MASARYK UNIVERSITY

FACULTY OF ECONOMICS AND ADMINISTRATION

Three Essays in Macroeconomics: A Comprehensive Framework in Macroeconomic Policy Evaluation

Habilitation Thesis

JAN ČAPEK

Brno 2021

MUNI
ECON

Contents

1 Introduction

The habilitation thesis includes three essays in macroeconomics with commentary. All essays present the results of empirical analyses of macroeconomic policy in a comprehensive framework. The first essay focuses on monetary policy, while the second and the third analyze fiscal policy effects. All studies are data-based and use time series of various macroeconomic variables.

The first essay in the collection, Čapek (2014), investigates the effect of real-time data on parameter estimates of monetary policy reaction functions. The second essay, Čapek and Crespo Cuaresma (2020), analyses the role played by data and specification choices as determinants of the size of the fiscal multipliers. The third essay, Čapek et al. (2021), estimates fiscal multipliers for Austria in a framework of model uncertainty emanating from the choice of a particular econometric model.

Čapek (2014) investigates the differences between parameter estimates of monetary policy reaction functions using real-time data and those using revised data. The model is a New Keynesian DSGE model of the Czech, Hungarian and Polish small open economies in interaction with the euro area. Unlike the related literature, this paper uses separate vintages of real-time data for all successive estimations. The paper reports several statistically significant differences between parameter estimates of monetary policy reaction functions based on real-time data and those based on revised data. The parameter whose estimate is the most affected by the usage of real-time data is the preference for output growth. This result is common across the countries in the study. The results suggest that real-time data matter when conducting a historical analysis of monetary policy preferences.

Čapek and Crespo Cuaresma (2020) analyse the role played by data and specification choices as determinants of the size of the fiscal multipliers obtained using structural vector autoregressive models. The results, based on over twenty million fiscal multipliers estimated for European countries, indicate that many seemingly harmless modelling choices have a significant effect on the size and precision of fiscal multiplier estimates. In addition to the structural shock identification strategy, these modelling choices include the definition of spending and taxes, the national accounts system employed, the use of particular interest rates or inflation measures, or whether data are smoothed prior to estimation. The cumulative effects of such arguably innocuous methodological choices can lead to a change in the spending multipliers of as much as

0.4 points.

Čapek et al. (2021) estimate fiscal multipliers for Austria in a framework of model uncertainty emanating from the choice of a particular econometric model. We present a comprehensive framework which allows to assess the effects of different multiplier definitions and choices related to the data, the model employed, and further technical choices associated with the specification of the model exert on fiscal multiplier estimates. The mean present-value government spending multiplier over all models entertained, based on over one thousand estimates, is 0.94. Estimates of the peak spending multiplier tend to be larger than present-value spending multipliers, with a mean value of 1.08. The value of the mean present-value tax multiplier is -0.76 and the mean peak tax multiplier is -0.58 for all specifications used.[1]

This part of the habilitation introduces the comprehensive concept of the submitted collection of works, part 2 delivers the literature review, part 3 brings a brief insight into the methodology, part 4 states the contribution of the research and specifies the applicant's contribution, and part 5 discusses the limitations of the research and possible avenues for further research. All mentioned parts of the habilitation including the references and the seventh supplement serve as the unifying commentary in accordance with Masaryk University Directive No. 7/2017, Section 5 (1) b), and the Directive of the Faculty of Economics and Administration of Masaryk University No. 4/2019, Section 6, paragraph 3. The first to the sixth supplement to this document constitute the collection of previously published works in accordance with Masaryk University Directive No. 7/2017, Section 5 (1) b), and the Directive of the Faculty of Economics and Administration of Masaryk University No. 4/2019, Section 6.

1.1 Comprehensive framework

A comprehensive framework can be understood as a means of communicating analysis results, which illustrates the results under various reasonable variants or settings. It is somewhat akin to sensitivity or robustness analyses, which are typically focused on showing that the main results are unaffected by reasonable variations. A comprehensive framework shows the results under many different scenarios and illuminates the effects of different scenarios on the results.

[1] The paragraph is adopted from the article's version, which has been accepted by the journal. For more information, see page 34.

The idea of communicating uncertainty in economics can be traced back to Morgenstern (1950, 1963), although the focus of Morgenstern's book was more on the accuracy and errors in economic statistics. Since then, a large proportion of mainstream economics became quantitative and studies started to propose policy recommendations relying on data-based analyses. As this practice became more prevalent in the 1970s and 1980s, numerous studies start to address the fact that the published empirical results were often too fragile to reasonable variation and that the uncertainty connected to the estimates was not adequately communicated.[2] One of the earliest studies, which was fully focused on the problem of fragility of results of empirical studies, was Leamer (1985), which opens the study with the following first sentence:

> ”A fragile inference is not worth taking seriously.”

Leamer then illustrates his case with an example of a study by Ehrlich (1975), which found that capital punishment deters murders. However, the results were deemed so fragile that a battery of follow-up articles emerged, which addresses various specific omissions in the original study. Nevertheless, Leamer (1985) finds these disorganized studies not particularly helpful to understanding the roots of prevailing uncertainty and calls for a so-called ”global sensitivity analysis”, which would address the complete (relevant) neighborhood of used assumptions and would communicate its effects on the results:

> ”In principle, a global sensitivity study should be carried out with respect to all dimensions of the model in one grand exercise.” (Leamer, 1985, p. 311)

In order to operationalize the notion of ”global sensitivity analysis” Leamer (1985) introduced ”Extreme bounds analysis”, which instigated some controversy. Sala-i Martin (1997) runs two million growth regressions to show that Extreme bounds analysis does not deliver useful results and proposes to assign some level of confidence to respective determinants of economic growth instead of labelling the variables as ”robust” or ”nonrobust”.

In current literature, the focus is placed on the question of how to best communicate the uncertainty of the analysis results. The failure to appropriately account for

[2] The literature also addressed the issue of identification strategies used to reach the results, which is not a focus of this commentary. See e.g. Sims (1980) or Leamer (1983) for seminal contributions in this topic or Angrist and Pischke (2010) for a more recent review.

accuracy and errors from Morgenstern's era is apparent in many studies 60 years later. Manski (2019, abstract) describes the problem as

> "A prevalent practice has been to report policy analysis with incredible certitude. That is, exact predictions of policy outcomes are routine, while expressions of uncertainty are rare."

Aikman et al. (2011) offers an illustration of how to report uncertainty in macroeconomics. Interestingly, the depiction in Figure 8 in the article is very similar to how we present the uncertainty connected to the estimation of fiscal multipliers in Čapek and Crespo Cuaresma (2020) and Čapek et al. (2021).

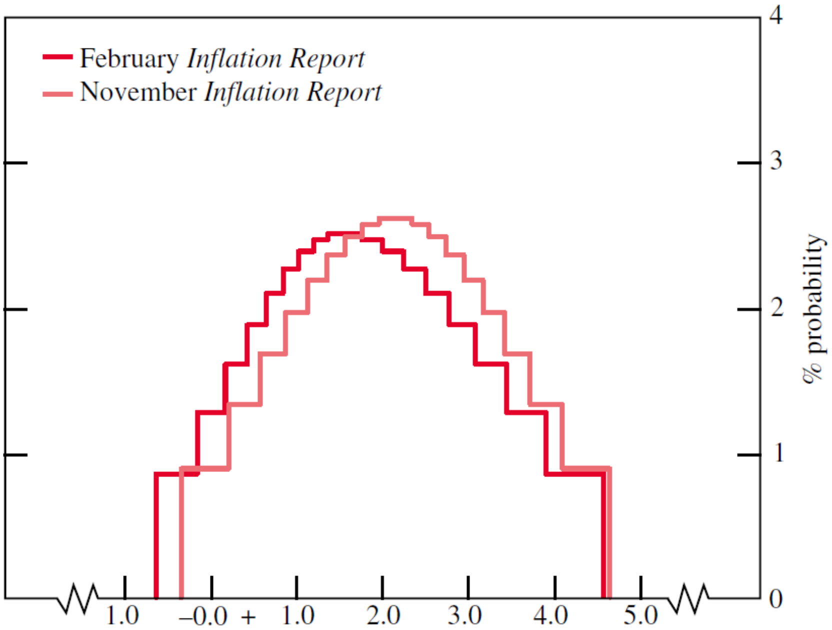


Figure 8. Distributional cross sections of inflation in 2012Q4, for February 2010 *Inflation Report* projection and November 2009 *Inflation Report* projection.

The body of literature on communicating uncertainty in science goes beyond economics - see e.g. Van der Bles et al. (2019) for a more generally focused review with case studies in economic statistics and climate change, or Hullman (2019) for a study focused more on journalism rather than scientific articles.

2 Literature review

This section introduces the literature relevant to the submitted essays and for further research.

Section 2.1 introduces the challenges of applying real-time data for macroeconomic analysis in general and section 2.2 then focuses on monetary policy evaluation with the use of real-time data. Section 2.3 reviews the literature of analysing fiscal policy effects, measured as fiscal multipliers. The literature on real-time data effects in fiscal policy, which is not covered in section 2.3, is available in a survey by Cimadomo (2016). The literature review is concluded in section 2.4 on macroeconomic forecasting, with and without the use of real-time data.

Sections 2.1 and 2.2 are relevant in the context of Čapek (2014), while section 2.3 addresses the literature followed in Čapek and Crespo Cuaresma (2020) and Čapek et al. (2021). Review in section 2.4 relates to further research (see part 5, page 22).

