 | [0.0010, 0.0014] |
| $\kappa$ | Gamma | 0.10 | 0.10 | 0.191 | [0.120, 0.322] |
| $\sigma$ | Gamma | 1 | 2 | 4.294 | [3.036, 7.891] |
| $\alpha$ | Beta | 0.90 | 0.05 | 0.901 | [0.776, 0.952] |
| $\gamma$ | Beta | 0.50 | 0.25 | 0.750 | [0.690, 0.854] |
| $\rho$ | Beta | 0.7500 | 0.10 | 0.688 | [0.618, 0.762] |
| $\phi_\pi$ | Gamma | 1.50 | 0.25 | 2.278 | [1.888, 2.687] |
| $\phi_y$ | Gamma | 0.50 | 0.15 | 0.589 | [0.389, 0.826] |

Note: The average acceptance ratio of the Metropolis algorithm across countries estimates is 0.219. The values reported correspond to the median values across country-specific estimations of the economic model.

## 4.2 Evaluation method of the output gap estimations

Based on each estimated New Keynesian model (one for each country in our sample), a Monte Carlo simulation is carried out, where, in each iteration, a sample (33 years long) of the model variables is simulated, and a corresponding "true" output gap is obtained. The simulated observable economic variables are then used as inputs for a set of competing econometric methods that estimate the output gap of the simulated model. For each iteration, the cross-correlation between each econometric estimate of the output gap and the simulated -actual- output gap of the model is calculated and recorded.

In other words, in each simulation, we use the economic model to obtain an output gap and other consistent economic variables with the former;[21] and then we feed the econometric methods -that includes our proposal- with the observable economic variables to generate an estimated output gap. Finally, we assess the estimates in terms of the co-movement between their estimated gap and the actual gap (simulated).

The methods compared are: (i) our proposal, a Permanent-transitory decomposition with a Lenza and Primiceri (2022) type adjustment for large shocks episodes with a known date (LP-Adjusted), (ii) a Permanent-Transitory decomposition via a BSVAR (PT-Decomp), (iii) a Hodrick-Prescott filter (HP), and a Christiano-Fitzgerald Band Pass filter (CF). The latter two filters are more frequent and widely available methods of estimation of the potential output, while the Permanent-Transitory decomposition is relatively more complex as it aims to achieve a structural identification for an SVAR based on the long-run forecasts of the output. Finally, our proposed method combines the structural long-run forecast identification approach with an adjustment of the estimation to account for the presence of very large shocks whose date is known.

Importantly, the identification method in the SVAR models methods (i) and (ii) (or LP-Adjusted and PT-Decomp) are the same. Thus, we are carrying two tests here, first, whether it is worthwhile to focus on a structural identification method despite its higher complexity, and second, if in addition, it is relevant to adjust the estimates of the model for the presence of very large shocks.

Finally, to make the experiment more relevant in our context of interest (an economy affected by an unprecedented scale shock), we incorporate a large-scale shock in each simulation that mimics the dynamics of the COVID-19 episode, that is, we apply a high-magnitude negative productivity shock at the end of each simulated sample, that is short-lived, and then after impacting the variables (e.g., the output) it allows them to approach their previous trend values by the last date but without depicting a full recovery.

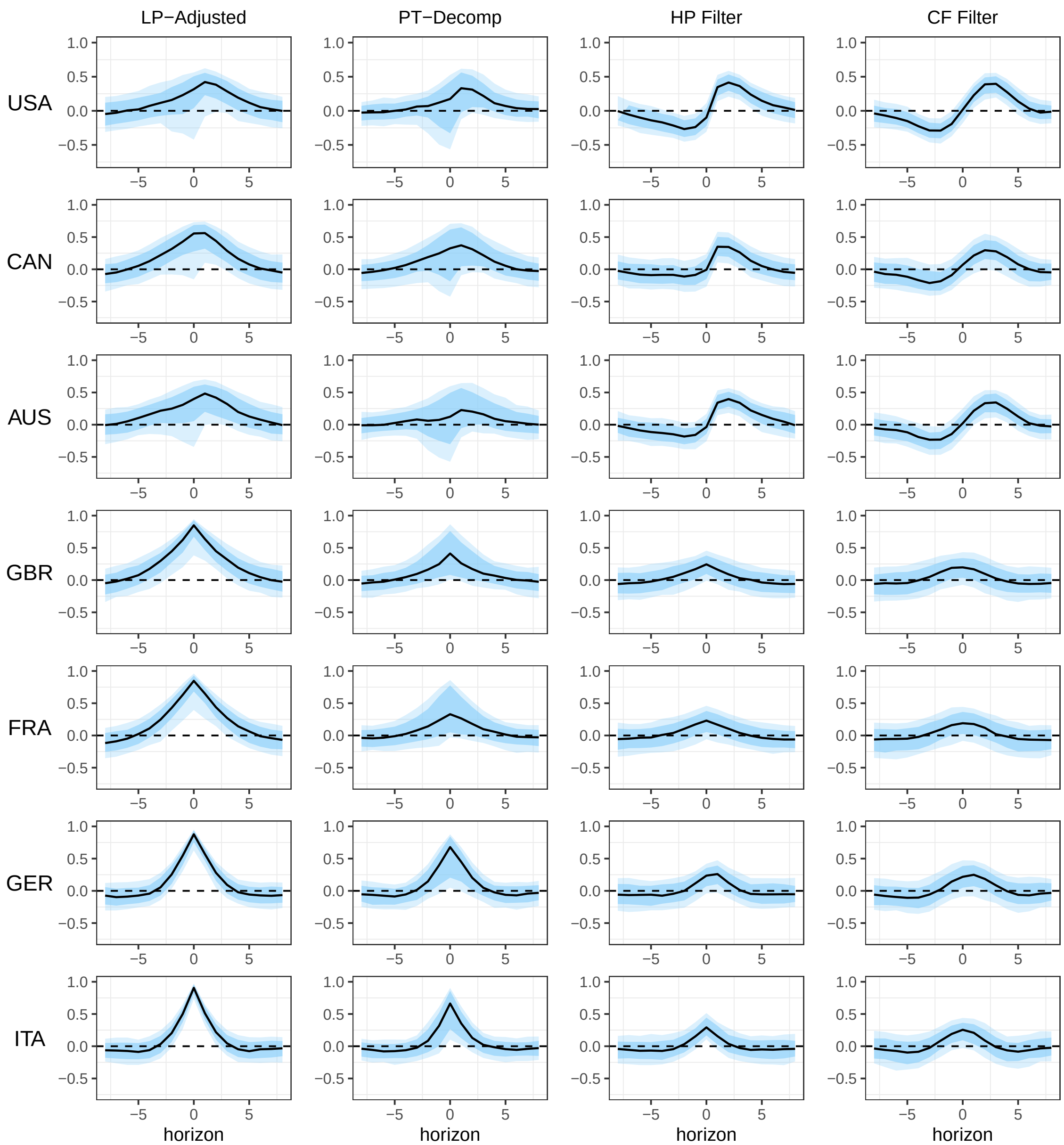
The results for all the countries in our sample are reported in Figure 7. We find two salient patterns. First, the standard filtering techniques (HP, and CF) can be distorted by the large shock and generate negatively correlated estimators with the target in some cases even in short horizons.[22]

[21] here by consistent we mean that these are generated within each iteration by the same shocks.

[22] As a check we repeated the experiment without the large-scale shock in which case the abrupt reversion to negative correlations does not occur. Such estimations and figures are available upon request.

In those cases, the feature of the HP and CF filters of yielding too-sharp and quick reversions of the gap in a single period after the large-scale shock is revealed to be transitory is the reason behind the deterioration of the output estimates. This drawback is corrected by the methods that decompose the permanent and transitory components of the output (LP-Adjusted and PT-Decomp.).

**Figure 7:** Cross-correlation between the output gap estimates and their simulated target




Note: median (black), 68% coverage, and 90% coverage percentiles of the cross-correlations between the output gap estimate of each method and the simulated output gap of the economic model.

Now, it should be mentioned that the reversion does not occur in all cases, and in fact, it will also not appear if we instead perform an experiment without large-scale shocks. However, even if this issue does not emerge, we still see that there is a cross-correlation gain when switching from the standard filters (HP, CF) to those using a permanent-transitory identification scheme.

**The added value of adjusting the estimator.** A second salient pattern we obtain is that when we use our proposal -i.e., we complement the identification in the BSVAR model with the Lenza and Primiceri (2022) adjustment- there is both a decrease in the implied volatility of the estimator (something we verified in prior sections), but we also get a substantial cross-correlation gain, and in the best cases, the peak correlation can jump up to beyond 0.8 from values surrounding 0.5 at best (depending on the alternative method considered). To see this, we can notice how the correlation in the first column is higher than in the other ones for all rows.

