DEUTSCHE
BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 35/2022

Robust real-time estimates of the German output gap based on a multivariate trend-cycle decomposition

Tino Berger
(University of Göttingen)

Christian Ochsner
(German Council of Economic Experts)

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978–3–95729–912–3
ISSN 2749–2958

Non-technical summary

Research Question

The German economic situation is of key importance for economic flourishing and stability in the European Union. Central banks and other policy institutions frequently use the output gap – the deviation of actual from potential real gross domestic product (GDP) – as a measure of the business cycle. Being intrinsically unobserved, the output gap has to be estimated. Economic disruptions require fast economic policy responses. Waiting until GDP data has been released one month after the end of a quarter can be economically costly. Therefore, a timely and reliable estimate of the output gap is needed.

Contribution

Our contribution is threefold. First, we provide a state-of-the-art estimate of the German output gap. We reconcile traditional production-function identification of the German output gap with the Beveridge-Nelson decomposition as a tool to disentangle trends and cycles. This decomposition is particularly appealing, as it provides a straightforward characterization of the cyclical component as well as the possibility to interpret the output gap in terms of the aggregates that inform cycle estimation. Second, by tracking the German output gap up to three months prior to the GDP data release, we provide timely information about the stance of the largest economy in the euro area. Finally, we contribute to the topic of the reliability of output gap estimates in data-rich environments: We investigate to what extent our estimate of the output gap is revised ex post.

Results

Our estimate of the output gap lines up reasonably well with established filter and production-function based measures. Furthermore, we find that fluctuations in the German output gap are mainly transmitted by business and consumer expectations. Moreover, in line with the related literature we show that external relations play a key role for the German output gap, whereas labour market aggregates are informative in times of large deviations from potential output. We find that the output gap estimated after the first month of a quarter and, thus, without knowing the current quarter's GDP is very close to its final estimate.

Nichttechnische Zusammenfassung

Fragestellung

Die deutsche Wirtschaftslage ist von zentraler Bedeutung für die Stabilität der Wirtschaft in der Europäischen Union. Zentralbanken und andere Institutionen verwenden häufig die Produktionslücke, d. h. die Abweichung des tatsächlichen vom potenziellen realen Bruttoinlandsprodukt (BIP), als Maß für den Konjunkturzyklus. Da die Produktionslücke nicht beobachtbar ist, muss sie aus BIP-Daten geschätzt werden. In Zeiten wirtschaftlicher Krisen ist eine schnelle wirtschaftspolitische Reaktion gefragt. Das Warten auf die Veröffentlichung der BIP-Daten einen Monat nach Quartalsende kann kostspielig sein. Daher ist eine rechtzeitige und zuverlässige Schätzung der Produktionslücke erforderlich.

Beitrag

Unsere Arbeit enthält drei Beiträge. Erstens liefern wir eine zeitgemäße Schätzung der deutschen Produktionslücke. Wir bringen die traditionelle Produktionsfunktionsidentifikation der deutschen Produktionslücke mit der Beveridge-Nelson-Zerlegung als Instrument zur Unterscheidung von Trends und Zyklen in Einklang. Diese Zerlegung ist besonders attraktiv, da sie eine einfache Charakterisierung der zyklischen Komponente sowie die Möglichkeit bietet, die Produktionslücke im Hinblick auf die Aggregate zu interpretieren, die für die Zyklusabschätzung maßgeblich sind. Zweitens liefern wir durch die Verfolgung der deutschen Produktionslücke bis zu drei Monate vor der Veröffentlichung der BIP-Daten zeitnahe Informationen über den Zustand der größten Volkswirtschaft im Euroraum. Schließlich leisten wir einen Beitrag zum Thema der Zuverlässigkeit der Schätzung von Produktionslücken in einem datenreichen Umfeld: Wir untersuchen, in welchem Umfang unsere Schätzung der Produktionslücke im Nachhinein revidiert werden muss.

Ergebnisse

Unsere geschätzte Produktionslücke stimmt gut mit etablierten Filter- und Produktionsfunktionsverfahren überein. Darüber hinaus stellen wir fest, dass die Schwankungen der deutschen Produktionslücke hauptsächlich durch die Erwartungen der Unternehmen und Verbraucher beeinflusst werden. Außerdem zeigen wir im Einklang mit der einschlägigen Literatur, dass die Außenbeziehungen eine Schlüsselrolle für die deutsche Produktionslücke spielen, während Arbeitsmarktaggregate in Zeiten großer Abweichungen vom Produktionspotenzial informativ sind. Schließlich stellen wir fest, dass die im ersten Monat eines Quartals geschätzte Produktionslücke, obwohl das BIP des laufenden Quartals dann noch unbekannt ist, kaum von der endgültigen Schätzung abweicht.

Bundesbank Discussion Paper No 35/2022

Robust Real-Time Estimates of the German Output Gap based on a Multivariate Trend-Cycle Decomposition*

Tino Berger[†1] and Christian Ochsner[‡2]

[1]University of Göttingen, Chair of Empirical International Economics
[2]German Council of Economic Experts

August 22, 2022

Abstract

The German economy is an important economic driver in the Euro-area in terms of gross domestic product, labour force and international integration. We provide a state of the art estimate of the German output gap between 1995 and 2022 and present a nowcasting scheme that accurately predicts the German output gap up to three months prior to a gross domestic product data release. To this end, we elicit a mixed-frequency vector-autoregressive model in the spirit of Berger, Morley, and Wong (forthcoming) who propose to use monthly information to form an expectation about the current-quarter output gap. The mean absolute error of our nowcast compared to the final estimate is very small (0.28 percentage points) after only one month of observed data. Moreover, we show that business and consumer expectations, international trade and labour market aggregates consistently explain large shares of variation in the German output gap. Finally, our procedure is very reliable, as it implies an output gap that is hardly revised ex post. This is particularly important for policymakers.

Keywords: output gap, Germany, nowcast, mixed frequency, vector-autoregression
JEL classification: E32, E37, C53

*We thank David Boll, Lasse Trienens and Lieve Vanhooren for research assistance. We are grateful to Elmar Mertens, Julia Richter, Maik Wolters and Malte Knüppel for great discussions. Moreover, we thank workshop participants at University of Göttingen, Deutsche Bundesbank, German Council of Economic Experts, the 15^{th} Ruhr Graduate School Doctoral Conference in Economics, the 30^{th} Annual Symposium of the Society for Non-Linear Dynamics and Econometrics and at the 8^{th} Annual Meeting of the International Association for Applied Econometrics for helpful comments. We acknowledge financial support from Deutsche Bundesbank. The usual disclaimers apply. In particular, the views expressed in the article at hand represent the authors' personal opinions and do not necessarily reflect the views of the German Council of Economic Experts, Deutsche Bundesbank or the Eurosystem. Note that the draft was written before Christian Ochsner joined the the scientific staff of the German Council of Economic Experts.

†Email: tino.berger@wiwi.uni-goettingen.de;
‡Email: christian.ochsner@uni-goettingen.de; corresponding author.

1 Introduction

The German economy is the fifth largest economy in the world and with a GDP weight of about 28% in 2021 the single largest economy in the euro area. Therefore, the German economic situation is of key importance for economic flourishing and stability in the European Union. For the conduct of monetary policy in the euro area, the European Central Bank relies on measurements of a euro area business cycle. The latter is driven to a large extend by the German business cycle, the focus of the paper at hand.

Central banks and other policy institutions frequently use the output gap, i.e. the deviation of actual from potential real gross domestic product (GDP) as a measure of the business cycle. Being intrinsically unobserved, the output gap has to be estimated. Numerous models and filtering approaches have been proposed to this end (most prominently, Hodrick and Prescott (1997) and more recently Hamilton (2018)). Moreover, estimating the output gap with the help of production-functions is wide-spread among policymakers across the globe. This approach is appealing, as it provides information on the structural determinants of the output gap. However, the vast majority of available procedures only yield retrospective insights into the output gap due to a significant delay in GDP data availability. As a recent exception, Berger, Morley, and Wong (forthcoming) (BMW (forthcoming), henceforth) propose to nowcast the output gap using a multivariate Beveridge-Nelson decomposition based on a mixed-frequency Bayesian vector-autoregressive model (VAR). They show that a model comprising economic aggregates available at monthly frequency implies a reasonable output gap for the U.S. economy well in advance current quarter GDP data is released.

We estimate the German output gap by means of the output gap using the approach of BMW (forthcoming). Going beyond their work, we present a reliable procedure to select the most relevant variables for estimating the German output gap in a multivariate model. Using an informational decomposition allows us to quantify the relative importance of each variable. Moreover, we analyze the contribution of each variable in each month within a quarter. The accuracy of the models' nowcasts are evaluated by comparing the nowcast after each month of a given quarter to the final estimate obtained using the full information set. In addition, we extensively discuss the role of data, parameter and specification revisions. In particular, we present a detailed analysis in the spirit of Orphanides and van Norden (2002) to demonstrate the reliability of our approach.