2.1 Real-time data

The unreliability of real-time macroeconomic data is a well-known issue and many studies have investigated the properties of data revisions. Orphanides and Norden (2002) report that revision of U.S. published data is not the main issue; it is the unreliability of end-of-sample trend estimates. These results are confirmed by Marcellino and Musso (2011) on euro area data, and by Ince and Papell (2013) on data for nine OECD countries. On the other hand, Cusinato et al. (2013) find that data revision and the end-of-sample problem contribute to uncertainty about the Brazilian output gap, but do not find any evidence that the former is less important than the latter. Investigating the empirical properties of U.S. macroeconomic data, Aruoba (2008) finds that the revisions are biased and predictable. For European countries, Giovannelli and Pericoli (2020) report that governmental forecasts of real GDP growth are biased. Rusnák (2013) reports that revisions of Czech GDP and its components are rather large. He also studies whether the revisions are "news" or "noise", i.e. whether the revisions are predictable or unpredictable, and ascertains in-sample predictability and out-of-sample unpredictability for most variables of interest. Interested readers are referred to Croushore (2011) for an extensive survey of the real-time data literature.

Real-time databases are available for the U.S.[3], OECD countries[4], and also euro area[5]. Problematic features of real-time data bring about difficulties for macroeco-

[3] Croushore and Stark (2001), https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/real-time-data-set-for-macroeconomists

[4] McKenzie (2006), https://stats.oecd.org/Index.aspx?DataSetCode=MEI_ARCHIVE

[5] Giannone et al. (2012), https://eabcn.org/eabcn-real-time-database

nomic policy evaluation and macroeconomic forecasting.

2.2 Monetary policy decision-making in real-time

There are many possible problems with macroeconomic data revisions and they are quantitatively of variable importance in different countries. However, the fact that the data revisions are large does not necessarily mean that they must create problems for economic agents. A researcher who wants to find out how significant the revisions are for decision-making in different situations must incorporate real-time data or data revisions into the decision-making process and observe if there are noteworthy differences in the results. A great deal of research effort has focused on the analytic consequences of bad real-time data quality for monetary policy. The literature goes back to Maravall and Pierce (1986), who investigate the conduction of monetary policy in real-time and ask whether the policy would have been different if final data had been available. The authors conclude that the answer to that question is no. A similar research question is also investigated by a stream of literature that focuses on the monetary rule under real-time conditions: Orphanides (2001) uses a Taylor-type rule to look at the effect of different policy recommendations using real-time data compared to final data and argues that monetary policy reaction functions estimated on final data provide a misleading description of historical policy. Similar results, i.e. that real-time data play a (significant) role, were also reached by Meirelles Aurelio (2005), Gerdesmeier and Roffia (2005), Gerberding et al. (2005), Horváth (2009), and Belke and Klose (2011). More recent literature also uses DSGE models with monetary rules as a tool for monetary policy investigation. Vázquez et al. (2010) and Casares and Vázquez (2016) find that monetary policy parameters are robust to real-time specification. On the other hand, Neri and Ropele (2012) show that there is indeed a statistically significant difference in policy parameters when real-time data are considered (rather than final data).

2.3 Fiscal policy effects

The estimation of fiscal multipliers (the ratio of the change in output to an exogenous change in government spending or taxes) is a central element for the evaluation of the macroeconomic effects of fiscal policy. Fiscal multipliers can be communicated and compared easily across different countries and time periods and the precision of their estimation contributes significantly to the quality of GDP growth predictions (Blanchard and Leigh, 2013).

The main bulk of the existing literature on the macroeconomic effects of fiscal interventions can be categorized as either model-based or empirical. Model-based approaches typically employ calibrated DSGE models to study the effects of fiscal stimuli in an internally-consistent theoretical framework. Kilponen et al. (2015), for instance, compare such estimates of fiscal multipliers across models and countries in Europe, while Barrell et al. (2012) focus on model-based fiscal multipliers in the context of fiscal consolidation. The advantage of the model-based approach lies in the ability to analyse counterfactual scenarios by simulating the dynamics of the model variables under different conditions. On the other hand, empirical approaches, mostly based on SVAR models, tend to be more data-driven and typically impose less stringent restrictions on the structure of the economic model. The availability of long time series for some countries allow for the use of modern identification methods such as the narrative approach (Ramey, 2011b) to extract exogenous fiscal shocks or the assessment of different regimes (Auerbach and Gorodnichenko, 2012) where fiscal multipliers may differ. In cases where such long time series are not available, countries are often pooled and the empirical analysis is conducted on a panel setting (Beetsma and Giuliodori, 2011; Ilzetzki et al., 2013), or fiscal multipliers for single economies with shorter time series are studied using SVAR models inspired by the seminal contribution by Blanchard and Perotti (2002).[6]

The estimates of fiscal multipliers tend to differ, sometimes strongly, from study to study (see the evidence presented in the meta-analysis provided by Gechert, 2015). These differences can be attributed to various identification strategies (Caldara and Kamps, 2017) as well as to other technical choices made in the analysis (Čapek and Crespo Cuaresma, 2020).

The interest in assessing the macroeconomic effects of fiscal policy in industrialized countries has gained renewed momentum since the Great Recession. Given the limited scope of action of monetary policy in the context of very low nominal interest rates, fiscal policy re-emerged as a policy of choice and a large literature has concentrated on investigating how fiscal policy affects macroeconomic variables and GDP in particular.[7]

[6] See e.g. Ramey (2016) for a review of the methods used for the identification of exogenous fiscal shocks.

[7] See e.g. Hebous (2011) or Ramey (2011a) for earlier surveys on the issue, or Ramey (2019) for a recent contribution.

There is little evidence on the size of fiscal multipliers for developed European small open economies.[8] Ravn and Spange (2012) enhance the Blanchard-Perotti methodology based on structural vector autoregression (SVAR) models to estimate spending multipliers for Denmark and obtain a point estimate of approximately 0.6 after four quarters. Jemec et al. (2011) investigate Slovenian fiscal policy employing a standard SVAR approach and estimate an impact spending multiplier of 1.5, which diminishes in subsequent periods. Unfortunately, not all studies investigating the effects of fiscal stimuli report the results in the form of multipliers (e.g. Afonso and Sousa, 2011, for Portugal or Benetrix and Lane, 2009, for Ireland). In addition to estimates for single countries, evidence from panel studies also exists. Ilzetzki et al. (2013) report that the subgroups of countries corresponding to high income, open, low-debt and fixed exchange rate countries have average spending multipliers of 0.4, 0, 0.2, and 0.6, respectively. The empirical evidence can be supplemented making use of the work by Barrell et al. (2012), where a model-based consumption multiplier of 0.5 is reported for Austria. Breuss et al. (2009) provides an overview of fiscal multipliers derived by Austrian forecasting institutions from large-scale macroeconometric models (within the tradition of the *Cowles commission approach*). Spending multipliers over the first year after the fiscal shock are typically below unity, first-year wage and income tax multipliers are below 0.5. Recent papers by Koch et al. (2019) and Schuster (2019) complement the existing results by simulating fiscal multipliers for Austria using calibrated New-Keynesian general equilibrium models and derive multipliers of comparable magnitudes.

2.4 Macroeconomic forecasting

The forecasting ability of Dynamic Stochastic General Equilibrium (DSGE) models has been a topic of debate in the academic literature over the last decade. Some of the existing empirical results suggest that the forecasting performance of DSGE models can reach (and in some cases surpass) that of econometric time series models (see e.g. Adolfson et al., 2007; Del Negro and Schorfheide, 2013; Wolters, 2015). The fact that DSGE models can be used by policy institutions not just for forecasting, but also for policy analysis, makes them a useful and versatile tool to address multiple questions on short-run and medium-run macroeconomic developments (Christiano et al., 2018;

[8] See the extensive summary of existing multiplier estimates in Mineshima et al. (2014) or the data used for the broad meta-analysis in Gechert (2015).

Lindé, 2018).

However, many studies indicate that the forecasting performance of DSGE and econometric time series models does not tend to be stable across countries or over time (Bjørnland et al., 2017; Kolasa and Rubaszek, 2015; Nalban, 2018). Consequently, the results on the superiority of certain modelling frameworks when it comes to out-of-sample prediction cannot be easily generalized. The results achieved with the use of US data, for instance, may only be partly relevant for a different country, and similar considerations can be taken with respect to the time frame used, the level of data revisions, the transformations of the data, and many other dimensions of the modelling exercise.

The existing literature also shows that forecasting performance may vary over time (Bjørnland et al., 2017) and can substantially change if real-time data are taken into consideration (Croushore and Stark, 2003).

	models		**data**		
	DSGE	*empirical*	*country*	*real-time*	*variables*
Adolfson et al. (2007)	5	(B)VAR, VECM	euro area	✗	15
Bjørnland et al. (2017)	✗	DFM, AR	33	✗	GDP
Cai et al. (2019)	5	✗	US	✓	GDP, inf.
Carriero et al. (2019)	1	(FA)(B)(V)AR	7	✗	7–14
Clark and Ravazzolo (2014)	✗	(B)(V)AR	US	✓	4
Diebold et al. (2017)	4	✗	US	✓	3
Gürkaynak et al. (2013)	SW	(B)(V)AR, RW	US	✓	3
Kolasa and Rubaszek (2015)	3	✗	US	✗	7
Kolasa et al. (2012)	SW	DSGE-VAR	US	✓	3
Mandalinci (2017)	✗	10	9	✗	inflation
Nalban (2018)	8	✗	Romania	✗	7
Panagiotelis et al. (2019)	✗	(B)(V)AR, DFM	Australia	✗	3
Wolters (2015)	4	BVAR	US	✓	3

Table 1: Classification of features of influential pieces in the macroeconomic forecasting literature. **Notes:** References with *DSGE*=✗ do not use DSGE models, whereas *empirical*=✗ means that the study does not use empirical models. *DSGE=SW* relates to the use of variants of Smets and Wouters (2003, 2007) model. In case of *variables=3*, the study deals with GDP (growth), a measure of inflation, and the interest rate. Literature with *variables=7* also adds consumption, investment, wages, and hours worked.