In summary, the results of this experiment, favor both the identification method for computing the output gap, but also the addition of the adjustment via the scaling factor as in our proposal, which ultimately will outperform the other alternatives for all countries. Thus, in this case, it is worthwhile to use a relatively more complex but structural estimation method, however, it is even better to complement it with the adjustment for the large-scale shock.

## 5 Concluding remarks

This study examined whether potential output models should be adjusted to account for rare, large-magnitude shocks, such as those experienced during the COVID-19 lockdown in 2020. It aimed to include a complete set of observations in the model while preventing observations of unprecedented magnitudes (that do not resemble the sample data-generating process) from affecting the quality of the resulting econometric modelling framework.

To address this question, we considered a baseline model incorporating ample information sources into a structural framework that allows for the application of an identification strategy that exploits the relationship between consumption and output to recover the permanent and transitory components of GDP as in Uhlig (2003, 2004). Based on this setup, we adjusted the model with a scaling factor of the residuals around the COVID-19 pandemic outbreak along the lines of Lenza and Primiceri (2022).

Our results indicate that only one structural error is enough to account for most of the long-run behaviour of GDP (and potential output) and that the remaining shocks majorly explain transitory fluctuations (i.e. the gap). At the same time, simulation exercises show that the adjusted model outperforms both simple filtering alternatives and similarly complex models that abstract from adjusting the large shock periods or that do so in alternative setups that do not explicitly account for outlying observations at the specific dates of the high-magnitude episodes (e.g. models with stochastic volatility). Concurrently, our setup prevents quick output gap reversals after downturns or drastic changes in the potential output after high-magnitude transitory observations. In that sense, while our setup aligns with the findings of recent studies on the scarring effects of economic downturns (e.g., Cerra, Fatás, and Saxena, 2023; Aikman, Drehmann, Juselius, and Xing, 2022), it still prevents the unprecedented-magnitude observations from affecting the resulting model substantially.

It is relevant to mention that we can make a good approximation of the potential GDP (and gap) by trading off the possibility of disentangling output dynamics into separate structural drivers (e.g. monetary, financial, global, supply, and demand, etc.) Not being able to carry out such a type of exercise is the cost of accessing our identification strategy, which is strictly concerned with an endogenous determination of the horizon profile of the structural shocks. In that spirit, a separation of the output dynamics into fundamental drivers is left for future research, where we can draw lessons from the results of this study that allow mitigating the approximation costs of *ad-hoc* changes in the term horizon of the shocks, common to other identification setups.

## References


Aikman, D., M. Drehmann, M. Juselius, and X. Xing (2022, October). The Scarring Effects of Deep Contractions. BIS Working Papers 1043, Bank for International Settlements.

Álvarez, L. J. and A. Gómez-Loscos (2018). A menu on output gap estimation methods. *Journal of Policy Modeling 40*(4), 827–850.

An, S. and F. Schorfheide (2007). Bayesian Analysis of DSGE Models. *Econometric Reviews 26*(2-4), 113–172.

Angeletos, G.-M., F. Collard, and H. Dellas (2020). Business-cycle anatomy. *American Economic Review 110*(10), 3030–70.

Barsky, R. B. and E. R. Sims (2011). News shocks and business cycles. *Journal of monetary Economics 58*(3), 273–289.

Benati, L. (2008). Investigating Inflation Persistence Across Monetary Regimes. *The Quarterly Journal of Economics 123*(3), 1005–1060.

Benati, L. and T. A. Lubik (2021). Searching for hysteresis. Technical report, Federal Reserve Bank of Richmond.

Berger, T., J. Morley, and B. Wong (2023). Nowcasting the output gap. *Journal of Econometrics 232*(1), 18–34.

Berger, T. and C. Ochsner (2022). Robust real-time estimates of the german output gap based on a multivariate trend-cycle decomposition. Technical report, Deutsche Bundesbank Discussion Paper No. 35/2022, available at SSRN working paper series.

Blanchard, O. J. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbances. *The American Economic Review 79*(4), 655–673.

Blinder, A. S. and J. B. Rudd (2013, June). *The Supply-Shock Explanation of the Great Stagflation Revisited*, pp. 119–175. University of Chicago Press.

Brignone, D. and M. Mazzali (2022). Evidence on the confounding nature of the main business cycle driver. Technical report, Available at SSRN working paper series.

Canova, F. (2020, June). FAQ: How do I measure the Output gap? CEPR Discussion Papers 14943, C.E.P.R. Discussion Papers.

Carriero, A., T. E. Clark, M. Marcellino, and E. Mertens (2022, 06). Addressing COVID-19 Outliers in BVARs with Stochastic Volatility. *The Review of Economics and Statistics 105*, 1–38.

Cerra, V., A. Fatás, and S. C. Saxena (2023, March). Hysteresis and business cycles. *Journal of Economic Literature 61*(1), 181–225.

Chen, M. J. and L. Gornicka (2020, February). Measuring Output Gap: Is It Worth Your Time? IMF Working Papers 2020/024, International Monetary Fund.

Cochrane, J. H. (1990). Univariate vs. multivariate forecasts of gnp growth and stock returns: Evidence and implications for the persistence of shocks, detrending methods.

Cochrane, J. H. (1994). Permanent and transitory components of gnp and stock prices. *The Quarterly Journal of Economics 109*(1), 241–265.

Dieppe, A., N. Francis, and G. Kindberg-Hanlon (2021). The identification of dominant macroeconomic drivers: coping with confounding shocks. Technical report, ECB Working Paper No. 2534.

Fontanari, C., A. Palumbo, and C. Salvatori (2020). Potential output in theory and practice: a revision and update of okun's original method. *Structural Change and Economic Dynamics 54*, 247–266.

Forbes, K., I. Hjortsoe, and T. Nenova (2018). The shocks matter: Improving our estimates of exchange rate pass-through. *Journal of International Economics 114*(C), 255–275.

Fornaro, L. and M. Wolf (2020). Covid-19 coronavirus and macroeconomic policy. Technical report, CEPR Discussion Paper No. DP14529.

Furlanetto, F., A. Lepetit, Robstad, J. F. Rubio-Ramírez, and P.-b. Ulvedal (2021, September). Estimating Hysteresis Effects. CEPR Discussion Papers 16558, C.E.P.R. Discussion Papers.

Giannone, D., M. Lenza, and G. E. Primiceri (2015). Prior selection for vector autoregressions. *Review of Economics and Statistics 97*(2), 436–451.

Guerrieri, V., G. Lorenzoni, L. Straub, and I. Werning (2022). Macroeconomic implications of covid-19: Can negative supply shocks cause demand shortages? *American Economic Review 112*(5), 1437–74.

Guisinger, A. Y., M. T. Owyang, and H. Shell (2018). Comparing Measures of Potential Output. *Review 100*(4), 297–316.

Hartwig, B. (2022). Bayesian vars and prior calibration in times of covid-19. *Studies in Nonlinear Dynamics  Econometrics*.

Hazell, J., J. Herreño, E. Nakamura, and J. Steinsson (2022, 02). The Slope of the Phillips Curve: Evidence from U.S. States*. *The Quarterly Journal of Economics 137*(3), 1299–1344.

Jordà, O., S. R. Singh, and A. M. Taylor (2020, January). The Long-Run Effects of Monetary Policy. NBER Working Papers 26666, National Bureau of Economic Research, Inc.

Keating, J. W. and V. J. Valcarcel (2015). The time-varying effects of permanent and transitory shocks to real output. *Macroeconomic Dynamics 19*(3), 477–507.

King, R. G., C. I. Plosser, J. H. Stock, and M. W. Watson (1987). Stochastic trends and economic fluctuations.

Lenza, M. and G. E. Primiceri (2022). How to estimate a vector autoregression after march 2020. *Journal of Applied Econometrics 37*(4), 688–699.

McLeay, M. and S. Tenreyro (2020). Optimal Inflation and the Identification of the Phillips Curve. *NBER Macroeconomics Annual 34*(1), 199–255.

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*(C).

Ng, S. (2021). Modeling macroeconomic variations after covid-19. Technical report, National Bureau of Economic Research.