The contribution is threefold. First, we provide a state-of-the-art estimate of the German output gap. German output gap fluctuations are monitored by the German council of economic experts ('Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung') at the behest of the federal government. As part of its mandate, the council presents a comprehensive summary of the main economic developments. The biannual expert reports make an important contribution to understanding the German economy. To approximate the output gap, the council estimates a production function. However, academic accounts of the German output gap are sparse. We reconcile traditional production-function identification of the German output gap with the Beveridge-Nelson decomposition as a tool to disentangle trends and cycles. The Beveridge-Nelson decomposition is particularly appealing, as it provides a straightforward characterization of the cyclical component as well as the possibility to interpret the output gap in terms of the aggregates that inform cycle estimation. Second, by tracking the German output gap up to three months prior to GDP data release, we provide timely information about the stance of the largest economy in the euro area that are crucial for the conduct of monetary and fiscal policy. The impact of large economic shocks, such as the COVID-19 pandemic or recent oil-market disturbances can be straightforwardly tracked within a month within a given quarter, rather than six weeks after the previous quarter. Finally, we contribute to the literature on output gap estimation in data-rich environments. It is well known that univariate approaches suffer from unreliability in real-time (see Orphanides and van Norden (2002)). On the contrary, recently proposed multivariate models of the output gap (e.g. Jarocinski and Lenza (2018), Morley and Wong (2020) and Barigozzi and Luciani (2021)) take advantage of a large set of information which improves not only economic interpretability, but also real-time forecasting performance. In the vein of this literature, we show that our model is reliable: The nowcasting scheme is very accurate and the output gap estimate is hardly revised ex post.

Our estimate of the output gap lines up reasonably well with established filter and production-function based measures. Furthermore, we find that fluctuations in the German output gap are mainly transmitted by expectations. Moreover, in line with the related literature we show that external relations play a key role for the German output gap (see e.g. Eickmeier (2007)), whereas labour market aggregates are informative in times of large deviations from potential output. Regarding the nowcast accuracy, we find that after the first month of a quarter, the nowcast has a mean absolute error of 0.28 percentage points from the final estimate of the output gap. Even in times of substantial volatility, our model yields robust results under real-world conditions without observing current quarter GDP data. For instance, the output gap in the COVID-19 induced recession in 2020Q2 was nowcasted to be −8.0% after the release of the May 2020 monthly data. It turned out to be −8.8% after the release of the entire quarter data (including revised GDP) at the end of August 2020.

Sections 2 and 3 present our empirical approach and the data, respectively. In Sections 4, 5 and 6, we discuss our estimate and the nowcasting performance in real-time. Section 7 examines robustness of our results to an even larger information set and Section 8 concludes.

2 Methodology

We estimate a model in the spirit of BMW (forthcoming), who propose a mixed-frequency Bayesian vector-autoregression (MF-BVAR) to obtain the output gap c_t as the cyclical component of output from a multivariate Beveridge-Nelson decomposition. Using the Beveridge-Nelson decomposition in a multivariate setting enables us to interpret the resulting estimate in terms of the information that is most relevant for disentangling trend (i.e. potential output) and cycle (i.e. output gap). More precisely, if y_t is a $K \times 1$ vector of macroeconomic observables with $K \times 1$ drift component μ, its Beveridge-Nelson trend τ_t is given as $\lim_{h\to\infty} \mathbb{E}_t\left(y_{t+h} - h\mu\right)$ and the cycle can be obtained as $c_t = y_t - \tau_t$. Assuming stationarity and zero mean, Morley and Wong (2020) show that the cycle c_t is given by

$$c_t = -\boldsymbol{F}(\boldsymbol{I}_K - \boldsymbol{F})^{-1}\boldsymbol{X}_t, \tag{1}$$

where $\boldsymbol{I}_K$ is the identity matrix of rank K and the remaining quantities derive from the vector-autoregressive model $\boldsymbol{X}_t = \boldsymbol{F}\boldsymbol{X}_{t-1} + \boldsymbol{H}\epsilon$ with $\epsilon \sim \mathcal{N}(0, \Sigma)$ (BMW forthcoming). In particular, $\boldsymbol{X}_t$ and ϵ are $K \times 1$ vectors of macroeconomic observables and innovations in t, respectively. $\boldsymbol{F}$ is a companion-form coefficient matrix and $\boldsymbol{H}$ maps the innovations into companion form. We stack all high-frequency (HF) variables above the lower frequency (LW) variables. For instance, assume one low-frequency period can be subdivided into $d = 1, ..., D$ equidistant high-frequency-periods, then we obtain $\boldsymbol{X}_t$ as

$$\boldsymbol{X}_t = \begin{bmatrix} \boldsymbol{x}^{HF}_{t-1+1/D} \\ \vdots \\ \boldsymbol{x}^{HF}_{t-d/D} \\ \boldsymbol{x}^{HF}_{t} \\ \boldsymbol{x}^{LF}_{t} \end{bmatrix},$$

where $\boldsymbol{x}_t$ indicate partitions of $\boldsymbol{X}_t$ for the high-frequency series in $d = 1, ..., D$ in t. To contain parameter proliferation, we follow Morley and Wong (2020) in adopting Bayesian shrinkage. We use a standard Minnesota prior for location and scale parameters, where the Minnesota-prior shrinkage parameter λ is chosen such that the one-step-ahead root mean squared error of output is minimized.[1]

The higher (in our case, monthly) frequency information is exploited in the spirit of Waggoner and Zha (1999) to update the vector-autoregression for the subsequent period. To this end, we note that (by positive-definiteness) the innovation covariance $\boldsymbol{\Sigma}$ obeys a representation $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}'$, where $\boldsymbol{B}$ is the lower-triangular Choleski factor (alternative decompositions might be employed). $\boldsymbol{B}$, by virtue of its triangular structure, is used as the contemporaneous impact multiplier for the new high-frequency information. Put differently, by pre-multiplying the relevant parameters in $\boldsymbol{B}$ to the observed high-frequency shocks, we can track their propagation through the system in time t during D high-frequency periods. More precisely, we observe the upper partitions of $\boldsymbol{\varepsilon}_{T+1}$ and use $\boldsymbol{B}$ to form an expectation about future innovations during the *entire* low-frequency period ahead. Hence, the expected innovations are non-zero *conditionally* on information observed in a given high-frequency interval. By means of the subsequent evaluation of the vector-autoregression, we obtain a forecast of $T + 1$ for the entire system. Taking advantage of this technique, BMW (forthcoming) elicit a nowcast of the output gap by iterating on Eq. (1):

$$c_{t+1} \quad = \quad -\boldsymbol{s}_k\boldsymbol{F}(\boldsymbol{I} - \boldsymbol{F})^{-1}\left[\boldsymbol{F}\boldsymbol{X}_t + \boldsymbol{H}\boldsymbol{\varepsilon}_{t+1}\right],$$

where $\boldsymbol{s}$ is a selection column vector. Finally, we can trace out variation in the cycle $c_{ij,t}$ of the higher-frequency series i to surprises in variable j in time t by means of

$$c_{ij,t} \quad = \quad -\sum_{d=1}^{D}\sum_{l=0}^{t-1} \boldsymbol{s}_k\boldsymbol{F}^{l+1}(\boldsymbol{I} - \boldsymbol{F})^{-1}\boldsymbol{H}\boldsymbol{s}'_j\boldsymbol{s}_j\boldsymbol{\varepsilon}_{t-1}. \tag{2}$$

[1] In our final specification, we obtain $\lambda = 0.18$. In order to obtain a stable shrinkage parameter, we estimate λ only with data until 2019Q4.

We emphasize that this 'informational decomposition' (Morley and Wong 2020) is not structural as the innovations $\boldsymbol{\varepsilon}$ are not necessarily orthogonal and economically interpretable. However, even though Eq. 2 does not permit a causal interpretation, it is a convenient instrument to shed light on the *transmitters* of variation in the output gap.

3 Data and model selection

Currently, the output gap literature is evolving from univariate and filter-based approaches to estimation in data-rich environments. For instance, Morley and Wong (2020) give an account of the US output gap using a similar approach as the paper at hand. Moreover, Jarocinski and Lenza (2018) and Barigozzi and Luciani (2021) estimate the output gap in a big data environment. While multivariate models have the advantage that more information can be analysed, they may be subject to over-fitting. Thus, we face a trade-off between eliciting a model that exploits all relevant information and a parsimonious specification.