To summarize the ground covered by some influential pieces in the literature studying the time-varying nature of macroeconomic forecasting, Table 1 categorizes some of the characteristics of the approaches used in macroeconomic forecasting exercises. The second and third columns present information about whether the piece uses *DSGE* and/or *empirical* (econometric) models. The following group of columns shows the *country* or group of countries of interest of the study, if *real-time* data were used and for which macroeconomic *variables* forecasting performance was assessed.

3 Methodology

From the practitioner's perspective and with relation to the topic of this commentary, models used for macroeconomic policy analysis can be categorized into two general classes: Dynamic Stochastic General Equilibrium (DSGE) models and Structural Vector AutoRegressive (SVAR) models.[9]

3.1 Dynamic Stochastic General Equilibrium models[10]

Dynamic Stochastic General Equilibrium (DSGE) models are typically understood as quantitative models of economic growth or business cycle, which are derived from microeconomic foundations of separate economic agents. First models, which can be traced back to e.g. Kydland and Prescott (1982), were following real business cycle (RBC) theory. As the name of the theory suggests, technology shocks were very important in explaining economic fluctuations and there was a limited role of monetary factors. These features made RBC models clearly unsuitable for (monetary) policy practitioners and were also at odds with empirical evidence (Friedman and Schwartz, 1963). These shortages paved way for New Keynesian (NK) models, which featured monopolistic competition, nominal rigidities, and short-run monetary non-neutrality (see review article Clarida et al., 1999). The models in the NK family got new features, which brought them closer to observed economic behavior and whose impulse response functions to various exogenous shocks were more realistic. This group of models can be represented by seminal works Smets and Wouters (2003, 2007).

The Great financial crisis (GFC) of 2007–2008 brought strong stimulus for further development of DSGE models, as the pre-crisis models seemed unable to predict the GFC. Reflecting on this failure, post-crisis models incorporate financial frictions,

[9]This distinction is made for the ease of the exposition in this commentary. From the econometrics point of view, DSGE models can be (under some conditions) understood as VARMA models and subsequently (under further conditions) as infinite or finite VAR models. See Giacomini (2013) for a survey. Also, the class of empirical models is addressed as VAR models, which is not accurate in cases like e.g. non-invertible MA (moving average) representation.

[10]The introduction of this section uses Christiano et al. (2018) and Galí (2008).

among other channels. Additionally, evidence shows that not only does financial frictions block need to be part of the model, but also financial data need to be among the model's observable variables (Christiano et al., 2014; Del Negro et al., 2015; Justiniano et al., 2010). The usefulness of various other model features is being investigated by modern literature, such as model non-linearities (like zero lower bound), heterogeneous agent models, and others.

Model used in Čapek (2014)

For illustrative purposes, this section introduces the log-linearized version of the New Keynesian DSGE model, which was used for estimation and policy analysis in Čapek (2014).

The model is a small open economy (SOE) model, with the Czech economy as the home country and the euro area as the foreign country. The model is adapted from Lubik and Schorfheide (2005).

One of the representative agents in the economy are households, which draw utility from (effective) consumption and disutility from labor. Households can also enter financial markets to bridge the time gap between pay-day and consumption. If we denote consumption at time t as c_t, effective consumption as $\mathscr{c}_t$,[11] marginal utility of real income as λ_t and the growth rate of the world-wide technology shock z_t, the log-linearized evolution of marginal utility of income is[12]

$$-\lambda_t = \frac{\tau}{1-h\beta}\mathscr{c}_t - \frac{h\beta}{1-h\beta}E_t(\tau\mathscr{c}_{t+1} + z_{t+1}), \tag{1}$$

where parameter τ denotes coefficient of relative risk aversion, h is habit (persistence) in consumption, and β is discount factor of future utility. The law of motion of the habit stock is

$$\mathscr{c}_t = \frac{1}{1-h}(c_t - hc_{t-1} + hz_t). \tag{2}$$

Denoting the nominal interest rate r_t and inflation π_t (note that $E_t\pi_{t+1}$ are one-period-

[11] In original model formulation, effective consumption $\mathscr{C}_t$ is defined as $\mathscr{C}_t = C_t - h\gamma C_{t-1}$. Taking log-deviations from the steady state, $c_t = \log C_t - \log C$ and $\mathscr{c}_t = \log \mathscr{C}_t - \log \mathscr{C}$.

[12] Notation convention: variables are letters with a subscript t, $t-1$, or $t+1$, according to the timing. Letters without the timing are parameters. E_t denotes the expectations operator (at time t).

ahead expectations of inflation formed at time t) yields Euler equation

$$-\lambda_t = -E_t\lambda_{t+1} - (r_t - E_t\pi_{t+1}) + E_t z_{t+1}. \tag{3}$$

We can define inflation as a weighted average of domestic ($\pi_{H,t}$) and imported ($\pi_{F,t}$) inflation

$$\pi_t = (1-\alpha)\pi_{H,t} + \alpha\pi_{F,t}, \tag{4}$$

where α is the import share parameter. Production is done by monopolistically competitive firms, which operate in competitive labor markets. Production function features exogenous labor-augmenting home-specific technology progress a_t. Firms face a Calvo-style pricing mechanism, with a fraction of firms $1-\theta_H$ setting price optimally and a fraction of firms θ_H having sticky prices. The log-linearized price-setting decision-making results in a New-Keynesian Phillips curve

$$\pi_{H,t} = \frac{1-\theta_H}{\theta_H}(1-\beta\theta_H)mc_{H,t} + \beta E_t\pi_{H,t+1}, \tag{5}$$

where $mc_{H,t}$ denotes the (domestic) marginal cost, which evolve according to

$$mc_{H,t} = -\alpha q_t - \lambda_t - a_t, \tag{6}$$

where q_t are terms of trade. Similar to producers, importers are also monopolistically competitive. Because importers can sell products with a mark-up, purchasing power parity need not hold in the short run. The result of importers' price-setting is importers' Phillips curve

$$\pi_{F,t} = \frac{1-\theta_F}{\theta_F}(1-\beta\theta_F)\psi_{F,t} + \beta E_t\pi_{F,t+1}. \tag{7}$$

Note the similarities to producers' Phillips curve (5): variables and parameters feature subscript F for goods imported from the foreign economy and $\psi_{F,t}$ represents the law of one price gap.

Foreign economy is modelled structurally and home-economy equations (1), (2), and (5) have analogous versions for the foreign economy:

$$-\lambda_t^* = \frac{\tau}{1-h\beta}\boldsymbol{c}_t^* - \frac{h\beta}{1-h\beta}E_t(\tau\boldsymbol{c}_{t+1}^* + z_{t+1}), \tag{8}$$

$$\boldsymbol{c}_t^* = \frac{1}{1-h}(c_t^* - hc_{t-1}^* + hz_t), \tag{9}$$

$$\pi_t^* = \frac{1-\theta^*}{\theta^*}(1-\beta\theta^*)(-\lambda_t^* - a_t^*) + \beta E_t \pi_{H,t+1}^*. \tag{10}$$

The star superscript (*) denotes foreign economy variables and parameters.

The equations derived from the behavior of households, produces, and importers are complemented with the following definitions and equilibrium conditions:

$$\Delta e_t = \Delta s_t + \pi_t - \pi_t^*, \tag{11}$$

$$q_t = q_{t-1} + \pi_{H,t} - \pi_{F,t}, \tag{12}$$

$$s_t = \psi_{F,t} - (1-\alpha)q_t, \tag{13}$$

$$\lambda_t = \lambda_t^* - s_t, \tag{14}$$

$$r_t - r_t^* = E_t \Delta e_{t+1}, \tag{15}$$

$$y_{H,t} = (1-\alpha)c_t + \alpha c_t^* + \alpha\eta(s_t - q_t) + g_{H,t}. \tag{16}$$

Equation (11) defines the depreciation rate of nominal exchange rate e_t with the use of real exchange rate s_t,[13] equation (12) is differenced version of terms of trade definition, and (13) mutually defines the real exchange rate and the law of one price gap. Equilibrium conditions are international risk-sharing equation (14), uncovered interest parity condition (15), and domestic market clearing condition (16), where $y_{H,t}$ represents domestic output, $g_{H,t}$ government expenditures and parameter η is intratemporal elasticity of substitution between home and imported consumption goods. Equilibrium conditions are completed with the market clearing condition for the foreign economy

$$y_t^* = c_t^* + g_t^*, \tag{17}$$

which is simpler than its domestic counterpart (16), because the foreign economy is a large open economy, which is, by definition, not influenced by a small open domestic economy.

The model is closed by specifying monetary policy with a Taylor-type interest rate

[13] Δ denotes the difference operator such that $\Delta e_t = e_t - e_{t-1}$.

rule for domestic and foreign economies

$$r_t = \rho_r r_{t-1} + (1-\rho_r)[\psi_1 \pi_t + \psi_2(\Delta y_{H,t} + z_t) + \psi_3 \Delta e_t] + \varepsilon_{r,t}, \tag{18}$$

$$r_t^* = \rho_r^* r_{t-1}^* + (1-\rho_r^*)[\psi_1^* \pi_t^* + \psi_2^*(\Delta y_t^* + z_t)] + \varepsilon_{r,t}^*, \tag{19}$$

where ρ_rs are backward-looking parameters and ψs are monetary policy preference parameters, which are at the core of investigation in Čapek (2014). These parameters show the weights that monetary policy places on different sources of instability as it reacts to the deviations of the variables from the steady state. Innovation $\varepsilon_{r,t}$ captures non-systematic parts of monetary policy.