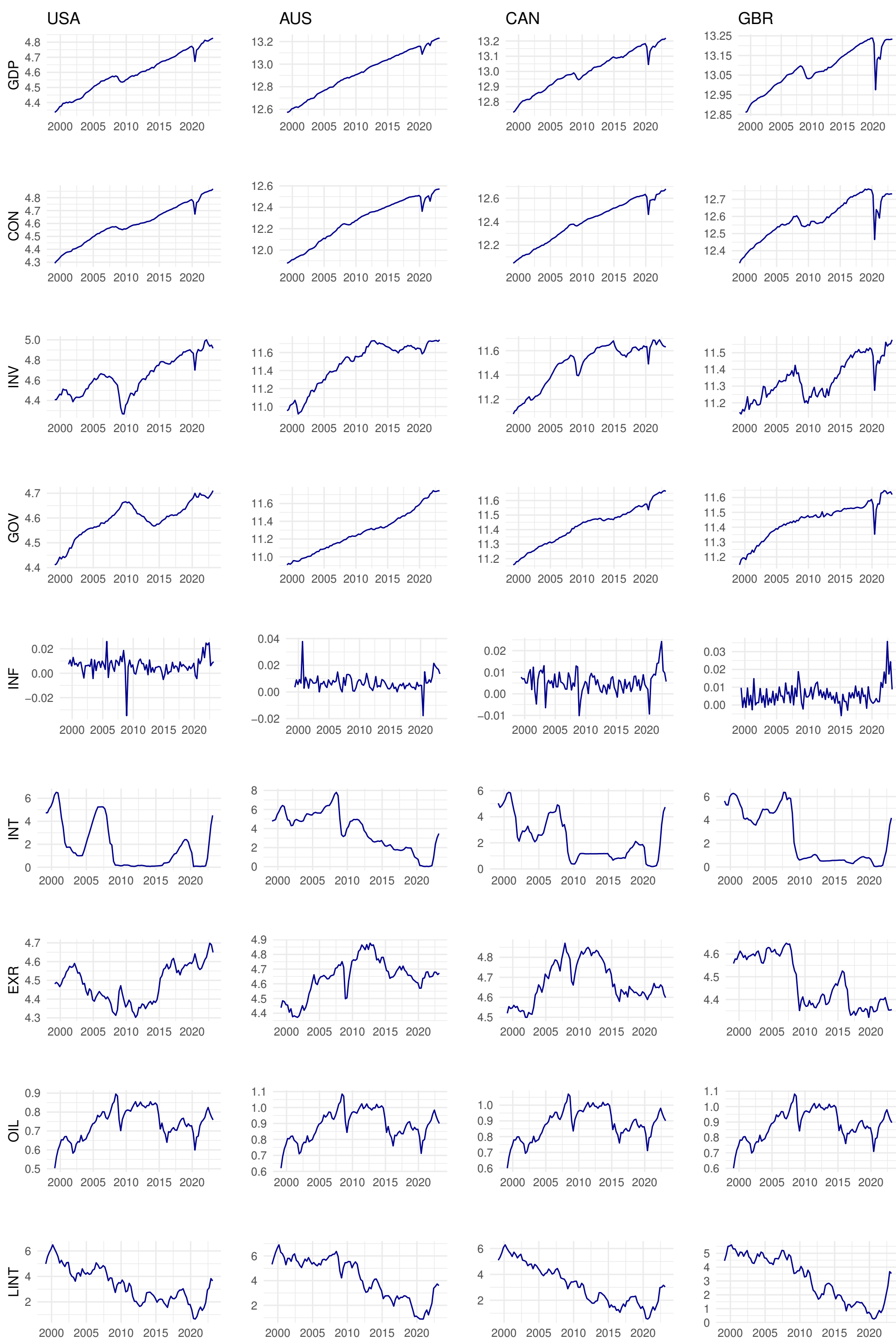
Uhlig, H. (2003). What moves real gnp? Unpublished manuscript.

Uhlig, H. (2004). What moves gnp? In *Econometric Society 2004 North American Winter Meetings*, Number 636 in 1. Econometric Society.

# A Baseline model: Results

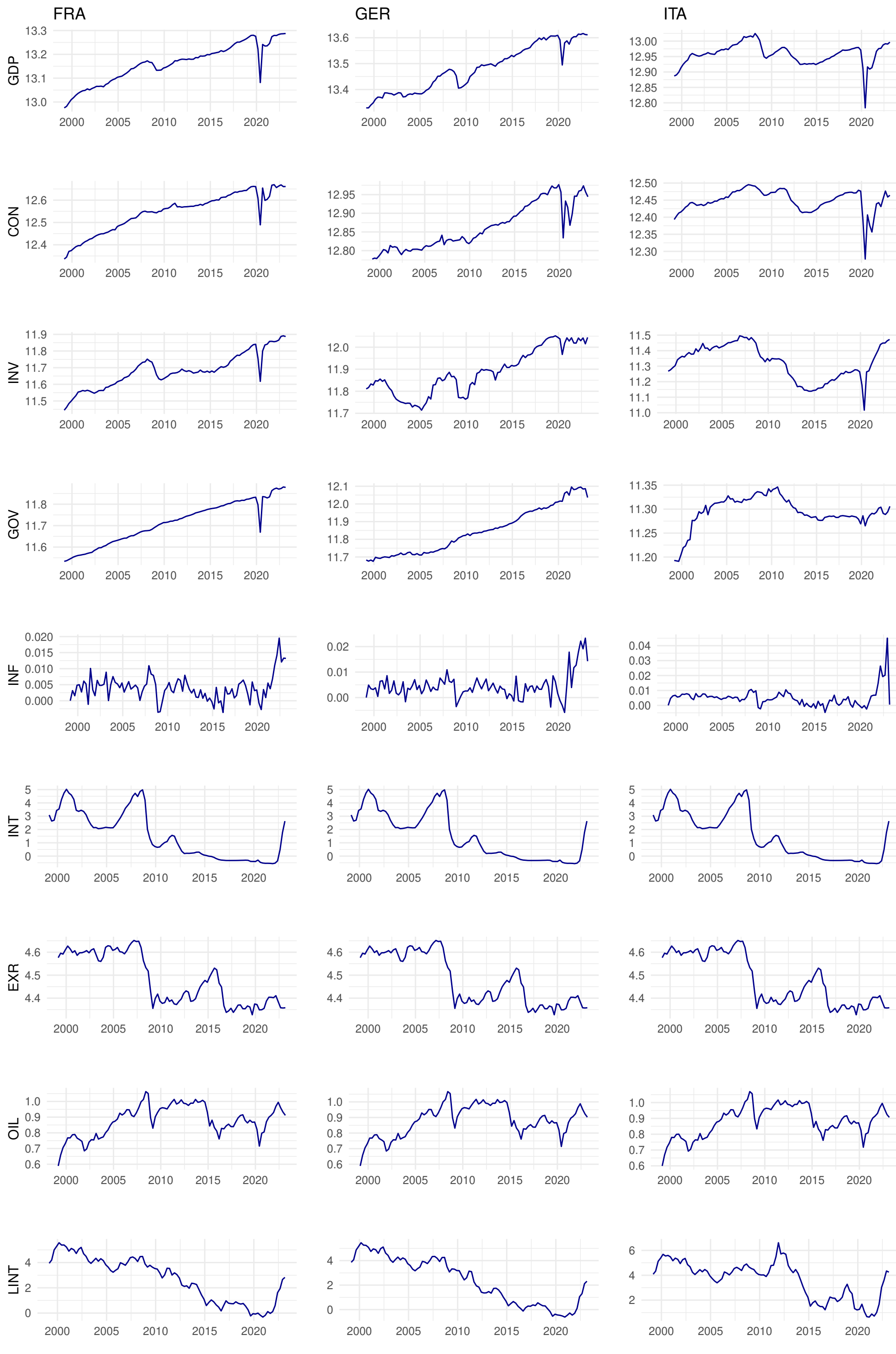
## A.1 Data

**Figure 8:** Data for USA, AUS, CAN and GBR




Notes: All data, except interest rates and inflation, are in logs.