Our variable selection procedure seeks to reconcile traditional economic approaches to estimating potential output with statistical information on variable relevance. To this end, we estimate a medium scale model including economic aggregates in the spirit of the production function approach to potential output estimation. That is, we assume output growth y_t (and thus, implicitly, the output gap) obeys a linearized Cobb-Douglas regime of the form

$$y_t = \alpha k_t + (1-\alpha) l_t + a_t,$$

where a_t is the Solow residual, k_t is capital formation, l_t denotes labour inputs and α is the substitution elasticity of capital. As the Solow residual cannot be subjected to direct analysis, we treat it as stochastic innovation. However, we can approximate capital k_t and labour l_t inputs by means of observable economic aggregates. Our choice of candidate variables is inspired by the literature on production function estimation. In particular, the included labour market aggregates largely derive from the EU commission's procedure on potential output estimation (Havik, McMorrow, Orlandi, Planas, Raciborski, Roeger, Rossi, Thum-Thysena, and Vandermeulen 2014). For approximating innovations to the capital stock, we propose to use a larger set of economic indicators broadly related to capital formation (e.g. investment and industrial production) and its costs (e.g. exchange and interest rates). Thus, we include various sectors of the real economy, the German labour market, external relations and financial markets. Subsequently, we estimate a candidate model and reduce the number of variables in accordance with a statistical criterion.

Table 1 depicts candidate variables, transformations and sampling frequencies. We sample data for the period of 1995Q1 until 2022Q1 in monthly frequency. If an indicator is not available on a monthly basis, we obtain it at quarterly frequency. The time period is constrained by data availability and by the German reunification (1990) which possibly caused a business cycle regime change that we omit from the model for our purposes. If not stated otherwise in Table 1, we obtain data from the Deutsche Bundesbank database. We partition our data-set into five variable blocks, which are ordered as shown above. All monthly series are ordered before all quarterly series. We emphasize that model invertibility is indispensable for our purposes (see Eq. 1). Thus, we apply convenient transformations to secure stationarity of each time series.

Variable	**Transformation**	**Frequency**
	CAPITAL	
External Relations		
CPB World Trade Monitor: World Trade Volume	rolling demean	M
Current Account: Exports	growth rates	M
Capital account Balance: Portfolio Investment	growth rates	M
Capital Account Balance: Direct Investment	growth rates	M
Real Effective Exchange Rate of the Euro against EERK-42	rolling demean	M
Finance		
Interbank Rate for Germany (obtained from FRED database)	growth rates	M
Total Share Prices for All Shares for Germany (obtained from FRED)	rolling demean	M
Term spread (1 year over 10 year government bonds)	growth rates	M
Non-financial private sector credit (obtained from BIS)	growth rates	Q
Fiscal Activity		
Government Consumption	growth rates	Q
Government Investment	growth rates	Q
Sentiment and Expectations		
OECD Consumer Opinion Surveys (obtained from FRED)	growth rates	M
ifo Business Climate Index (obtained from ifo Institute)	growth rates	M
ifo Business Expectations Index (obtained from ifo Institute)	growth rates	M
Real Economy		
Consumer Price Index	growth rates	M
Construction Permits	growth rates	Q
Industrial Production	growth rates	M
New Orders of Consumption Goods	growth rates	Q
New Orders of Investment Goods	growth rates	Q
New Orders of Input Goods	growth rates	Q
New Orders of Industrial Goods	growth rates	Q
Real Gross Domestic Product	growth rates	Q
Resource Price Index (excl. Energy)	growth rates	Q
Resource Price Index (only Energy)	growth rates	Q
	LABOUR	
Hours in Construction	growth rates	M
Labour Compensation Index	growth rates	Q
Labour Unit Costs	growth rates	Q
Labour Market Stabilization Policy ('Kurzarbeit' policy)	growth rates	M
Unemployment Rate	growth rates	M
Working Population	growth rates	M

Table 1: Variable blocks and data transformations. 'Growth rates' denotes the transformation 100 $\times$ first differences of natural logarithms. 'rolling demean' denotes a rolling demean (backward moving average) filter with a 40-quarters window. 'M' and 'Q' denote monthly and quarterly frequency, respectively.

Parsimony is the second most important priority after economic plausibility. To reduce the size of the model in the interest of parsimony, we proceed as follows. Subsequent to estimation of the model implied by Table 1, we compute the standard deviations of the informational decomposition contributions (see Eq. 2) to approximate the explanatory relevance of the variables for the output gap (following Morley and Wong (2020)). Higher standard deviations imply higher relevance of a given economic aggregate. Moreover, model selection is a reoccurring problem, as the model requires re-estimation after each quarterly period. Thus, it seems reasonable to stick to specifications that consistently implied decent estimates in the past. Therefore, we estimate the full model as shown in Table 1 for every quarter since 2008Q4. We then average over the time dimension of the normalized standard deviations of the contributions to explained variation for each variable.[2] Thereafter, we estimate a model that includes all variables with an average, normalized standard deviation of explained contribution to variance equal to or larger that of GDP growth (i.e. unity). Thus, we drop all variables from the model with relatively low explanatory power.

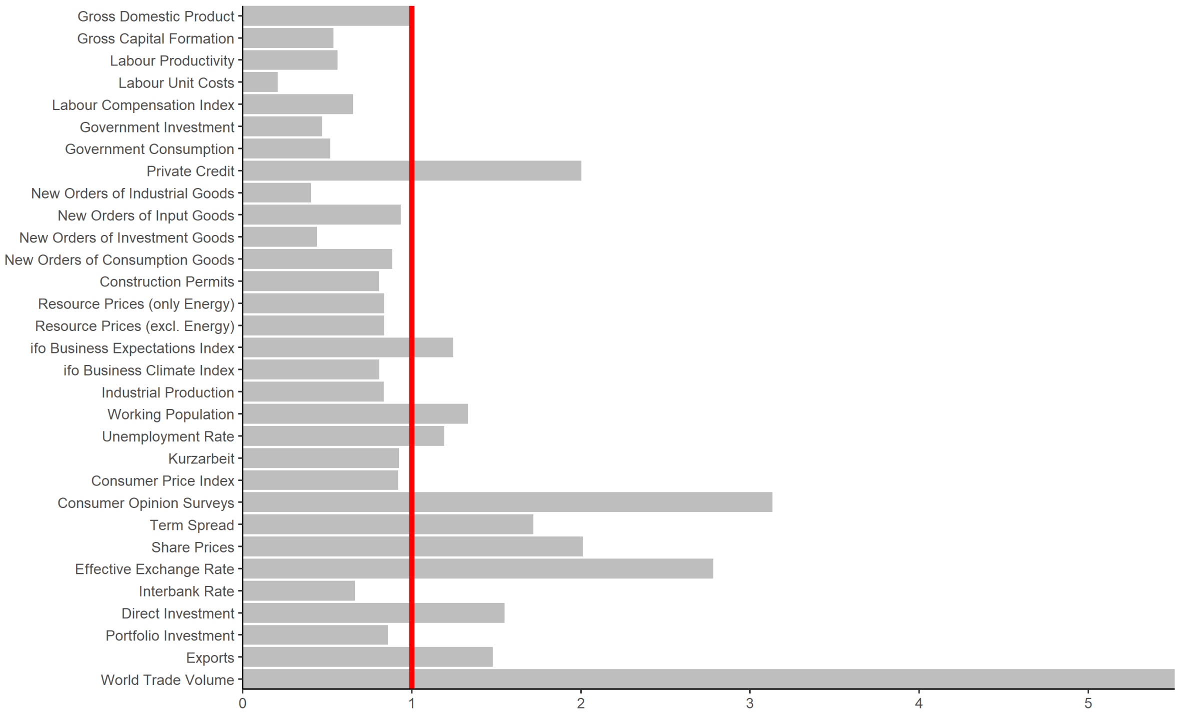

Figure 1: Normalized time-averages of standard deviations of informational decomposition. The red line indicates the standard deviation of the GDP growth contribution to output gap variation.

[2]We normalize by means of the standard deviation of GDP growth.

4 The German output gap

We now turn discussing the economic properties of the proposed output gap. In a first step, we discuss our estimate in more detail and assess its plausibility. Subsequently, we examine the contributions of the individual variables to informing our estimate in a reduced-form framework.

4.1 Estimate of the output gap

Figure 2 depicts the German output gap for 1995Q1 until 2022Q1.

As can be seen, the German economy suffered from substantial slack before the beginning of the millennium. This was likely due to high adjustment costs resulting from German reunification and the global millennium recession. Subsequent to the millennium recession and after the Dotcom-bubble burst in 2002, we observe strong overheating (+4%) prior to the financial crisis (2005 – 2008). Unsurprisingly, the European sovereign debt and banking crisis (2010 – 2015) coincides with sluggish mean-reversion tendencies of the output gap which fluctuated slightly below zero at the time.