When the model is taken to the data, it is supplemented with autoregressive AR(1) processes for domestic and foreign technology progress (a_t, a_t^*), government expenditures $(g_{H,t}, g_t^*)$, and the growth rate of the world-wide technology shock z_t.

The log-linearized version of the model consists of 24 equations, 7 exogenous processes (model innovations), and 20 parameters.

3.2 Structural Vector AutoRegressive models

We can nest the set of empirical models used to estimate policy effects in the stacked form of a dynamic factor model, following Stock and Watson (2016). A set of q dynamic factors are stacked to yield r static factors in the vector F_t and, abstracting from further deterministic terms, a factor-augmented VAR (FAVAR) structure is be given by

$$\begin{pmatrix} \underset{n\times 1}{Y_t} \\ \\ \underset{m\times 1}{X_t} \end{pmatrix} = \begin{pmatrix} \underset{n\times n}{\mathbf{I}} & \underset{n\times r}{\mathbf{0}} \\ \\ \underset{m\times n}{\mathbf{\Lambda}^Y} & \underset{m\times r}{\mathbf{\Lambda}^F} \end{pmatrix} \begin{pmatrix} \underset{n\times 1}{\tilde{F}_t} \\ \\ \underset{r\times 1}{F_t} \end{pmatrix} + \begin{pmatrix} \underset{n\times 1}{\mathbf{0}} \\ \\ \underset{m\times 1}{e_t} \end{pmatrix} \tag{20}$$

$$\underset{(n+r)\times(n+r)}{\mathbf{\Phi}(L)} \begin{pmatrix} \underset{n\times 1}{\tilde{F}_t} \\ \\ \underset{r\times 1}{F_t} \end{pmatrix} = \begin{pmatrix} \underset{(n+q)\times(n+q)}{\mathbf{I}} \\ \\ \underset{(r-q)\times(n+q)}{\mathbf{0}} \end{pmatrix} \underset{(n+q)\times 1}{\eta_t} \tag{21}$$

$$\underset{(n+q)\times(n+q)}{\mathbf{A}} \underset{(n+q)\times 1}{\eta_t} = \underset{(n+q)\times(n+q)}{\mathbf{B}} \underset{(n+q)\times 1}{\varepsilon_t} \tag{22}$$

where equation (20) is the measurement equation, equation (21) is the transition equation, and equation (22) is the identification equation, while the (matrix) lag polynomial $\mathbf{\Phi}(L)$ is given by $\mathbf{\Phi}(L) = \mathbf{I} - \mathbf{\Phi}_1 L - \cdots - \mathbf{\Phi}_p L^p$ for matrices $\mathbf{\Phi}_l$, $l = 1, \ldots, p$. The variables in Y_t are assumed to be measured without error by the observed factors $\tilde{F}_t$. X_t contains m observed time series (not contained in Y_t) summarizing information about other macroeconomic phenomena. Variables in X_t are assumed to depend on observed factors $\tilde{F}_t$, unobserved factors F_t and an idiosyncratic component e_t, with matrix $\mathbf{\Lambda}^F$ comprising the corresponding factor loadings. Equation (22) specifies the relationship between reduced-form (η_t) and structural shocks (ε_t). If the number of unobserved factors r is set to zero, the model collapses to a standard SVAR model which can be utilized to implement the methods in Blanchard and Perotti (2002) or Perotti (2004) for structural shock identification. The unobserved factors of the model (F_t) are estimated as principal components and the identification of the model is reached once matrices $\mathbf{A}$ and $\mathbf{B}$ are chosen (see Stock and Watson, 2016).

Various identification methods can be used to retrieve the structural shocks in ε_t. The method pioneered by Blanchard and Perotti (2002) relies on exact restrictions imposed on the error terms of a VAR model which includes GDP, government expenditure and taxes through an identification scheme based on lags in the implementation of fiscal policy. More modern methods (Rubio-Ramírez et al., 2010) use sign restrictions that constrain the direction of the response of variables to particular shocks.

In case of application of the methodology to fiscal policy analysis, once the structural shocks have been identified, government spending and tax multipliers can be computed. In line with recent literature (e.g. Caggiano et al., 2015; Gechert and Rannenberg, 2014; Ilzetzki et al., 2013; Mountford and Uhlig, 2009), present-value (or discounted cumulative) multipliers at lag T can be calculated as

$$\text{present-value spending multiplier} = \frac{\sum_{t=0}^{T}(1+i)^{-t} y_t}{\sum_{t=0}^{T}(1+i)^{-t} g_t} \frac{1}{g/y}, \tag{23}$$

where y_t is the response of output at time t (in logs), g_t denotes the response of government expenditures at time t (in logs) and g/y is the average share of government expenditures in GDP over the sample. The multiplier is discounted with the interest rate i. The tax multiplier is calculated analogously, after substituting government expenditures in equation (23) with taxes.

If we concentrate on the non-cumulative reaction of GDP, such effects can be summarized using the so-called peak multipliers (see e.g. Blanchard and Perotti, 2002;

Caggiano et al., 2015; Fragetta and Gasteiger, 2014; Ramey, 2011b),

$$\text{peak spending multiplier} = \frac{\max_{t=0,\ldots,H}\{y_t\}}{\max_{t=0,\ldots,H}\{g_t\}}\frac{1}{g/y}. \tag{24}$$

4 Contribution

In all articles, specific macroeconomic policies have been investigated with the use of a comprehensive framework, which was in all cases conducted by the applicant.

4.1 Čapek (2014)

The most notable contribution is that, unlike all literature cited in the article, separate vintages of data are used for the estimation, not just a single series of real-time data. Such an approach should better mimic the time series that are actually available to the decision-maker. Note that the ”traditional” approach requires just two estimates: one with the time series of ”real-time data” and the other with the time series with the fully revised data. In contrast, using actual vintages requires re-estimation for each period under investigation. For example, in the model with the Czech economy, 40 separate estimates were needed in order to calculate the baseline results.

A further contribution is the use of a comprehensive framework to communicate the uncertainty of the results. Specifically, the revision is split into data revision and trend revision, as both can affect the results, but both have different origins and consequences. The analysis covers methodological approaches (recursive or rolling estimates), data variations (CPI seasonally adjusted or not), or model changes (Taylor rule contains quarter-on-quarter or year-on-year GDP growth). In order to fully communicate the uncertainty, the whole posterior distribution of the parameter, which resulted from Bayesian estimation of the model, is used for reporting (the measure used for reporting is the lowest level of significance at which the posterior mode is out of the Highest Posterior Density interval bands for the most different estimates). By offering all these options and variants, the reader can get a complex understanding of the presented results.

Compared to the literature, Čapek (2014) uses a more complex DSGE model. Early literature, e.g. Orphanides (2001), uses just one equation of the Taylor rule. Later literature, e.g. Neri and Ropele (2012), already uses a DSGE model, albeit rather stylized. Finally, an empirical contribution lies in delivering the results for CEE (Central and Eastern European) economies, which were not previously available.

4.2 Čapek and Crespo Cuaresma (2020)

The article aims to assess how the size and precision of fiscal multipliers obtained using structural VAR (SVAR) models depend on the different methodological choices that need to be made when specifying them. The main contribution of the article stems from conducting a comprehensive analysis and presenting the results with the use of a meta-regression.

Systematic analysis of the role played by data composition, data transformations, the methodology of fiscal data collection or the specific formulation of the reduced-form VAR model was absent in the existing literature on fiscal multipliers, either in meta-analysis pieces like Gechert (2015) and Rusnák (2011), or in more systematic empirical approaches like Caldara and Kamps (2008). There are several reasons for the missing evidence. From the meta-analysis perspective, there are so many possible combinations of these characteristics that there are simply not enough studies yet to have been able to cover the variability needed to identify their effects on the estimates of fiscal multipliers. In addition, from the perspective of the practitioner, some of these characteristics are often considered innocuous and do not tend to be reported in the published pieces.

Article's results indicate that many seemingly inconsequential choices affect the value of the estimated multipliers as well as the precision with which they are estimated. An implication of the analysis is that, when structural VARs are used to estimate fiscal multipliers, it is important for researchers to document their choices in detail, even for aspects of the research design that may seem innocuous.

Author contribution statement[14]

- Conception or design of the work - Jan Čapek (general paper's idea; SVAR models estimation and identification), Jesús Crespo Cuaresma (Bayesian averaging, weighted least squares for meta-regression)
- Data collection - Jan Čapek

[14]For both co-authored articles in the collection, the author contribution statements are authorized by all co-authors. Separate paper sheets with handwritten signatures are included in the habilitation file.

- Computation of the results - Jan Čapek
- Data analysis and interpretation - Jan Čapek
- First drafting the article - Jan Čapek
- Revisions of the article - Jan Čapek, Jesús Crespo Cuaresma

4.3 Čapek et al. (2021)

There is little evidence on the size of fiscal multipliers for developed European small open economies. Following the literature review (see section 2.3), Austria was selected as the country for the study. A pure empirical assessment of fiscal multipliers specifically for Austria, as a stereotypical small open economy within the group of industrialized countries, did not exist. In this contribution, we provide for the first time a rigorous analysis of fiscal multiplier estimates in a small open economy (Austria) incorporating the uncertainty related to specification choice in several dimensions including that related to the particular variables included in the model, shock identification strategies, data preparation or the analytical structure of the model. Given the importance of economic openness to determine the size of the fiscal multiplier, such an exercise allows the results to be interpreted in the framework of theoretical models of fiscal policy effects in small open economy settings.