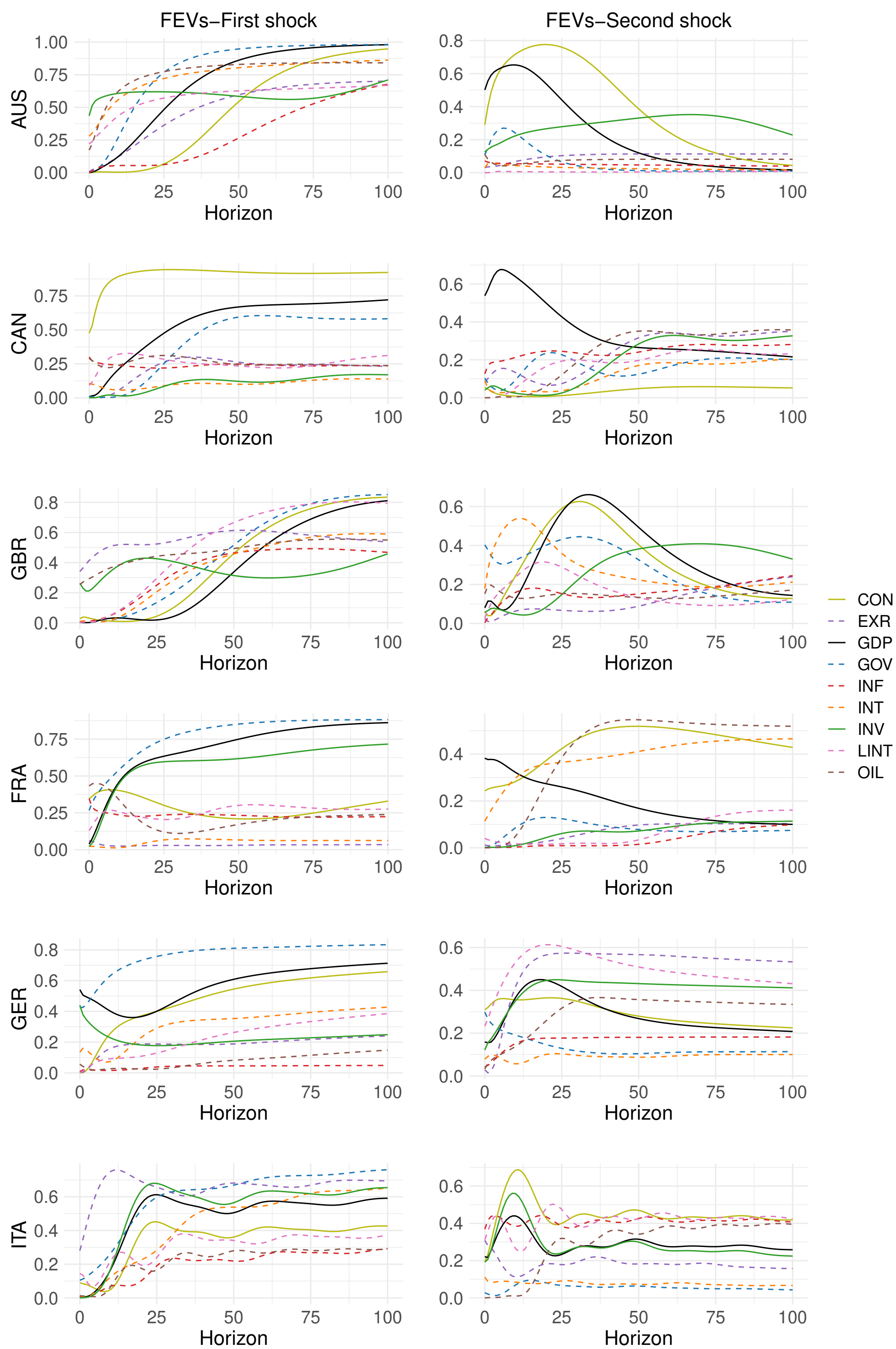
**Figure 9:** Data for FRA, GER and ITA




Notes: All data, except interest rates and inflation, are in logs.

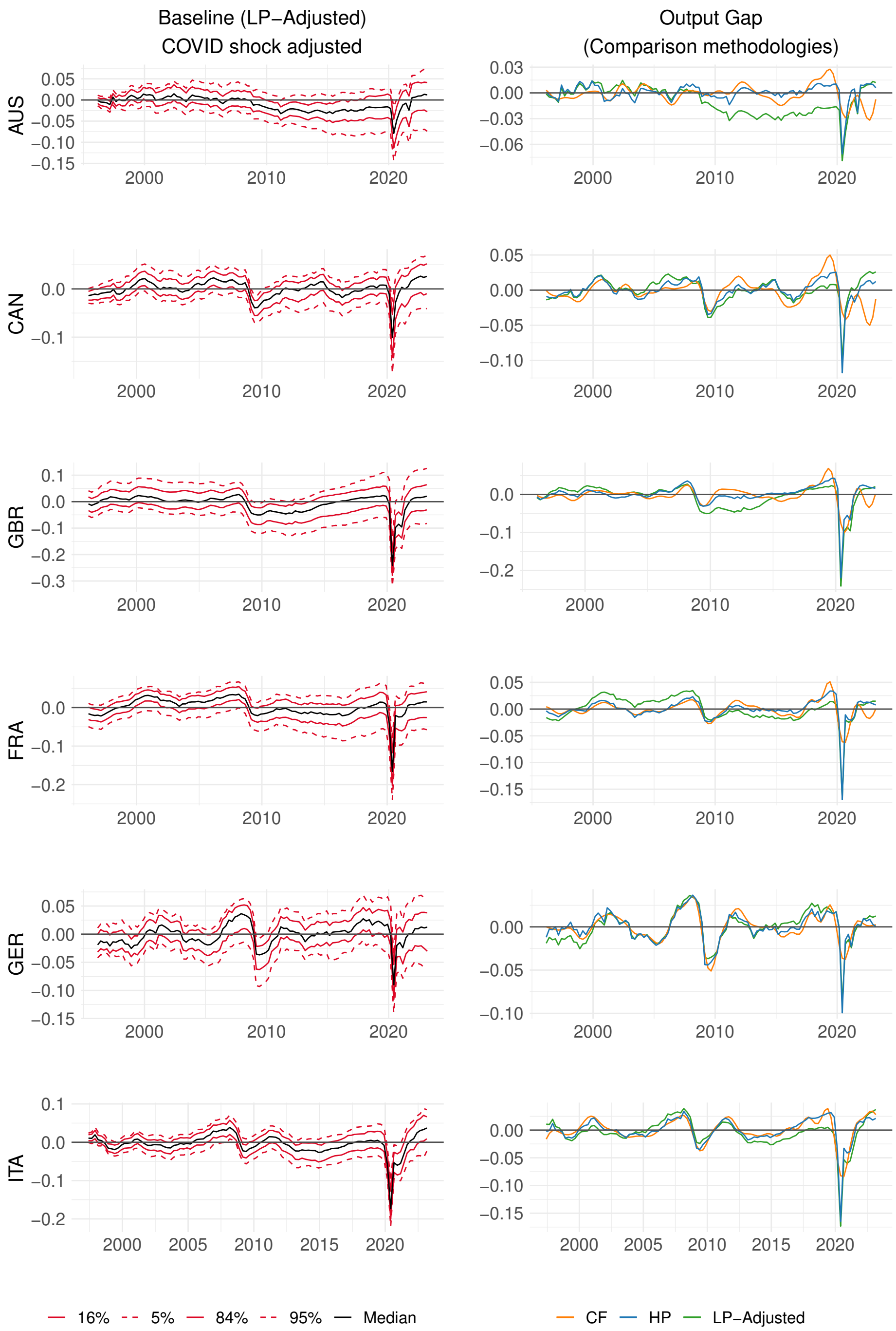
## A.2 Contribution of FEV explanation over each variable for the rest of six economies (The first two largest shocks that maximize FEV of GDP for each country)