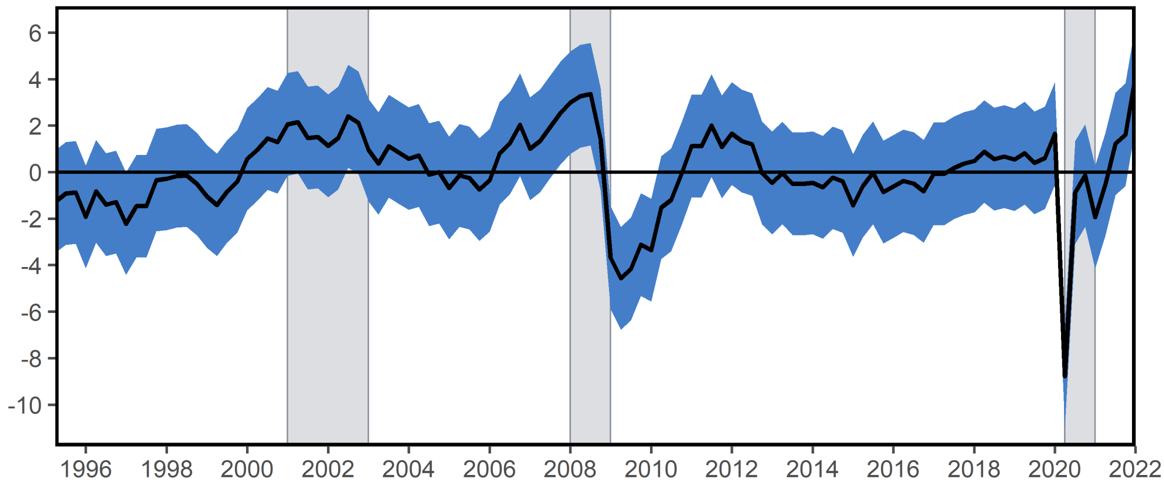

Figure 2: German output gap in percentage deviation from potential output, 1995Q1 - 2022Q1 (black). Blue and orange shaded areas indicate 90% credible sets following Kamber et al. (2018) for the mean estimate and the nowcast, respectively. Grey areas indicate recessions according to the German council of economic experts and the COVID-19 pandemic.

As a result of the COVID-19 shock the output gap dropped from the pre-pandemic level of 1.17 to −8.78% at the end of 2020Q2. Thus the −9.95% decline of the output gap is similar to the decline of German GDP growth (−10%) in 2020Q2. This implies that the COVID-19 shock only marginally affected potential output but is accounted for by a massive decline in the output gap. Interestingly, even the second and third 'lockdown' episodes (2020M11 – 2021M3 and 2021M4 – 2021M5) exerted – relatively – small contractionary pressures of about −1.5% on the German output gap. In the second half of 2021, the German output gap is still well below the zero mean while we see massive overheating at the beginning of 2022. The latter is most likely due to the Russian attack on Ukraine and the encompassing pressures on global energy, food and material supplies as well as due to inflationary pressures.

4.2 Is our output gap estimate plausible?

We turn to discussing the plausibility of the our results. To this end, Figure 3 depicts the comparison of our output gap estimate (black) with the GDP-based output gaps implied by the Hodrick and Prescott (1997) filter (green) with the smoothing parameter set to 1600 (as is common for quarterly data) and the Hamilton (2018) filter (orange) with $p = 4$ lags. Moreover, Figure 3 depicts a comparison of our estimate (black) to the official output gaps estimated by the German council of economic experts (dashed red) and AMECO (dashed blue). The latter two estimates are obtained from models that use a production function approach to approximate potential output. Both are only available at yearly frequency.

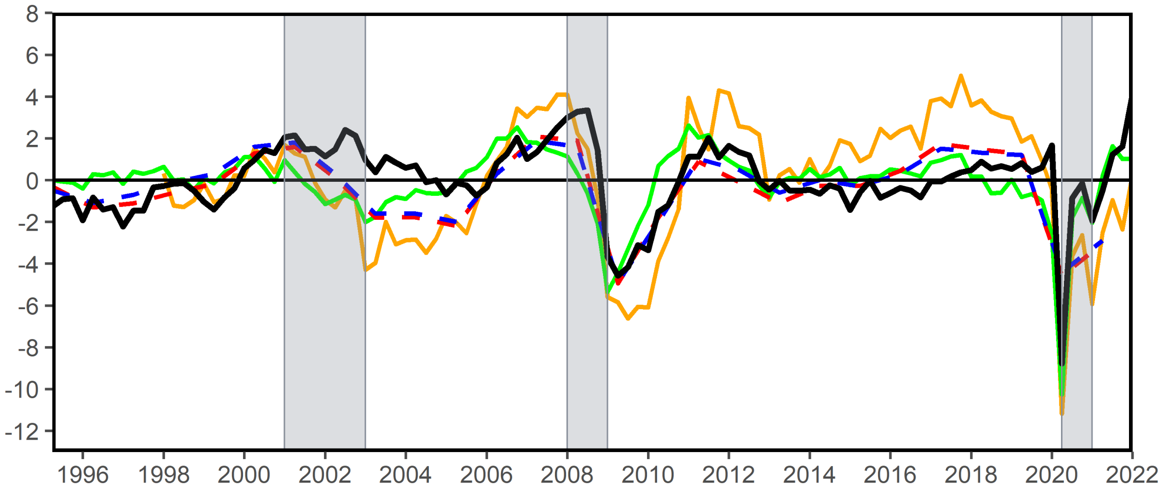

Figure 3: Comparison of our output gap measure (black) with established alternatives: One-sided Hodrick and Prescott (1997) filter (green) and Hamilton (2018) filter (orange). All estimates are reported in growth rates. Red and blue dashed lines are the (yearly-frequency) estimates by the German council of economic expert and AMECO, respectively. For further notes see Figure 2.

First of all, we note that all estimates are reasonable similar during most of the sample period. Our estimate differs in magnitude compared to the Hamilton regression filter during and prior to the two large recessions (2008 and 2020), as do all alternative estimates. In particular, we note that the magnitudes for the Hamilton-filtered estimate appear relatively large. For instance during 2016 – 2018, the Hamilton filter indicates an overheating almost as substantial as prior to the Great Recession in 2008. Given that no economic narrative is available to support this conjecture, this seems surprising. Thus, the

Hamilton filter seems to yield an 'outer bound' estimate of the German business cycle.[3]

Moreover, we observe that the HP-filtered estimates have a significant tendency to indicate recessions and more subtle economic downturns earlier (about one year, generally) than alternative estimates. This is exemplified prior to the financial crisis in 2008 and even more in advance of the COVID-19 health crisis. In case of the latter, the HP-filter produces results that indicate mean-reversal in mid 2017, which seems debatable (at least with respect to the shown magnitudes) in absence of a plausible economic narrative. For instance, German GDP grew between 2017 and 2018 about 0.5% to 1% almost each quarter compared to the previous quarter (except 2018Q1 and 2018Q3 with about −0.4% each). Given the HP-filtered negative growth rates of the output gap, this would imply substantial and implausibly high growth rates for potential GDP in 2017 and 2018. Morley and Wong (2020) suggest that a decent output gap estimate should be correlated positively with future inflation and negatively with future output growth. Thus, we compute correlations between the output gap estimate and the future quarter-on-quarter growth rates of output and consumer price index. We find that for the Hamilton filter, results are inconclusive (Pearson correlation coefficients of −0.11 for inflation and −0.26 for output growth). The coefficients of the HP-filtered estimate line up a little better with the economic expectation (coefficients of 0.09 and −0.16 with inflation and output growth, respectively). Our model compares well to the HP-filter (correlations of 0.01 and −0.50 with inflation and output growth, respectively). Summing up, we conclude that our estimate is economically at least as plausible as the output gaps obtained from filtering GDP by means of the procedures proposed in Hamilton (2018) and Hodrick and Prescott (1997).

Moreover, we compare our proposed output gap to production-function based estimates. Our estimate implies about the same overheating tendencies prior to the financial crisis (2007) and prior to the COVID-19 shock (2019) as do both output gaps by AMECO (blue dashed) and the German council of economic experts (red dashed). Furthermore, all three models indicate sluggish regression to the mean in 2020 and 2021 at about the same pace and to about the same levels. We take this as evidence that our model yields reasonably similar approximations to production-function based approaches. However, note that our model indicates slightly more overheating during and less slack before the financial crisis of 2008. We conjecture that both differences can partly be explained with reference to the underlying conceptions of the output gap. Whereas our estimate includes information on financial markets, this is not incorporated in the production function approaches. This aspect likely explains the positive output gap during the 2001-2003 recession. The aforementioned alternative output gaps focus on real economic activity (without considering financial transactions and imbalances), whereas our estimate is best understood as a real indicator that takes into account *all* economic activity in Germany, including finance. Nevertheless, incorporating financial information in the course of estimating the output gap is important (as pointed out by Borio, Disyatat, and Juselius (2013) and Berger, Richter, and Wong (2022)) when it comes to judging the sustainability of output growth, e.g. due to financial imbalances. Overall, we are confident that our model yields a plausible estimate for the German output gap.