The approach of this study is to present a consistent framework that encompasses a wide range of reasonable settings and choices which are routinely used in the empirical literature on fiscal multipliers.

The analysis expands the methodological setting put forward in Čapek and Crespo Cuaresma (2020) in several respects. First of all, by concentrating on a single economy, we gain comparability in the multiplier estimates, which correspond to the responses to fiscal impulses within the same institutional and historical setting. Furthermore, we expand the set of econometric specifications and modelling choices in Čapek and Crespo Cuaresma (2020) by including new models based on factor-augmented VAR structures and using out-of-sample predictive ability as a model selection tool. The focus on a single small open economy allows us to link the results in a more direct manner to the methodological framework provided by economic theory, in particular when interpreting the results of the analysis, and allows for the assessment of additional sources of model uncertainty as compared to Čapek and Crespo Cuaresma

(2020). This is the case, for example, for the composition of government spending and tax aggregates, or for the calculation of the values of tax and spending elasticities required for several identification techniques. In our analysis, we also contribute to the literature by identifying structural fiscal shocks in models where subcomponents of spending and tax revenues are used, making use of elasticities of disaggregated components of the fiscal variables to output and the price level obtained using the fiscal forecasting model by the Austrian Fiscal Advisory Council (2014).

Author contribution statement

- Conception or design of the work - Jan Čapek (structural (FA)VAR models estimation and identification), Jesús Crespo Cuaresma (forecasting performance), Johannes Holler & Philip Schuster (several spending and revenues compositions; spending and tax composition-specific elasticities)
- Data collection - Jan Čapek
- Computation of the results - Jan Čapek (structural (FA)VAR models estimation and identification), Johannes Holler & Philip Schuster (calculation of composition-specific elasticities with the fiscal forecasting model of the Austrian Fiscal Advisory Council)
- Data analysis and interpretation - Jan Čapek
- First drafting the article - Jan Čapek
- Revisions of the article - Jan Čapek, Jesús Crespo Cuaresma, Johannes Holler, Philip Schuster

5 Discussion

5.1 Limitations of the research

There are many limits to the research presented in the three submitted articles. For example, the universality of the findings in Čapek (2014) could have been broadened in case that more DSGE models were estimated. Current analysis results do not make it clear if the change in the preferences in real-time settings is not a feature of the specific DSGE model used.

In an ideal case, the results in Čapek and Crespo Cuaresma (2020) could be more relevant if we estimated more types of empirical models, like Markov-switching models following e.g. Auerbach and Gorodnichenko (2012), factor-augmented models following as in Čapek et al. (2021), or various other specifications. Also, it would be interesting to see similar results not only using empirical models, but also with the use of DSGE models. However, it is important to note that in case of Čapek and Crespo Cuaresma (2020), the computational burden is quite high already. As the title puts it, ”We just estimated twenty million fiscal multipliers”. Considering additionally just one of the above-mentioned variants would make it double.

In contrast, the analysis in Čapek et al. (2021) did not face big computational restrictions and results for more model and data specifications are therefore reported. Admittedly, the contribution is limited to the case of a single country, but the idea of the article is to deliver a comprehensive framework tailored to the economy in question. As such, it would not be useful to re-estimate *the same* specification for a different country.

5.2 Further research

There are many possible avenues of further research. One can perhaps start in the previous section and address the limitations of the research in the articles. In the following two subsections, I am going to focus on further research, which is attempted, or underway.

Fiscal multipliers

Following the results of the research in the area of fiscal multipliers, the Czech Fiscal Council became interested in collaboration on estimating the multipliers in a comprehensive framework, similar to the analysis conducted in Čapek et al. (2021). To that end, a project called ”Estimation of fiscal multipliers for the Czech economy” has been submitted to the Technology Agency of the Czech Republic (TACR) in the call Éta 5. The goals of the project were stated as:

> The current situation of accelerating government debt increases the need to inform the public about the risks of recurrent and high budget deficits. The Czech Fiscal Council reports about the management of public institutions. The key to assessing fiscal policies' impact is knowledge of fiscal multipliers,

> the systematic estimate of which for the Czech economy does not yet exist. The project aims to use econometric methods to estimate fiscal multipliers for the Czech economy, which will then be used by the application guarantor (Office of the Cz. Fisc. Council) to create analyzes of compliance with budgetary responsibility rules. The societal impact lies in improving the quality of information on public finances' sustainability, which is currently an important and widely discussed topic.

Despite a positive evaluation of both reviewers and the project rapporteur, the project did not receive funding. However, the interest of the Czech Fiscal Council still persists and the project may be resubmitted in some future calls of TACR focused on applied research in social sciences.

Macroeconomic forecasting

As the literature review in section 2.4 shows, macroeconomic forecasting literature brings interesting results for specific research questions, but in some cases lacks comparability. This opens the possibility of a research project aimed at delivering a comprehensive framework in macroeconomic forecasting.

Regarding the selection of the literature presented in Table 1 (page 12), the research proposal mainly relates to the work by Gürkaynak et al. (2013), which focuses on investigating the effects of considering different sub-samples on the predictive ability of macroeconomic models using real-time data, as well as to Nalban (2018), who analyses the forecasting power of a battery of DSGE models, and Mandalinci (2017), who offers a multi-country approach and compares many empirical model specifications. However, the results of these three selected articles are not readily comparable. Mandalinci (2017) entertains only empirical models, whereas Nalban (2018) concentrates only on DSGE specifications. Also, whereas Mandalinci (2017) offers the results for 9 countries, both Gürkaynak et al. (2013) and Nalban (2018) focus on one country only. In order to overcome this lack of comparability and to provide a shift of the research frontier, the project strives to run a comprehensive selection of modern methods and approaches on a battery of countries to allow for as broad comparability of the results as possible.

The goal of the project is to conduct a systematic investigation of the stability of the out-of-sample forecasting performance of theoretical and empirical models over time in order to reach conclusions about their applicability. The contribution lies primarily in the systematic approach, which allows for broad comparability, a feature that is missing in the current literature. In addition, new econometric models, whose predictive ability has not yet been explicitly compared to DSGE models, will be entertained in the research project. Systematic evaluation of forecasting performance can also help to identify model specifications that allow some models to perform better than others. In the case of structural models, these findings can unveil reasons for the good or bad forecast performance of used models.

The project has been submitted to Czech Science Foundation in a standard call and has been granted funding. Currently, the project is in its first year.

References

Adolfson, M., Lindé, J., & Villani, M. (2007). Forecasting Performance of an Open Economy DSGE Model. *Econometric Reviews*, 26(2-4):289–328.

Afonso, A. & Sousa, R. M. (2011). The macroeconomic effects of fiscal policy in Portugal: a Bayesian SVAR analysis. *Portuguese Economic Journal*, 10(1):61–82.

Aikman, D., Barrett, P., Kapadia, S., King, M., Proudman, J., Taylor, T., De Weymarn, I., & Yates, T. (2011). Uncertainty in macroeconomic policy-making: art or science? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1956):4798–4817.

Angrist, J. D. & Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives*, 24(2):3–30.

Aruoba, S. B. (2008). Data revisions are not well behaved. *Journal of money, credit and banking*, 40(2-3):319–340.

Auerbach, A. J. & Gorodnichenko, Y. (2012). Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy*, 4(2):1–27.

Austrian Fiscal Advisory Council (2014). Fiscal Forecasts of the Austrian Fiscal Advisory Council: Forecasting Principles and Methods.

Barrell, R., Holland, D., & Hurst, I. (2012). *Fiscal Consolidation: Part 2. Fiscal Multipliers and Fiscal Consolidations,* (OECD working papers ed.). Organisation for Economic Co-Operation and Development (OECD).

Beetsma, R. & Giuliodori, M. (2011). The Effects of Government Purchases Shocks: Review and Estimates for the EU. *The Economic Journal*, 121(550):F4–F32.

Belke, A. & Klose, J. (2011). Does the ECB rely on a Taylor rule during the financial crisis? Comparing ex-post and real time data with real time forecasts. *Economic analysis and policy*, 41(2):147–171.

Benetrix, A. S. & Lane, P. R. (2009). The impact of fiscal shocks on the Irish economy. *The Economic and Social Review*, 40(4):407.

Bjørnland, H. C., Ravazzolo, F., & Thorsrud, L. A. (2017). Forecasting GDP with global components: This time is different. *International Journal of Forecasting*, 33(1):153–173.

Blanchard, O. & Perotti, R. (2002). An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *The Quarterly Journal of Economics*, 117(4):1329–1368.

Blanchard, O. J. & Leigh, D. (2013). Growth forecast errors and fiscal multipliers. *The American Economic Review*, 103(3):117–120.

Breuss, F., Kaniovski, S., & Schratzenstaller, M. (2009). Macro-economic effects of the fiscal stimulus measures in Austria. *Austrian Economic Quarterly*, 14(4):205–216.

Caggiano, G., Castelnuovo, E., Colombo, V., & Nodari, G. (2015). Estimating Fiscal Multipliers: News From A Non-linear World. *The Economic Journal*, 125(584):746–776.

Cai, M., Del Negro, M., Giannoni, M. P., Gupta, A., Li, P., & Moszkowski, E. (2019). DSGE forecasts of the lost recovery. *International Journal of Forecasting*, 35(4):1770 – 1789.

Caldara, D. & Kamps, C. (2008). What are the effects of fiscal shocks? A VAR-based comparative analysis. Technical report, ECB Working Paper No. 877.