**Figure 10:** Contribution of FEV explanation over each variable for the rest of six economies



## A.3 Baseline results: Output gap for the other six economies

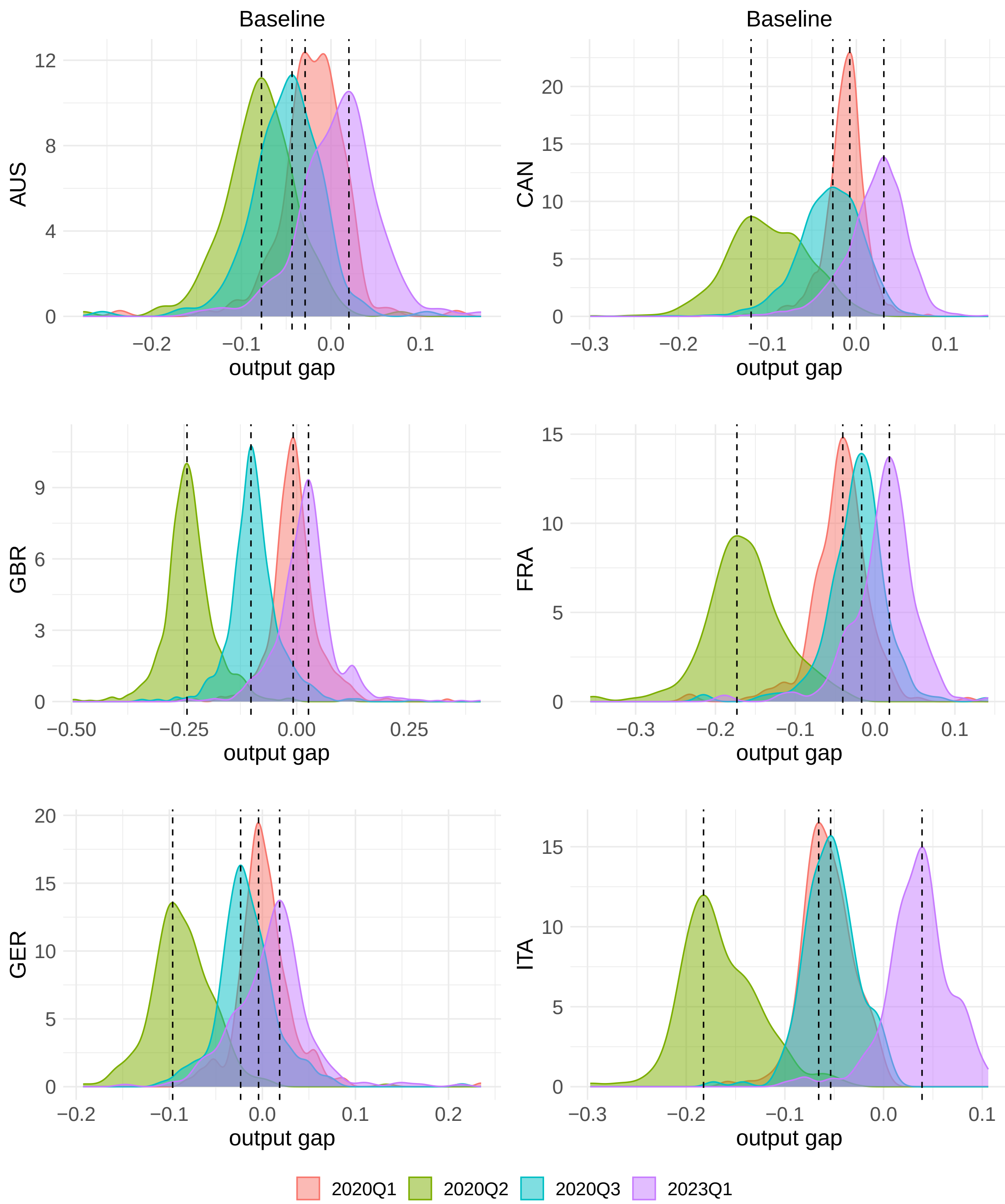
**Figure 11:** Baseline results: Output gap for the other six economies




Notes: Left panel: The Figures show the results of the baseline BSVAR model with Permanent-Transitory decomposition and Covid-19 adjustment (LP-Adjusted). The solid black line represents the median estimates. The solid and dotted red lines represent the percentiles of 5%, 16%, 84% and 95%, respectively. Right panel: The Figures show the results of the baseline model (LP-Adjusted) in comparison to the Hodrick-Prescott (HP) and Christiano-Fitzgerald (CF) univariate filters.