[3] Quast and Wolters (2022) offer a convincing explanation for the possibly spurious dynamics implied by the baseline Hamilton filter.

4.3 Informational decomposition

The results described above raise a number of questions regarding the key determinants of German business cycle fluctuations. The German output gap is much less researched than, say, the United States output gap. Therefore, even non-structural information is valuable to understand German business cycle fluctuations. Subsequently, we aim to contribute towards closing this research gap. Figure 4 shows the informational decomposition of the German output gap.

Although we refrain from drawing causal conclusions, Figure 4 yields interesting insights into the reduced-form contributions of our five variable blocks to business cycle variation. The capital inputs block accounts for roughly 86% of the variation (expectations 24%; external relations 35%; the real economy, i.e. GDP, 5%; financial 22%), whereas the labour block explains about 14% of the variation in the German output gap. The high relevance of the international aggregates does not come as a surprise, as the (very open) German economy is shaped by its external relations (Eickmeier 2007). In fact, Figure 4 unambiguously shows that international economic aggregates shape the output gap dominantly throughout the entire sample period. That is, their importance does not seem to be regime-dependent.

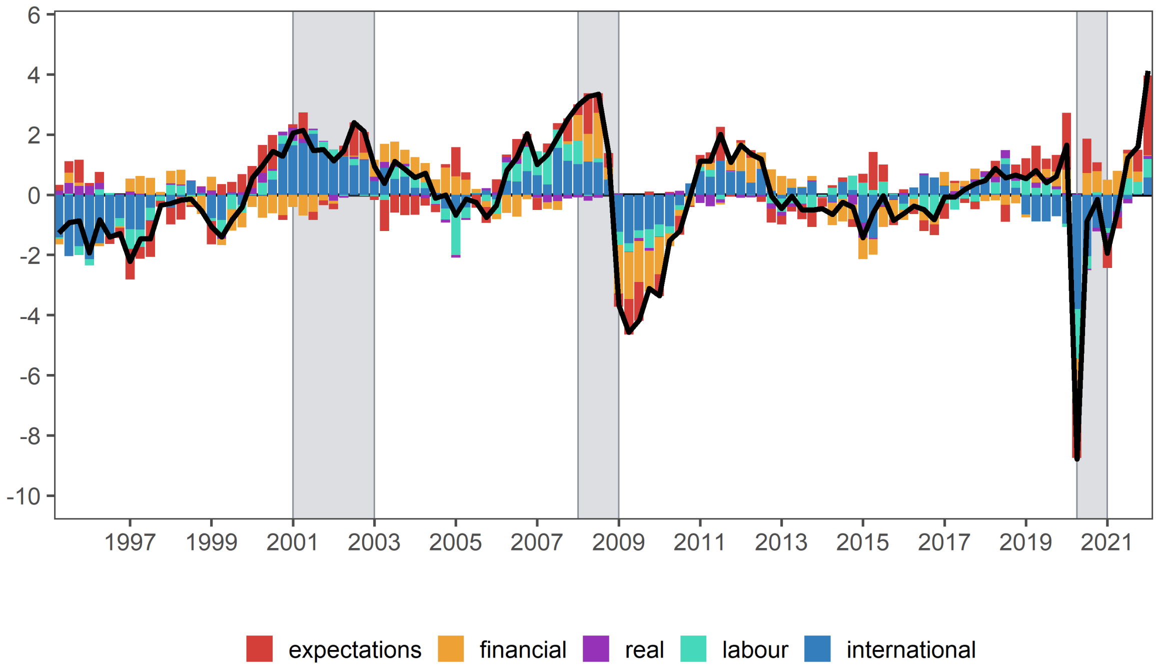

Figure 4: Informational decomposition of the German output gap. The financial block comprises private credit, share prices, the term spread, the real economy block summarizes contributions from gross domestic product, the labour block is made up of unemployment and the working population, the expectations block comprises consumer sentiment and business expectations and the international block contains the direct investment, the exchange rate, exports and world trade. For further notes see Figure 2.

Financial market information is mostly relevant in times of large overheating (2006-07) or substantial economic slack (e.g. during the late European banking and sovereign debt crisis in 2012-15), less so when the output gap reverts to its mean. Similar results for the US have been reported by Berger, Richter, and Wong (2022). However, this does not imply that the underlying *structural* real or financial shocks are irrelevant for the German business cycle, as the informational decomposition remains silent on this issue. Moreover, expectations explain large amounts of variation in the output gap. In particular, the strong increase in 2022Q1 can largely be explained by variation that is transmitted via expectations. Furthermore, we observe that labour market innovations contribute large amounts of variation in times of large and spontaneous contraction (e.g. 2020Q2). However, we emphasize that this interpretation does point to labour market shocks in a structural sense.[4]

To assess the cyclicality of the variable blocks' contributions to the German output gap more broadly, we compute correlations of the contributions to explained variation and the estimated output gap. The results are summarized in Table 2. Any sensible output gap estimate should be correlated positively on average with its explained variation by both capital and labour inputs, as displayed in Figure 4. As we see in Table 2, this is the case. In fact, the correlation coefficients are relatively large. In particular, the shares of variation in the output gap explained by the labour market and international as well as financial aggregates behave (strongly) pro-cyclically. In the short-run this also holds for expectations. Moreover, if expansionary shocks are transmitted via these variable blocks, the next-quarter output gap will increase on average. Conversely, we cannot infer causal chains at this point: As the output gap widens (closes), larger (smaller) shocks are transmitted by expectations as well as international, financial and labour aggregates. The variation explained by GDP itself is hardly informative for the current or near-future output gap.

	CAPITAL				LABOUR
	Expectations	*External Relations*	*Finance*	*Real economy*	
c_t	0.61*	0.78*	0.47*	-0.01	0.64*
c_{t+1}	0.19*	0.58*	0.22*	-0.04	0.44*
c_{t+2}	0.12	0.48*	0.07	-0.09	0.35*
c_{t+3}	0.07	0.40*	-0.03	-0.09	0.27*
c_{t+4}	-0.08	0.27*	-0.20*	-0.05	0.14

Table 2: Pearson correlations between average contributions to explained variation with the current and next-four-quarters German output gaps c_t as well as c_{t+1}, c_{t+2}, c_{t+3} and c_{t+4}. * marks correlation coefficients in excess of $2/\sqrt{T}$, which roughly corresponds to a significance level of 95%.

Interestingly, shock transmission from the international block exhibits quite substantial correlations with the four-quarters-ahead output gap. This finding is unique to international aggregates. For example, shock transmission from the labour market (which is associated with the largest contemporary and one-quarter-ahead correlations) are not

[4]We present a structural historical decomposition on Figure 10 in Appendix A.

too informative about the output gap beyond three horizons. We interpret this finding to point to the special relevance – and potential vulnerability – of the German economy to shocks transmitted by international aggregates. With regard to expectations, finance and the real economy, correlations at farther horizons confirm our previous conclusions.

5 Nowcasting performance in the baseline model

In times of economic disruption, a fast economic policy response is asked for. Waiting until GDP data has been released one month after the end of a quarter can be economically costly. Therefore, a timely estimate of the output gap is needed. We assess the nowcasting abilities of our approach. We proceed by analysing the nowcasting qualities of our model rigorously.

Table 3 shows the mean absolute forecast errors (MAE) for our model and given monthly indicators (first row for each indicator) compared to the end-of-quarter output gap estimate for the baseline model. Moreover, in order to assess the relevance of the individual variables, we compare two forecasts by means of the Diebold-Mariano procedure (Diebold and Mariano 1995): one forecast obtained from a unrestricted model with all variables and another forecast from a restricted model with the monthly indicator of interest omitted. More precisely, we test whether the forecast of the latter is superior in terms of mean absolute error than the forecast of the former model.

	1^{st} **month**	2^{nd} **month**	3^{rd} **month**
Share prices	0.39	0.28	0.12^{**}
Term spread	0.38	0.26^{*}	0.11^{**}
Consumer opinion surveys	0.34^{**}	0.16^{**}	0.06^{**}
Business expectations	0.28	0.15	0.04
Unemployment rate	0.31^{*}	0.15^{**}	0.05^{**}
Working population	0.29^{*}	0.15	0.05
World trade volume	0.43^{**}	0.28	0.15
Exports	0.41	0.28	0.15
Direct investment	0.40	0.28	0.14
Effective exchange rate	0.40	0.27	0.14

Table 3: Within-quarter mean absolute forecast errors associated with monthly variables rounded to two decimal places for all four models, given the full-sample parameters. ** and * indicate Diebold-Mariano p-values (based on mean absolute error) equal to or smaller than 0.05 and 0.10, respectively. Variables are ordered by expected release.