Caldara, D. & Kamps, C. (2017). The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers. *The Review of Economic Studies*, 84(3):1015–1040.

Čapek, J. (2014). Historical Analysis of Monetary Policy Reaction Functions: Do Real-Time Data Matter? *Finance a Uver: Czech Journal of Economics & Finance*, 64(6).

Čapek, J. & Crespo Cuaresma, J. (2020). We just estimated twenty million fiscal multipliers. *Oxford Bulletin of Economics and Statistics*, 82(3):483–502.

Čapek, J., Crespo Cuaresma, J., Holler, J., & Schuster, P. (2021). Fiscal multipliers in a small open economy: the case of Austria. *Oxford Economic Papers*. gpab027.

Carriero, A., Galvao, A. B., & Kapetanios, G. (2019). A comprehensive evaluation of macroeconomic forecasting methods. *International Journal of Forecasting*, 35(4):1226–1239.

Casares, M. & Vázquez, J. (2016). Data revisions in the estimation of DSGE models. *Macroeconomic Dynamics*, 20(7):1683–1716.

Christiano, L. J., Eichenbaum, M. S., & Trabandt, M. (2018). On DSGE Models. *Journal of Economic Perspectives*, 32(3):113–40.

Christiano, L. J., Motto, R., & Rostagno, M. (2014). Risk shocks. *American Economic Review*, 104(1):27–65.

Cimadomo, J. (2016). Real-time data and fiscal policy analysis: A survey of the literature. *Journal of economic Surveys*, 30(2):302–326.

Clarida, R., Gali, J., & Gertler, M. (1999). The science of monetary policy: a new Keynesian perspective. *Journal of economic literature*, 37(4):1661–1707.

Clark, T. E. & Ravazzolo, F. (2014). Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Croushore, D. (2011). Frontiers of real-time data analysis. *Journal of economic literature*, 49(1):72–100.

Croushore, D. & Stark, T. (2001). A real-time data set for macroeconomists. *Journal of econometrics*, 105(1):111–130.

Croushore, D. & Stark, T. (2003). A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter? *Review of Economics and Statistics*, 85(3):605–617.

Cusinato, R. T., Minella, A., & Júnior, S. d. S. P. (2013). Output gap in Brazil: a real-time data analysis. *Empirical Economics*, 44(3):1113–1127.

Del Negro, M., Giannoni, M. P., & Schorfheide, F. (2015). Inflation in the great recession and new keynesian models. *American Economic Journal: Macroeconomics*, 7(1):168–96.

Del Negro, M. & Schorfheide, F. (2013). DSGE model-based forecasting. In: *Handbook of economic forecasting*, volume 2, pages 57–140. Elsevier.

Diebold, F. X., Schorfheide, F., & Shin, M. (2017). Real-time forecast evaluation of DSGE models with stochastic volatility. *Journal of Econometrics*, 201(2):322 – 332.

Ehrlich, I. (1975). The Deterrent Effect of Capital Punishment: A Question of Life and Death. *The American Economic Review*, 65(3):397–417.

Fragetta, M. & Gasteiger, E. (2014). Fiscal Foresight, Limited Information and the Effects of Government Spending Shocks. *Oxford Bulletin of Economics and Statistics*, 76(5):667–692.

Friedman, M. & Schwartz, A. J. (1963). *A monetary history of the United States, 1867-1960*. Princeton University Press.

Galí, J. (2008). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.

Gechert, S. (2015). What fiscal policy is most effective? A meta-regression analysis. *Oxford Economic Papers*, 67(3):553–580.

Gechert, S. & Rannenberg, A. (2014). Are Fiscal Multipliers Regime-Dependent? A Meta Regression Analysis. *IMK Working Paper*, (139).

Gerberding, C., Seitz, F., & Worms, A. (2005). How the Bundesbank really conducted monetary policy. *The North American journal of economics and finance*, 16(3):277–292.

Gerdesmeier, D. & Roffia, B. (2005). The relevance of real-time data in estimating reaction functions for the euro area. *The North American Journal of Economics and Finance*, 16(3):293–307.

Giacomini, R. (2013). The relationship between DSGE and VAR models. *VAR Models in Macroeconomics–New Developments and Applications: Essays in Honor of Christopher A. Sims*.

Giannone, D., Henry, J., Lalik, M., & Modugno, M. (2012). An area-wide real-time database for the euro area. *Review of Economics and Statistics*, 94(4):1000–1013.

Giovannelli, A. & Pericoli, F. M. (2020). Are GDP forecasts optimal? Evidence on European countries. *International Journal of Forecasting*, 36(3):963–973.

Gürkaynak, R. S., Kısacıkoğlu, B., & Rossi, B. (2013). Do DSGE Models Forecast More Accurately Out-Of-Sample than VAR Models? In: *VAR Models in Macroeconomics–New Developments and Applications: Essays in Honor of Christopher A. Sims*, pages 27–79. Emerald Group Publishing Limited.

Hebous, S. (2011). The effects of discretionary fiscal policy on macroeconomic aggregates: a reappraisal. *Journal of Economic Surveys*, 25(4):674–707.

Horváth, R. (2009). The time-varying policy neutral rate in real-time: A predictor for future inflation? *Economic Modelling*, 26(1):71–81.

Hullman, J. (2019). Why authors don't visualize uncertainty. *IEEE transactions on visualization and computer graphics*, 26(1):130–139.

Ilzetzki, E., Mendoza, E. G., & Végh, C. A. (2013). How big (small?) are fiscal multipliers? *Journal of Monetary Economics*, 60(2):239–254.

Ince, O. & Papell, D. H. (2013). The (un) reliability of real-time output gap estimates with revised data. *Economic Modelling*, 33:713–721.

Jemec, N., Kastelec, A., & Delakorda, A. (2011). *How Do Fiscal Shocks Affect the Macroeconomic Dynamics in the Slovenian Economy*. Prikazi in analize. Banka Slovenije.

Justiniano, A., Primiceri, G. E., & Tambalotti, A. (2010). Investment shocks and business cycles. *Journal of Monetary Economics*, 57(2):132–145.

Kilponen, J., Pisani, M., Schmidt, S., Corbo, V., Hledik, T., Hollmayr, J., Hurtado, S., Júlio, P., Kulikov, D., Lemoine, M., et al. (2015). Comparing fiscal multipliers across models and countries in Europe. Technical report.

Koch, S., Monárová, Z., & Reiter, M. (2019). Macro-economic effects of the fiscal stimulus measures in Austria. *Mimeo, Institute for Advanced Studies, Vienna*.

Kolasa, M. & Rubaszek, M. (2015). Forecasting using DSGE models with financial frictions. *International Journal of Forecasting*, 31(1):1–19.

Kolasa, M., Rubaszek, M., & Skrzypczyński, P. (2012). Putting the New Keynesian DSGE Model to the Real-Time Forecasting Test. *Journal of Money, Credit and Banking*, 44(7):1301–1324.

Kydland, F. E. & Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370.

Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review*, 73(1):31–43.

Leamer, E. E. (1985). Sensitivity analyses would help. *The American Economic Review*, 75(3):308–313.

Lindé, J. (2018). DSGE models: still useful in policy analysis? *Oxford Review of Economic Policy*, 34(1-2):269–286.

Lubik, T. & Schorfheide, F. (2005). A Bayesian look at new open economy macroeconomics. *NBER macroeconomics annual*, 20:313–366.

Mandalinci, Z. (2017). Forecasting inflation in emerging markets: An evaluation of alternative models. *International Journal of Forecasting*, 33(4):1082–1104.

Manski, C. F. (2019). Communicating uncertainty in policy analysis. *Proceedings of the National Academy of Sciences*, 116(16):7634–7641.

Maravall, A. & Pierce, D. A. (1986). The transmission of data noise into policy noise in US monetary control. *Econometrica: Journal of the Econometric Society*, pages 961–979.

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

McKenzie, R. (2006). Undertaking revisions and real-time data analysis using the OECD main economic indicators original release data and revisions database.

Meirelles Aurelio, M. (2005). Do we really know how inflation targeters set interest rates? *FRB of Kansas City Working Paper No. RWP0502*.

Mineshima, A., Poplawski-Ribeiro, M., & Weber, A. (2014). Size of fiscal multipliers. In: *Post-Crisis Fiscal Policy*, pages 315–372. MIT Press.

Morgenstern, O. (1950). *On the accuracy of economic observations*. Princeton, NJ (USA) Princeton Univ. Press.

Morgenstern, O. (1963). *On the accuracy of economic observations*. Princeton, NJ (USA) Princeton Univ. Press. 2nd Ed.

Mountford, A. & Uhlig, H. (2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*, 24(6):960–992.

Nalban, V. (2018). Forecasting with DSGE models: What frictions are important? *Economic Modelling*, 68:190 – 204.

Neri, S. & Ropele, T. (2012). Imperfect Information, Real-Time Data and Monetary Policy in the Euro Area. *The Economic Journal*, 122(561):651–674.

Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review*, 91(4):964–985.

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

Panagiotelis, A., Athanasopoulos, G., Hyndman, R. J., Jiang, B., & Vahid, F. (2019). Macroeconomic forecasting for Australia using a large number of predictors. *International Journal of Forecasting*, 35(2):616–633.

Perotti, R. (2004). Estimating the effects of fiscal policy in OECD countries. Technical Report December, IGIER Working Paper No. 276.

Ramey, V. A. (2011a). Can Government Purchases Stimulate the Economy? *Journal of Economic Literature*, 49(3):673–685.

Ramey, V. A. (2011b). Identifying government spending shocks: It's all in the timing. *The Quarterly Journal of Economics*, 126(1):1–50.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. In: *Handbook of macroeconomics*, volume 2, pages 71–162. Elsevier.