## A.4 Distribution of the output gap estimation during COVID-19 shock and 2023Q1 for the rest six economies

**Figure 12:** Distribution of the output gap estimation for selected quarters



## A.5 Forecast Error Variance decomposition (FEVs)

**Figure 13:** Forecast Error Variance decomposition (FEVs) by variable(USA)

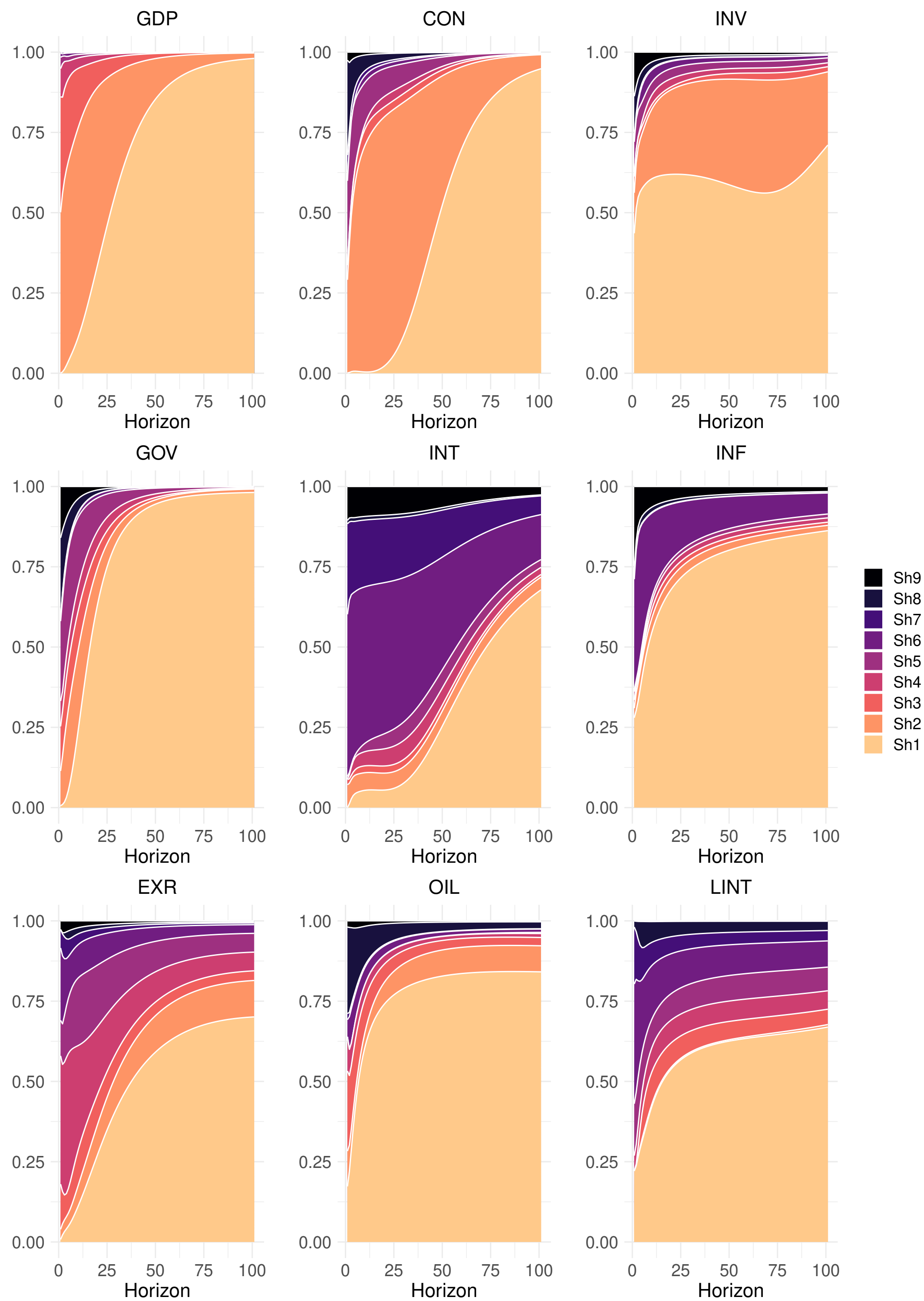


**Figure 14:** Forecast Error Variance decomposition (FEVs) by variable(AUS)

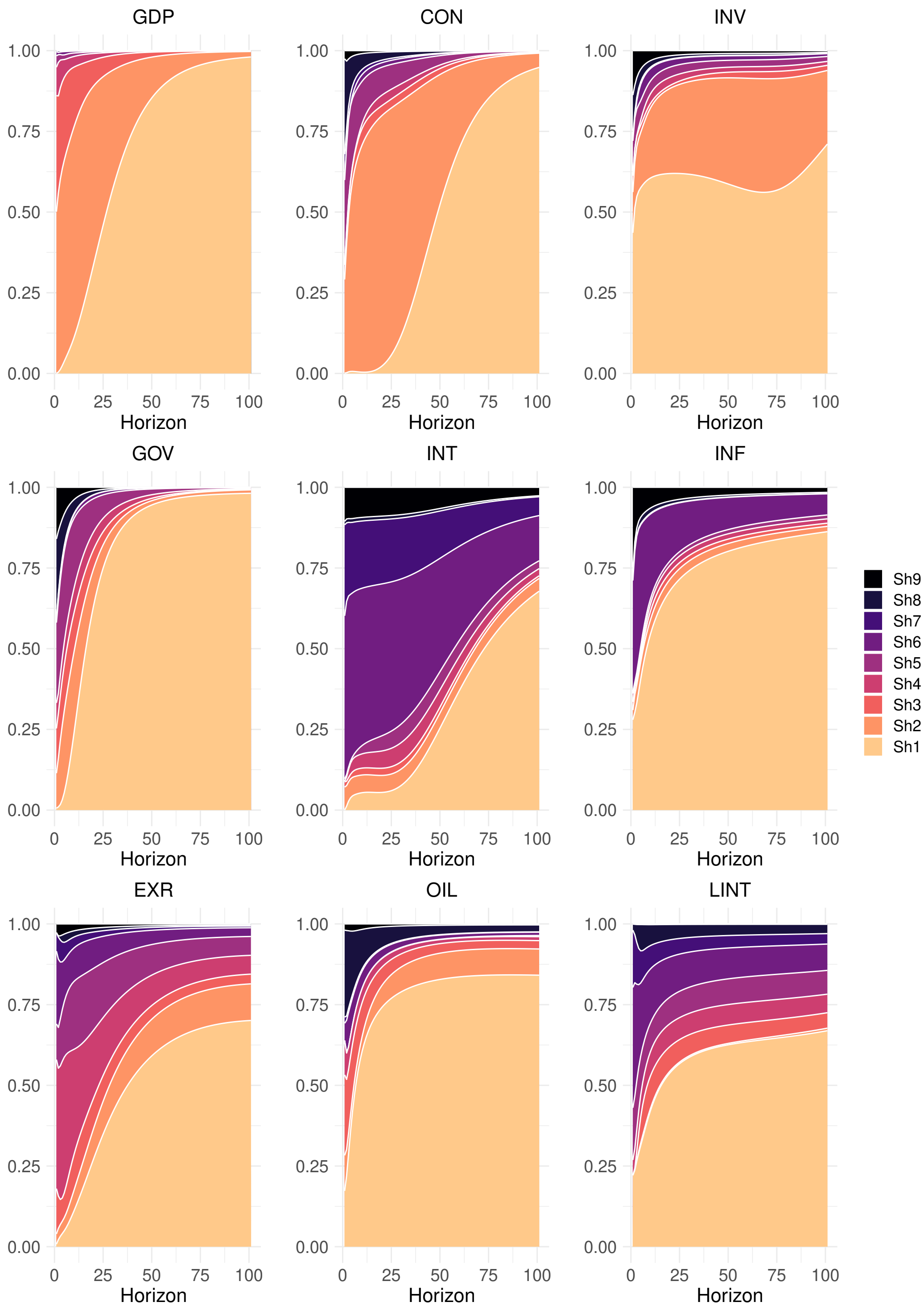


**Figure 15:** Forecast Error Variance decomposition (FEVs) by variable(CAN)

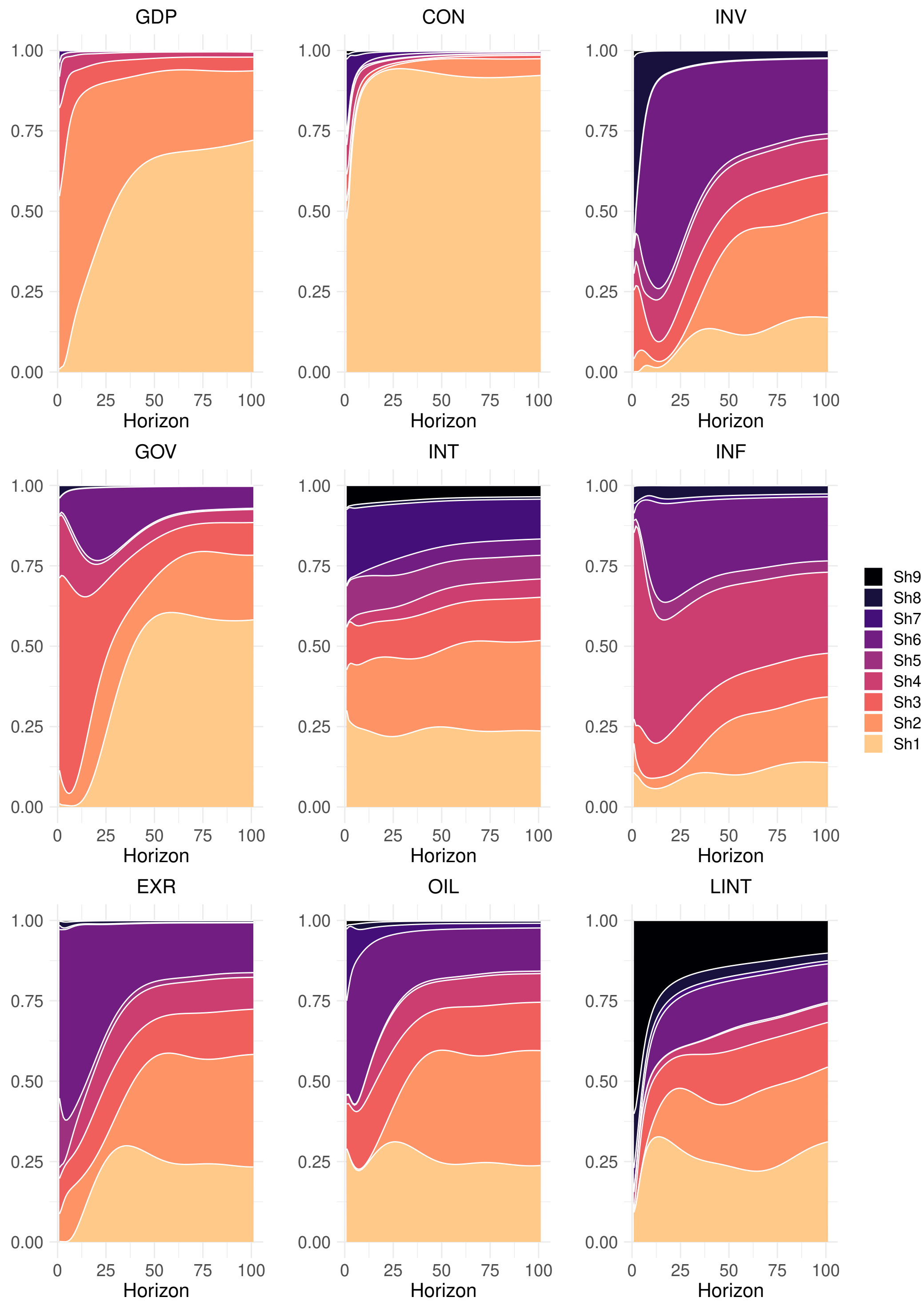