The unconditional forecast (i.e., a forecast unconditional on current-quarter information) implies a mean absolute forecast error of 0.51. From Table 3 we see that the mean absolute forecast error is moderate throughout compared to an unconditional forecast with within-quarter information. Effectively, the additional information incorporated throughout the course of the quarter cuts the nowcast error almost in half (from 0.51 to

0.28 percentage points). At the end of the third month, the mean absolute forecast error is negligible (0.04). We emphasize that this still is several weeks prior to a GDP data release in the subsequent quarter. Observing the world trade volume, the working population, unemployment and consumer opinion help to improve forecast precision in the first month. Subsequently, consumer opinion, the term spread and the unemployment rate help to improve the forecast in the second month, whereas consumer opinion, share prices, the term spread and the unemployment rate do so in the third month of a given quarter.

Following BMW (forthcoming), Table 4 depicts correlations of the within-quarter output gap nowcasts (unconditionally and after a given month) and the final estimate (left block) as well as model-implied and realized output growth (right block).

	Output gap	Output growth
No information	0.92	0.91
First month	0.97	0.94
Second month	0.99	0.96
Third month	1.00	0.96

Table 4: Correlations of the within-quarter nowcasts with the final estimate, model-implied and real-time output growth

As can be seen, the model benefits from the high degree of persistence in the output gap (see first row of Table 4), but our estimate substantially improves upon an unconditional forecast after only a single month. Observing GDP hardly increases the correlations after seeing data for three complete months. This picture hardly changes when we consider correlations between model-implied and realized output growth: Our specification implies plausible output growth rates and observing GDP data after three month adds only little information.

6 Real-time reliability

In this Section, we analyze the real-time performance of our model. As Orphanides and van Norden (2002) point out, real-time estimates of the output gap are chronically unreliable. We extensively discuss whether our baseline model is subject to this charge. We begin with assessing our model's nowcasting abilities in a real-world setting (i.e. including model selection). Subsequently, we investigate estimation revisions and decompose the ex post output gap revision into effects resulting from quarter-on quarter changes in model selection (i.e. different variables), parameter revisions (i.e. different sample periods for the baseline specification) and data revisions[5] (for a given specification and given parameters).

[5]Unfortunately, we cannot obtain real-time data for the majority of the monthly indicators from Deutsche Bundesbank. Thus, we limit ourselves to investigating the effects of GDP data revisions, which are available since 2005.

6.1 Nowcasting in real-time

The COVID-19 health crisis has brought about the most devastating economic shock since World War II to many advanced economies across the world. Germany is no exception. According to our estimates, the German output gap was at historical low (-8.8%). In this situation it is crucial for policy makers to obtain real-time insight into the state of the economy to adjust or maintain stabilization measures. In times of such large (exogenous) shocks neither the unconditional forecast nor the ex post estimate are particularly useful for the conduct of stabilization policy. The former can – by construction – not forecast large shocks, and the latter is obtained far too late for a timely policy response. The mixed frequency sampling of our approach allows us to nowcast the output gap in a timely matter.

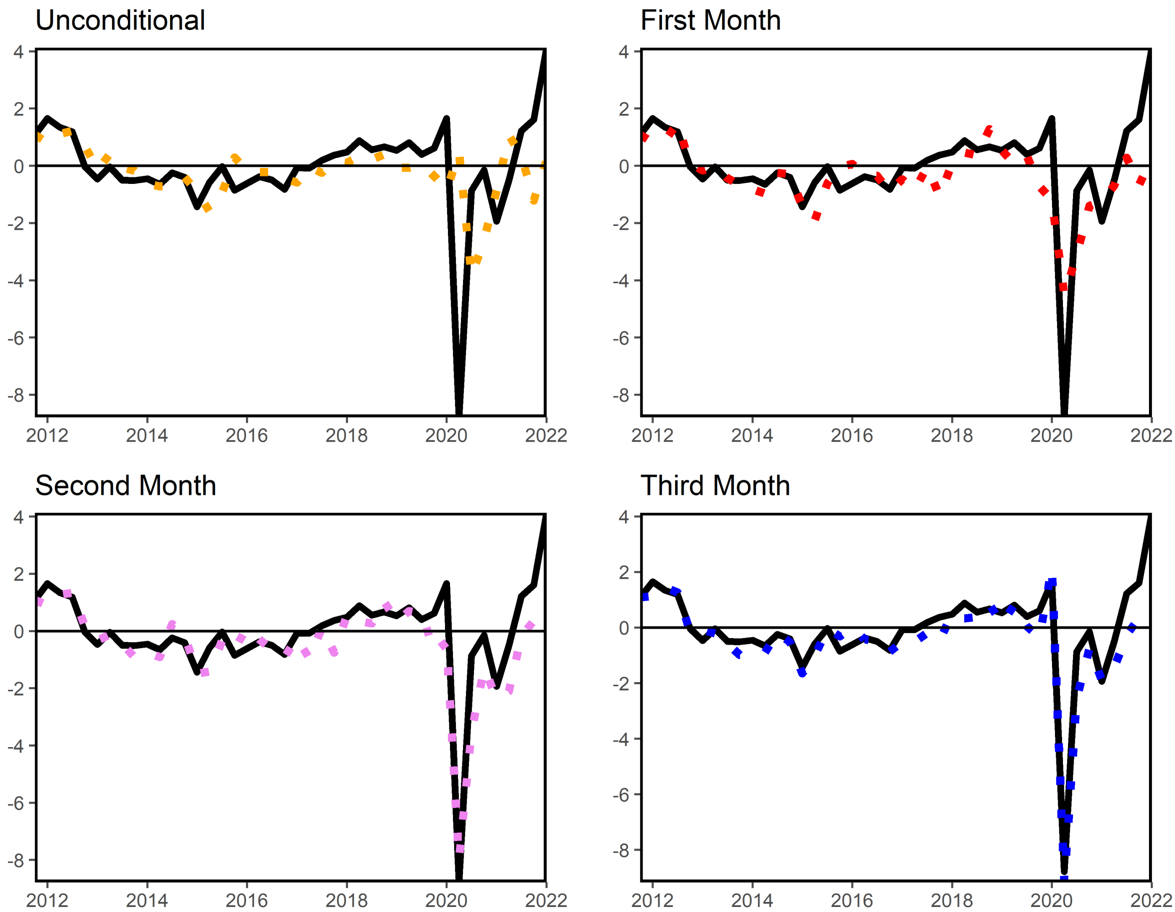

Figure 5: Nowcasts since 2012. Ex post estimate (black), unconditional forecast (orange, upper left panel) and nowcasts from the perspective of month 1 (red, upper right panel), month 2 (pink, lower left panel), month 3 (blue, lower right panel). Vintage GDP data is used, parameters are estimated and the specification is chosen in real-time (i.e. every quarter).

Figure 5 compares the unconditional forecast, the nowcasts after one month, two months, three months, and the final estimate of the output gap since 2012 (coinciding with the European banking and sovereign debt crisis) in real-time. That is, we estimate the model using vintage GDP data, automatically update the specification (i.e. the set of included aggregates) and re-estimate the parameters in every quarter. The orange dotted line (upper left panel) is the unconditional forecast after observing all data until the previous quarter. As can be seen, it is mean-reverting – and for large shocks (e.g. COVID-19) distant from the final estimate. However, considering the case of 2020Q2, once data for April 2020 is brought in (red dotted line, upper right panel), we note that the nowcast improves substantially. Once the nowcast incorporates data for May 2020, the difference is even smaller and with all monthly, but no quarterly information (blue dotted line), the real-time estimate is -9.2% and very close to the final (ex post, i.e. full sample) estimate (-8.8%). Thus the model provided a reasonable quantification of the COVID-19 shock after observing all 2020Q2 monthly data, without seeing private credit nor GDP. These conclusions generalize to the entire sample period under investigation with the interesting exception of 2022Q1. We conjecture that this behaviour occurs due to the relatively short sample period. Put differently, the model has been specified and estimated on data that does not include such sudden and large upward output gap expansions. Overall, our model yields stable nowcasts, even under real-world conditions including revised data, revised parameters and revised model specification.