Ramey, V. A. (2019). Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? *Journal of Economic Perspectives*, 33(2):89–114.

Ravn, S. H. & Spange, M. (2012). The effects of fiscal policy in a small open economy with a fixed exchange rate: The case of Denmark. Technical report, Danmarks Nationalbank Working Papers.

Rubio-Ramírez, J. F., Waggoner, D. F., & Zha, T. (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *The Review of Economic Studies*, 77(2):665–696.

Rusnák, M. (2011). Why Do Government Spending Multipliers Differ? A Meta-Analysis.

Rusnák, M. (2013). Revisions to the Czech National Accounts: Properties and Predictability. *Finance a Uver*, 63(3):244.

Sala-i Martin, X. X. (1997). I just ran two million regressions. *The American Economic Review*, 87(2):178–183.

Schuster, P. (2019). On fiscal multipliers in New Keynesian small open economy models. *Mimeo*.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, pages 1–48.

Smets, F. & Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European economic association*, 1(5):1123–1175.

Smets, F. & Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American economic review*, 97(3):586–606.

Stock, J. H. & Watson, M. W. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In: *Handbook of macroeconomics*, volume 2, pages 415–525. Elsevier.

Van der Bles, A. M., Van Der Linden, S., Freeman, A. L., Mitchell, J., Galvao, A. B., Zaval, L., & Spiegelhalter, D. J. (2019). Communicating uncertainty about facts, numbers and science. *Royal Society open science*, 6(5):181870.

Vázquez, J., María-Dolores, R., & Londoño, J. M. (2010). Data Revisions and the Monetary Policy Rule: An analysis based on an extension of the basic New Keynesian model. *The University of the Basque Country, Mimeo*.

Wolters, M. H. (2015). Evaluating Point and Density Forecasts of DSGE Models. *Journal of Applied Econometrics*, 30(1):74–96.

List of supplements

1. Essay 1: Čapek (2014): Čapek, J. (2014). Historical Analysis of Monetary Policy Reaction Functions: Do Real-Time Data Matter? *Finance a Uver: Czech Journal of Economics & Finance*, 64(6), http://journal.fsv.cuni.cz/mag/article/show/id/1311

2. Online Appendix to Essay 1: Čapek (2014)

3. Essay 2: Čapek and Crespo Cuaresma (2020): Čapek, J. & Crespo Cuaresma, J. (2020). We just estimated twenty million fiscal multipliers. *Oxford Bulletin of Economics and Statistics*, 82(3):483–502. https://onlinelibrary.wiley.com/doi/full/10.1111/obes.12351

4. Online Appendix to Essay 2: Čapek and Crespo Cuaresma (2020)

5. Essay 3 (**published** version): Čapek et al. (2021): Čapek, J., Crespo Cuaresma, J., Holler, J., & Schuster, P. (2021). Fiscal multipliers in a small open economy: the case of Austria. *Oxford Economic Papers*. gpab027, https://doi.org/10.1093/oep/gpab027

6. Online Appendix to Essay 3: Čapek et al. (2021)

7. Essay 3 (**accepted** version): Čapek, J., Crespo Cuaresma, J., Holler, J., & Schuster, P. (2021). Fiscal multipliers in a small open economy: the case of Austria. Accepted in *Oxford Economic Papers*.

The **published** version of Essay 3 contains errors (the text is from the next-to-last step in the revision process in the journal while the majority of figures are from the accepted version).[15] After consultation with the Office for Science, Research and Doctoral Studies (and respective office at University level), I also enclose the **accepted** version of the article as a part of the commentary to the habilitation thesis, which is included as a seventh supplement.

[15] The process of issuing a corrigendum is underway with the publisher.

JEL Classification: C11, E58, C82
Keywords: real-time data, revision, DSGE model, Bayesian estimation, recursive estimation

Historical Analysis of Monetary Policy Reaction Functions: Do Real-Time Data Matter?*

Jan ČAPEK—Faculty of Economics and Administration, Masaryk University, Brno, Czech Republic (capek@econ.muni.cz)

Abstract
This paper investigates the differences between parameter estimates of monetary policy reaction functions using real-time data and those using revised data. The model is a New Keynesian DSGE model of the Czech, Hungarian and Polish small open economies in interaction with the euro area. Unlike the related literature, this paper uses separate vintages of real-time data for all successive estimations. The paper reports several statistically significant differences between parameter estimates of monetary policy reaction functions based on real-time data and those based on revised data. The parameter whose estimate is the most affected by the usage of real-time data is preference for output growth. This result is common across the countries in the study. The results suggest that real-time data matter when conducting a historical analysis of monetary policy preferences.

1. Introduction

The unreliability of real-time macroeconomic data is a well-known issue and many studies have investigated the properties of data revisions. Orphanides and van Norden (2002) report that revision of U.S. published data is not the main issue; it is the unreliability of end-of-sample trend estimates. These results are confirmed by Marcelino and Musso (2011) on euro area data, and by Ince and Papell (2013) on data for nine OECD countries. On the other hand, Cusinato *et al.* (2013) find that data revision and the end-of-sample problem contribute to uncertainty about the Brazilian output gap, but do not find any evidence that the former is less important than the latter. Investigating the empirical properties of U.S. macroeconomic data, Aruoba (2008) finds that the revisions are biased and predictable. Rusnák (2013) reports that revisions of Czech GDP and its components are rather large. He also studies whether the revisions are "news" or "noise", i.e. whether the revisions are predictable or unpredictable, and ascertains in-sample predictability and out-of-sample unpredictability for most variables of interest.

Clearly, there are many possible problems with data revisions, and they are quantitatively of variable importance in different countries. However, the fact that the data revisions are large does not necessarily mean that they must create problems for economic agents. A researcher who wants to find out how significant the revisions are for decision-making in different situations must incorporate real-time data

* I am grateful to two anonymous referees and Roman Horváth for their helpful comments and suggestions. Computational resources were provided by the MetaCentrum under the program LM2010005 and the CERIT-SC under the program Centre CERIT Scientific Cloud, part of the Operational Program Research and Development for Innovations, Reg. no. CZ.1.05/3.2.00/08.0144.
A companion paper investigates the differences between the remaining parameter estimates (not pertaining to monetary policy) using real-time data, and those using revised data, on a bigger set of countries. See Čapek (2015).

or data revisions into the decision-making process and observe if there are noteworthy differences in the results. A great deal of research effort has focused on the analytic consequences of bad real-time data quality for monetary policy. The literature goes back to Maravall and Pierce (1986), who investigate the conduction of monetary policy in real-time and ask whether policy would have been different if final data had been available. The authors conclude that the answer to that question is no. A similar research question is also investigated by a stream of literature that focuses on monetary rule under real-time conditions: Orphanides (2001) uses a Taylor-type rule to look at the effect of different policy recommendations using real-time data compared to final data and argues that monetary policy reaction functions estimated on final data provide a misleading description of historical policy. Similar results, i.e. that real-time data play a (significant) role, were also reached by Aurelio (2005), Gerdesmeier and Roffia (2005), Gerberding *et al.* (2005), Horváth (2009), and Belke and Klose (2011). More recent literature also uses DSGE models with monetary rules as a tool for monetary policy investigation. Vázquez *et al.* (2010) and Casares and Vázquez (2012) find that monetary policy parameters are robust to real-time specification. On the other hand, Neri and Ropele (2012) show that there is indeed a statistically significant difference in policy parameters when real-time data are considered (rather than final data). Interested readers are referred to Croushore (2011) for an extensive survey of the real-time data literature.

This paper follows the aforementioned literature in exploring the analytic consequences of using real-time versus revised data for monetary policy decision-making. Prior to the main analysis, it may be interesting to look briefly at the statistical properties of data revision that may play a role in the main analysis. Section 3 of the paper investigates the statistical properties of data revisions for Czech, Hungarian and Polish GDP growth and inflation. The data revision analysis is followed (in Section 4) by the main part of the paper, which is an investigation of the influence of using real-time data for a historical analysis of Czech, Hungarian and Polish monetary policies. Theoretically speaking, using the most recent revised data for ex-post analyses on a historical time-sample may be misleading because revised data were not available at that time. This study uses a small-scale monetary macroeconomic DSGE model to analyze the importance of real-time data, focusing on the differences in implied decision-making by the monetary authority. In the model, the monetary authority's decision-making is approximated by a Taylor-type monetary rule, whose parameters can be interpreted as the monetary authority's preferences. Therefore, the operational goal of the paper is to investigate the differences between parameter estimates of monetary policy reaction functions using real-time data and using revised data.

The analysis proceeds from a Bayesian estimation of model parameters and its results are also presented in terms of the statistical significance of the differences in parameter estimates. Note that, unlike all cited literature, this paper uses separate vintages of data for the estimation, not just a single series of real-time data. Such an approach should better mimic the time series that are actually available to the decision-maker.

2. Methodology and Preliminaries

2.1 Model

This paper uses a New Keynesian (NK) Dynamic Stochastic General Equilibrium (DSGE) model. The model is derived from microeconomic behavior of particular economic agents. These include domestic and foreign households, domestic and foreign producers, domestic importers and domestic and foreign monetary authorities. Most of the model assumptions are adopted from Lubik and Schorfheide (2006).[1]

In baseline setting, domestic and foreign monetary authorities follow Taylor-type rules (the variant where central monetary authorities care about year-on-year rather than quarter-on-quarter changes is also investigated):

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left[\psi_1 \pi_t + \psi_2 \left(\Delta y_t + z_t\right) + \psi_3 \Delta e_t\right] + \varepsilon_{r,t}$$
$$r_t^* = \rho_r^* r_{t-1}^* + \left(1-\rho_r^*\right)\left[\psi_1^* \pi_t^* + \psi_2^* \left(\Delta y_t^* + z_t\right)\right] + \varepsilon_{r,t}^*$$

where r_t is the nominal interest rate at time t, ρ_r is a backward-looking parameter, π_t is inflation, Δy_t is the growth rate of real output, z_t is the growth rate of a worldwide technology shock, Δe_t is the depreciation of the domestic currency, ψs are the preference parameters and $\varepsilon_{r,t}$ is a monetary policy shock.[2] Variables and parameters with a star superscript refer to a foreign economy.