**Figure 16:** Forecast Error Variance decomposition (FEVs) by variable(GBR)

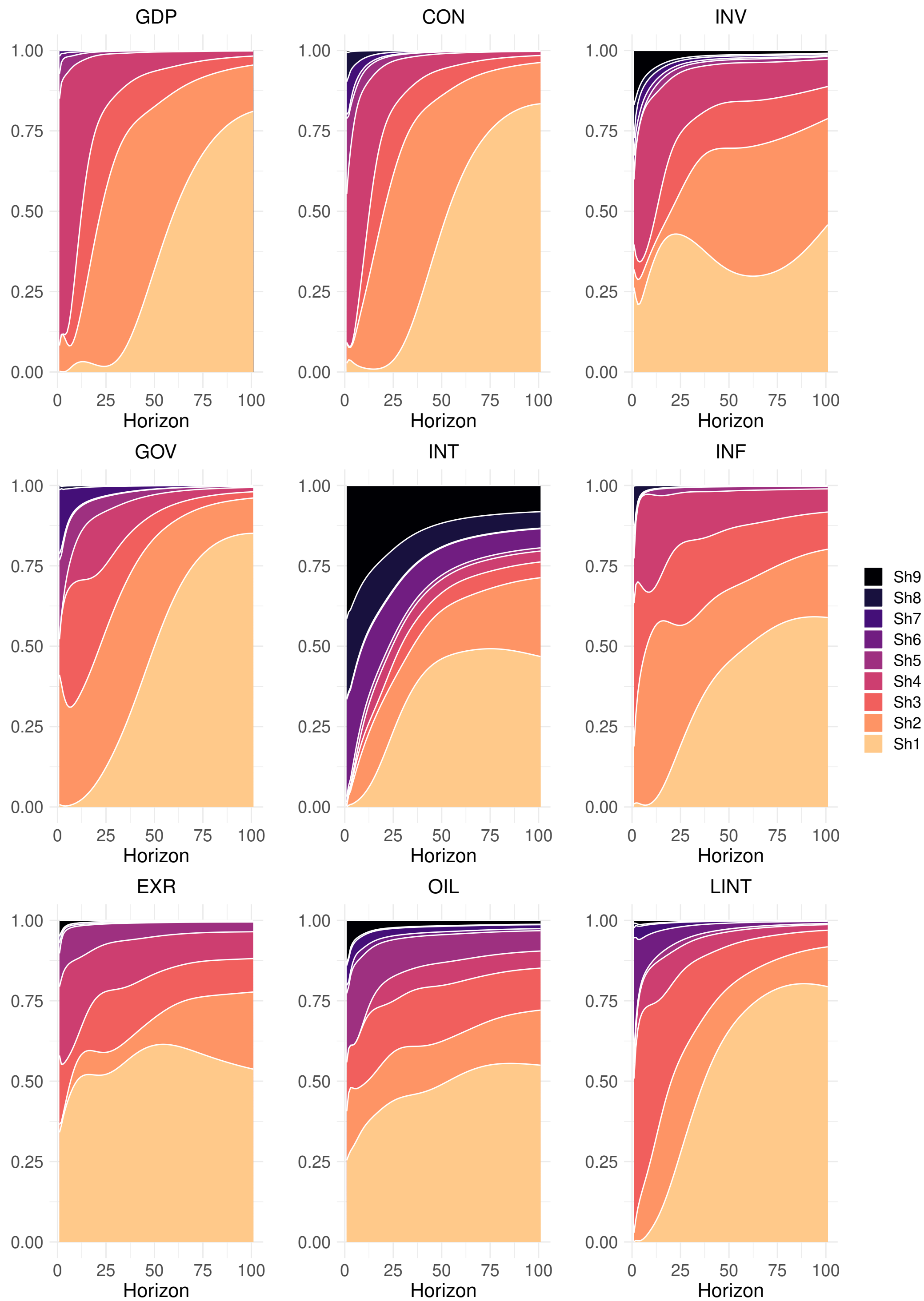


**Figure 17:** Forecast Error Variance decomposition (FEVs) by variable(FRA)

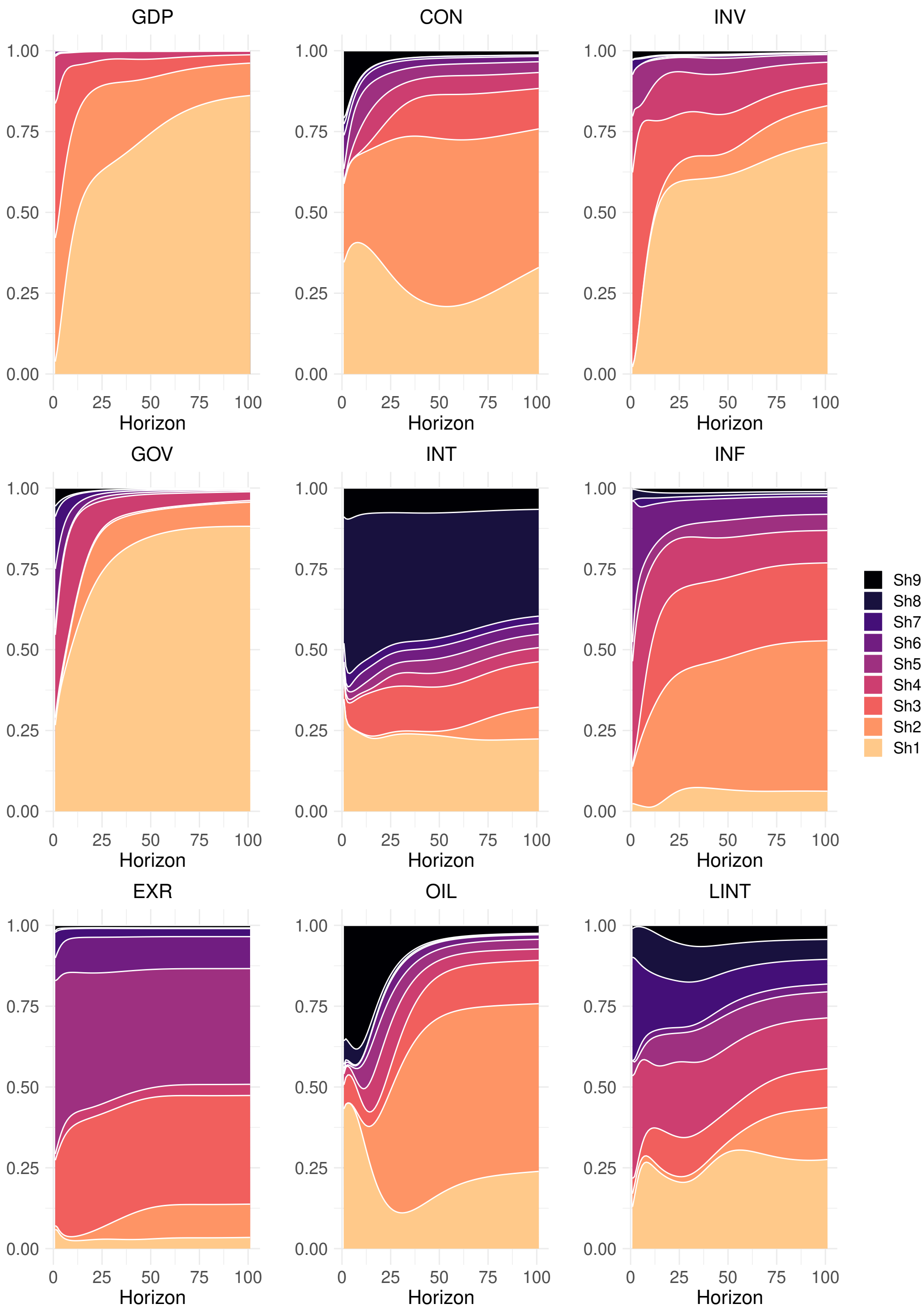


**Figure 18:** Forecast Error Variance decomposition (FEVs) by variable(GER)

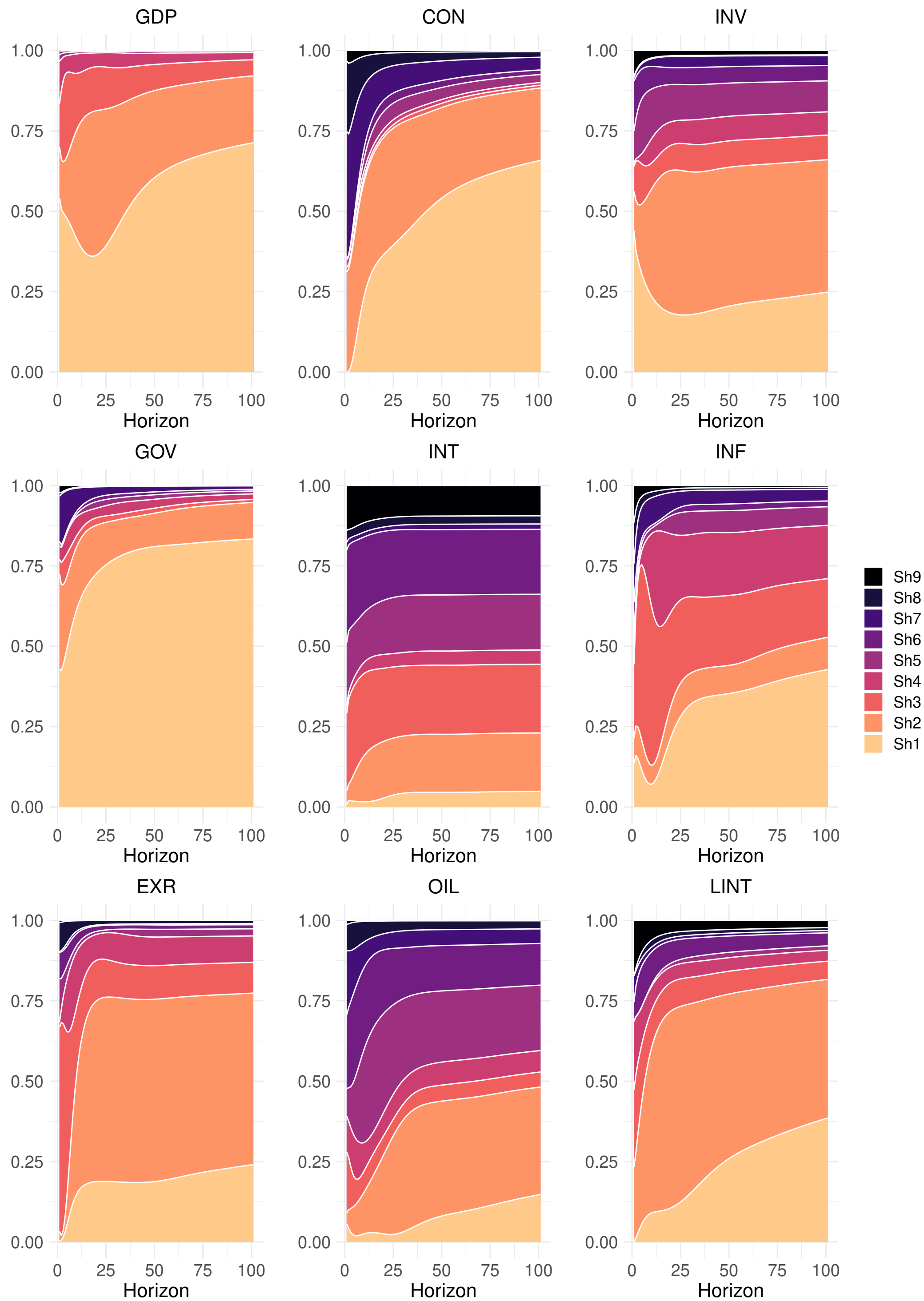
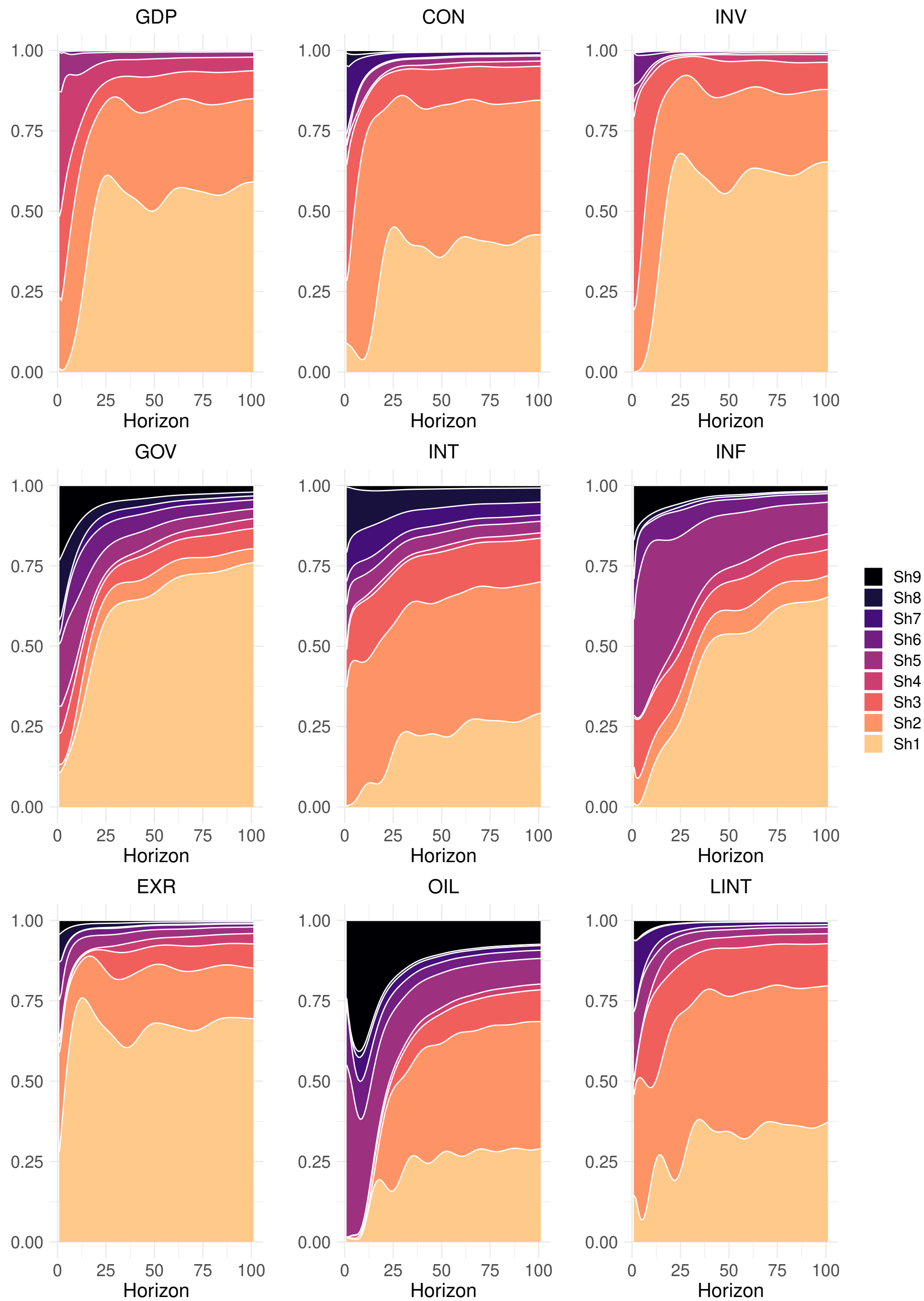
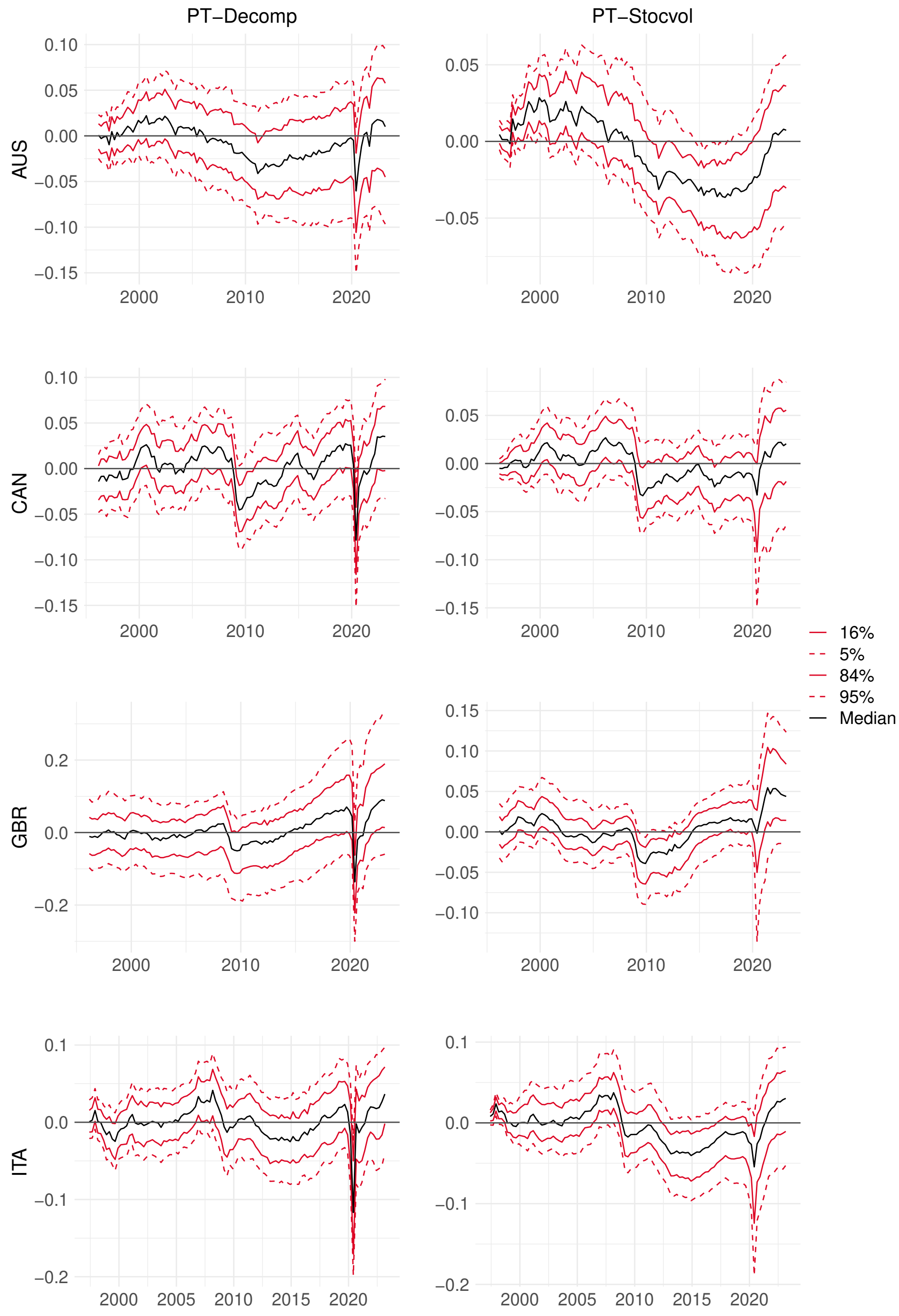


**Figure 19:** Forecast Error Variance decomposition (FEVs) by variable(ITA)



# B Other models: No COVID-19 adjustent and Stochastic Volatility.

**Figure 20:** Output gap for selected countries (PT-Decomp and PT-Stocvol)




Notes: Left panel: The Figures show the results of the BSVAR model with Permanent-Transitory decomposition but without Covid-19 adjustment (PT-Decomp). Right panel: The Figures show the results of the BSVAR model with Permanent-Transitory decomposition and Stochastic volatility but without Covid-19 adjustment (PT-Stocvol). The solid black line represents the median estimates. The solid and dotted red lines represent the percentiles of 5%, 16%, 84% and 95%, respectively.