6.2 Examining output gap revisions

A reliable model of the output gap not only implies decent nowcasts, but also small revisions. To investigate how much the estimated output gap is revised ex post, we pseudo-decompose the revisions into contributions from specification revisions (i.e. different cross-sectional information that is chosen automatically over time), parameter revisions (i.e. the effect of more information in the time dimension on the baseline specification) and GDP data revisions. We evaluate the associated revisions in the spirit of Orphanides and van Norden (2002). That is, we obtain the difference between the final full-sample estimate (i.e. the benchmark) and (pseudo-)real-time estimates. Subsequently, we compute various loss statistics. For instance, if revisions are small, the mean and the standard deviation of the difference should be small as well, it should not be autocorrelated (to make sure the revision is not systematic) and no extreme deviations should occur. The results are summarized in Table 5.

	Mean	Sd	Min	Max	RMSD	AR(1)
y_t - y_t^1	-0.01	0.05	0.04	-0.07	0.07	-0.01
y_t - y_t^2	0.01	0.04	0.04	-0.13	0.07	0.01
y_t - y_t^3	0.58	0.64	0.61	-0.33	1.53	0.52

Table 5: Summary and loss statistics in the spirit of Orphanides and van Norden (2002) of the difference of the baseline output gap y_t and the result taking into account parameter revisions (y_t^1), GDP data-revisions (y_t^2) as well as real-time model selection (y_t^3). From left to right: mean, standard deviation, minimum, maximum, root mean squared difference and the AR(1) regression coefficient of y_t - $y_t^{\{1,2,3\}}$, none of which is significantly different from zero at conventional levels.

Table 5 yields two key results. First of all, revisions in GDP data and the baseline model parameters are, on average, irrelevant for output gap revisions. That is, given our baseline specification, the estimated gap is hardly, if at all, revised ex post, indicated by the small mean revisions (−0.01 and 0.01, respectively). Secondly, if a different model specification is chosen, the output gap may be revised. However, the revision due to specification changes is still quite small and on average hardly different from zero. In the next step, we examine the revisions due to GDP data updates, parameter estimations and model selection in more detail across the time dimension.

6.2.1 Model selection in real-time

Our model selection procedure is designed to yield a stable specification. It ensures that only aggregates that consistently explain at least as much reduced-form variation over several periods in the output gap as GDP growth enter the model. However, as we run model selection since 2008Q4, the implied cross-section might in fact change from time to time. Figure 6 shows the implications for output gap revisions. As is apparent, they are larger than the previously discussed sources for revisions. Especially at the beginning of 2022, the revision is obvious. However, even in times of large economic stress (e.g. 2020), the specification is very stable. Overall, the revisions due to changes in the specification are small.

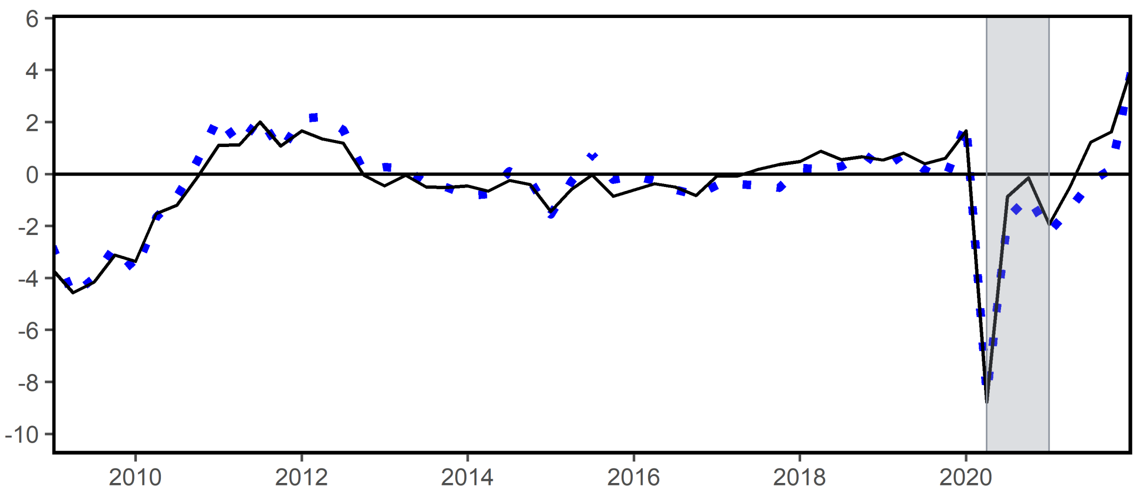

Figure 6: Black sold line: Ex post estimate. Blue dotted line: Estimate with real-time model selection. For further notes, see Figure 2.

6.2.2 Parameter revisions in pseudo real-time

After the previous quarter has been observed, our model needs updating in order to incorporate all available information. That is to say, even if the specification does not change, the parameters of the autoregressive vector as well as the shrinkage factor λ need to be estimated again. In the following, we examine the effect of parameter revisions on our estimate and nowcasting performance. To this end, we estimate the model until 2008Q3 as an initialization. Then we elicit a pseudo real-time update (ignoring data-revisions and specification changes for the moment) by adding in full-quarter new information of the subsequent quarters sequentially. Thus, we re-estimate the model every quarter conditionally on ex post data, the specification and the entire-quarter information. Figure 7 depicts the pseudo real-time output gap obtained from this procedure. Clearly, the difference between the real-time estimate and the ex post output gap is negligible in the entire period under investigation.

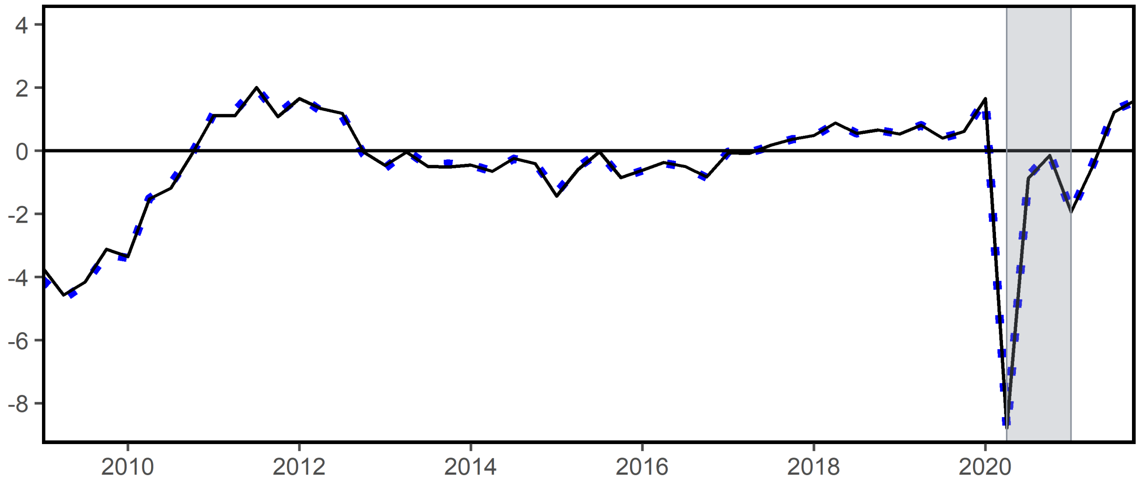

Figure 7: Black sold line: Ex post estimate. Blue dotted line: Estimate with real-time parameters in the baseline specification. For further notes, see Figure 2.

6.2.3 GDP data revisions in pseudo-real-time

In the final real-time analysis, we investigate the relevance of data revisions. We re-estimate the model for each quarter, given the baseline specification and the full-sample monthly information. Note that there are two sources of revisions in GDP. First of all, as we use GDP data chained in previous year prices, the level of GDP in the past years is adjusted in each first quarter of a given year. We expect this effect to be small, as we employ GDP in growth rates. Moreover, since 2005, GDP data is revised in the next quarter after the initial release. From Figure 8, we clearly see that these revisions do not play a role.

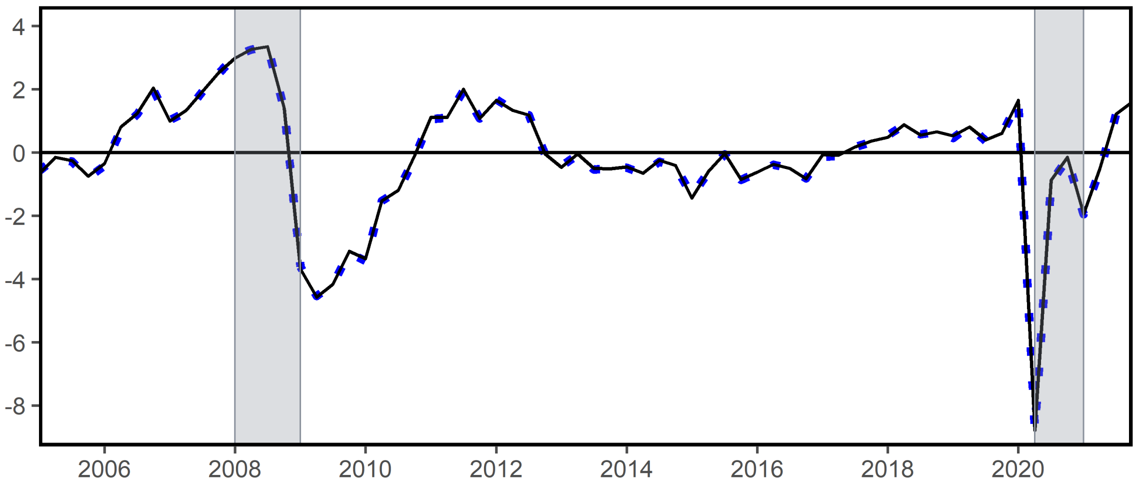

Figure 8: Black sold line: Ex post estimate. Blue dotted line: Estimate with GDP revisions. For further notes, see Figure 2.

7 Robustness: Larger information set

In this section, we investigate the effects of including a larger set of variables in our model. We estimate a model that comprises all variables of the baseline specification and all variables with a standard deviation of explained shares of variation larger than the median standard deviation of explained shares of variation of the aggregates that are not included in the baseline model. That is to say, in addition to the baseline variables, we include the consumer price index, Kurzarbeit and indices for new order for consumption goods and input goods. Figure 9 depicts the German output gap estimated from the alternative model. We note that for most of the period under examination, the two estimates are very similar. The difference for 2022Q1 is, most likely, due to the inclusion of the consumer price index.

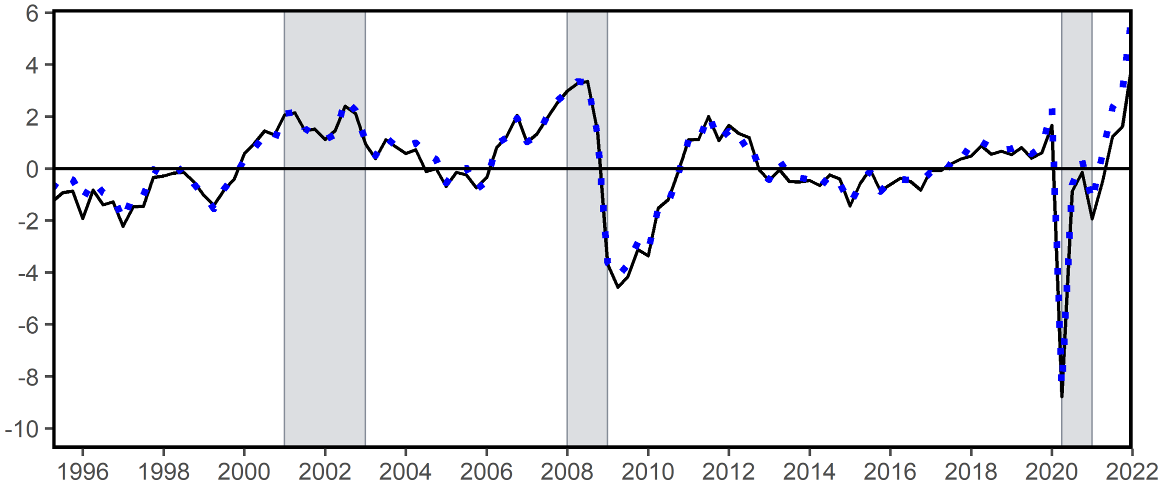

Figure 9: Black sold line: Ex post estimate. Blue dotted line: Estimate based on larger information set. For further notes, see Figure 2.

8 Conclusion

We have provided an in-depth analysis of the international and domestic determinants of the German output gap between 1995 and 2022 by means of a medium-size mixed-frequency vector-autoregressive model that exploits monthly information to evaluate the expectation associated with a multivariate Beveridge-Nelson decomposition. We showed that substantial shares of variation in the German business cycle are explained by expectations, the labour market and international economic aggregates. Moreover, we demonstrated that our model fairly accurately predicts the German output gap up to three months prior to a German gross domestic product data release. In particular, observing consumer sentiment and the labour market allows to produce a decent nowcast of the German output gap.

References

Barigozzi, M. and M. Luciani (2021): "Measuring the Output Gap using Large Datasets," *The Review of Economics and Statistics*, 1–45.

Berger, T., J. Morley, and B. Wong (forthcoming): "Nowcasting the output gap," *Journal of Econometrics.*

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.

Borio, C., F. P. Disyatat, and M. Juselius (2013): "Rethinking potential output: Embedding information about the financial cycle," BIS Working Papers 404, Bank for International Settlements.

Diebold, F. and R. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

Eickmeier, S. (2007): "Business cycle transmission from the US to Germany—A structural factor approach," *European Economic Review*, 51, 521–551.

Hamilton, J. D. (2018): "Why You Should Never Use the Hodrick-Prescott Filter," *The Review of Economics and Statistics*, 100, 831–843.

Havik, K., K. McMorrow, F. Orlandi, C. Planas, R. Raciborski, W. Roeger, A. Rossi, A. Thum-Thysena, and V. Vandermeulen (2014): "The Production Function Methodology for Calculating Potential Growth Rates Output Gaps," *Economic Papers*.

Hodrick, R. J. and E. C. Prescott (1997): "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 1–16.

Jarocinski, 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.

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.

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.

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.

Quast, J. and M. H. Wolters (2022): "Reliable real-time output gap estimates based on a modified Hamilton filter," Tech. rep.

Waggoner, D. and T. Zha (1999): "Conditional Forecasts In Dynamic Multivariate Models," *The Review of Economics and Statistics*, 81, 639–651.

Appendix

A Structural historical decomposition

In this Appendix, we briefly present a structural historical decomposition. Structural identification of our model is a challenging task, as the mere size due to the mixed-frequency setup, complicates the analysis tremendously. Traditional identification schemes based on sign restrictions and more recent data-driven alternatives can hardly cope with a system of this size. For instance, the rotation space for identification based on sign restrictions is vast even for a relatively small number of restrictions. Thus, we base our brief analysis on a Choleski factor of the reduced form error covariance and leave more sophisticated structural identification to future research. Figure 10 depicts a structural decomposition.

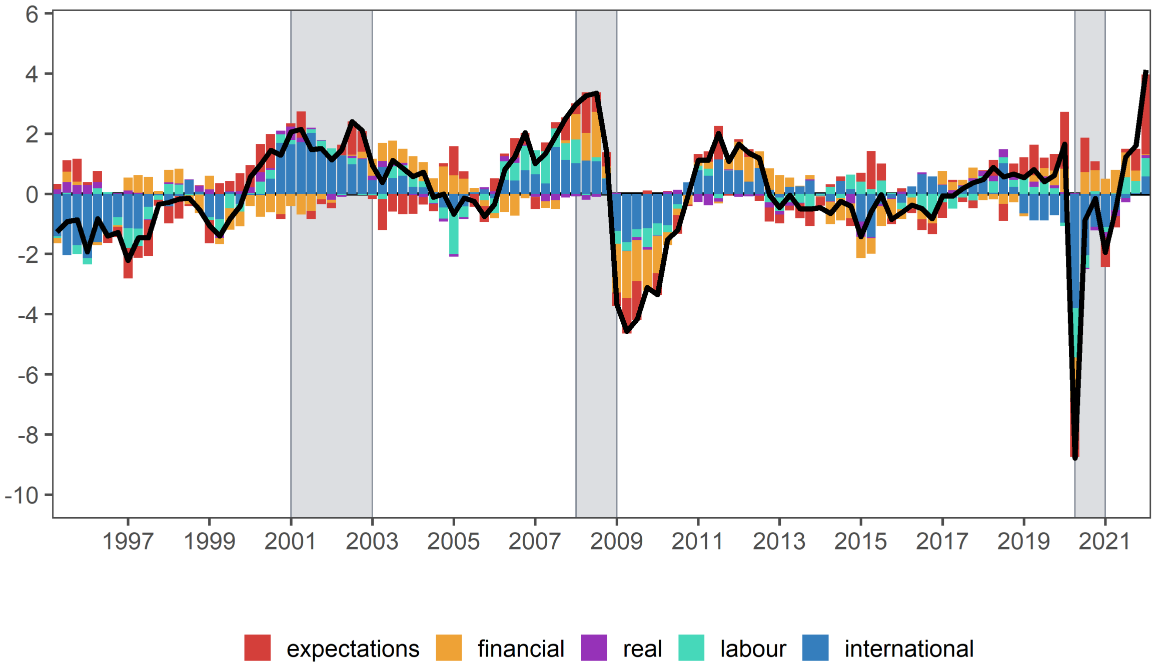

Figure 10: Structural historical decomposition based on Choleski decomposition of the German output gap. For further notes see Figure 4.