The model is in a small open economy (SOE) setting, so it presumes two countries—a small open economy influenced by a big closed economy. Also, the model incorporates an exchange rate channel, and the two countries in the model should therefore have different currencies. Finally, the formulation of the interest rate rule corresponds best with the inflation targeting regime.

Given the typical scope of the journal, this study concentrates on the countries of the Visegrád Four, of which three countries meet the restrictions given by the model formulation.[3] The small open economy is therefore the Czech, Hungarian or Polish economy. In all three cases, the euro area with 12 countries is the large economy.

2.2 Data

The observed variables were chosen in accordance with Lubik and Schorfheide (2006), where the authors use quarterly data for seven observable variables—output growth, CPI inflation and the three-month nominal interest rate, all for domestic and foreign economies, and the growth rate of bilateral nominal exchange rate

The data were acquired from OECD databases. GDP and the Consumer Price Index were acquired from a real-time database,[4] while the interest rates and the exchange rates were acquired from Key Economic Indicators.[5]

[1] See the *Appendix* for an extensive model description and log-linearized model form.

[2] See *Table 4* for a list of estimated parameters.

[3] Slovakia does not have independent monetary policy.

[4] http://stats.oecd.org/mei/default.asp?rev=1

[5] http://stats.oecd.org/Index.aspx?DataSetCode=KEI. Note that the Hungarian interest rate has several missing observations in the OECD database. The series was substituted from the Eurostat database (the data are virtually the same) where there is only one missing observation in 2004q3 which was linearly interpolated.

Real-time data is usually structured in vintages. A "vintage" is the quarter when the data become available or the time of publishing. For example, if the Czech Statistical Office releases an estimate of GDP growth for the last quarter of 2012 sometime in April of 2013, it is said that the data for the fourth quarter of 2012 are of the April 2013 vintage.

The term "real-time data" pertains to data that become available right after collection. Real-time macroeconomic data are usually available 3–4 months after the end of a quarter, typically being the first estimates published for that quarter. As time passes, new vintages become available and revised data become more accurate estimates of real values. The most recent vintage is referred to as the "final" value. The difference between the final and real-time data is a "total revision".

The original OECD real-time dataset for CPI has 180 *monthly* vintages (February 1999–January 2014) that cover 36 (January 1996–December 1998) to 215 *monthly* observations (January 1996–November 2013). The original real-time dataset for GDP also has 180 *monthly* vintages that cover 11 (1996Q1–1998Q3) to 71 *quarterly* observations (1996Q1–2013Q3).[6]

GDP in constant prices that was *not* seasonally adjusted in the dataset has been seasonally adjusted in Demetra[7] program using the Tramo & Seats method. Computed growth rates are quarter-on-quarter log-differences, the monthly Consumer Price Index was also seasonally adjusted. The third month in each quarter was used to compute quarter-on-quarter inflation (as log-differences). Both datasets were truncated so that there are at least 30 quarterly observations for the estimation.

Finally, there is the issue of which *monthly* vintages to select for estimation in each *quarter*. Since the model is quarterly, there are three choices available: to use January/April/July/October vintages, February/May/August/November vintages or March/June/September/December vintages. There are many difficulties with this choice and arguably none of the options is ideal. The approach used in this article is to use the vintage set that ensures the highest number of *balanced* real-time data subsets. A balanced real-time data subset in this context means that it includes the same number of quarterly observations for both GDP growth and inflation. This approach is convenient in that we do not need to discard any existing quarterly observations, nor do we need to estimate (nowcast) any non-existent observations. However, there are also drawbacks to this approach. First, since we focus only on quarterly observations, we disregard any monthly observations of inflation that may be available. Second, this approach does not discriminate between flash estimates (of GDP) and "regular" releases of national accounts data.[8,9] Third, the unbalanced real-time data subsets are not addressed—the affected vintages are simply not used for the estimation and no results are reported for that vintage. For the Czech and Polish economies, the January/April/July/October selection ensures the highest number

[6] Missing observations occur in some vintages for some countries. Such vintages are not used for the estimation.

[7] http://circa.europa.eu/irc/dsis/eurosam/info/data/demetra.htm

[8] I've conducted a sensitivity analysis (on Czech data and baseline setting) and used all three possible choices of sets of monthly vintages (in this case, ragged ends were cut) and the similarity of the results do not suggest there is a problem of vintage choice.

[9] I would like to thank an anonymous referee for pointing out these possible problems.

Table 1 Data, Trend, and Total Revision

Variable	Data revision	Trend	Total revision
domestic GDP growth	yes	constant	= data revision
domestic inflation	yes	constant	= data revision
domestic interest rate	no	HP filter	= trend revision
foreign GDP growth	yes	constant	= data revision
foreign inflation	yes	constant	= data revision
foreign interest rate	no	linear	= trend revision
nom. exchange rate growth	no	constant	= none

Notes: *Domestic* economies are Czech, Hungarian, and Polish. *Foreign* economy is euro area.

of balanced real-time data subsets (36 for CZ and 37 for PL).[10] The Hungarian data were probably published with different timing and ensure only one balanced real-time subset for the January/April/July/October vintages. On the other hand, the March/June/September/December vintages offer 21 balanced subsets and this setup is therefore used for the estimation on Hungarian data.

There are no real-time datasets for the remaining observable variables and truncated time series are therefore used for the estimation. The interest rates are three-month interbank rates without any transformation.[11] Quarterly nominal exchange rates were collected as "USD monthly averages" and transformed into domestic currency vs. euro in direct quotation, which means that the rise of its value reflects the depreciation of the domestic currency.[12] The growth rate of the nominal exchange rate was calculated as log-differences.

If the originally published data need to be detrended prior to use, then the issue of trend recomputation comes into play. When new data become available, their influence can be seen in two directions. First, the new vintage delivers more accurate data for historical periods—this influence is often referred to as "data revision". Secondly, the new data point for the new period enables a more accurate estimation of the trend for historical periods—this influence is "trend revision". The sum of data and trend revisions yields total revision.

Table 1 summarizes which macroeconomic variables are subject to data revision and which are subject to trend revision. Domestic and foreign GDP growth and inflation are part of the real-time database and are therefore subject to data revision. However, these variables are stationary and detrending them only requires deducting the means. Trend revision is negligible for constant trend, which is

[10] The following table shows the number of (un)balanced data subsets for Czech data and baseline model setting.

Vintage set	Balanced	Unbalanced	Total
January/April/July/October	36	5	41
February/May/August/November	1	39	40
March/June/September/December	19	21	40

[11] Note that several missing values were filled in with Eurostat data, which are consistent with OECD data.

[12] This selection and computation was used because the OECD dataset does not contain currencies quoted in EUR. The calculations were cross-checked against Eurostat datasets and the series are virtually the same.

the reason why it is omitted in this study. Interest rates are not subject to data revision, but are not stationary. The interest rates in the euro area are detrended by a (linear) time trend. Domestic interest rates are even less regular and are detrended by a Hodrick-Prescott filter. Nominal exchange rate growth is not subject to data revision and the series is stationary, which means that this series is not subject to any revision. Due to the availability of variables in the real-time database and the choice of detrending methods, no variable is subject to both types of revision.

In order to conveniently distinguish between data and trend revisions, the concept of so-called "quasi real-time data" is usually introduced.[13] Quasi real-time data are constructed with knowledge of the latest vintage but not of future values. The researcher therefore knows what data revision for today's value will occur tomorrow but she does not know any values for tomorrow's time period. Therefore, quasi real-time data isolates trend revision. Quasi real-time data minus real-time data is data revision and final data minus quasi real-time data is trend revision.

All observable variables enter the model as quarter-on-quarter growth rates (interest rates are quarterly) and *per quartal*.[14]

2.3 Recursive Estimates and Statistical Significance

This study undertakes a recursive analysis in order to analyze the influence of the use of real-time data on the differences in monetary authorities' preference parameter estimates in the course of time. The recursive analysis is conducted in such a way that the first observation is always the same and the last observation shifts by a quarter each time a new estimation is carried out. A logical implication is that the time frame of the estimate grows. The series of such estimates may be intuitively perceived as the exploration of the information in the newly added data. Although this intuition is not entirely correct, this paper concentrates on the influence of real-time data rather than on weaknesses of recursive estimation.

The log-linearized DSGE model is estimated using Bayesian methods. A numerical-optimization procedure is used to maximize the posterior. At least 1,000,000 draws from posterior density are generated with a random-walk Metropolis-Hastings algorithm, after which the convergence is checked according to Brooks and Gelman (1998) convergence diagnostics. If the chain does not converge, 1,000,000 more samples are added and the convergence is rechecked until convergence is reached.[15] Then 90% of the original sample is discarded and the rest is used for posterior analysis. The estimation is carried out in the Dynare software.[16] A Monte Carlo-based optimization routine is used for computing the mode so that different estimates all reach suitable acceptation rate. Parameters' prior densities are the same for all estimates.

Note that, unlike the cited literature, data enter the estimation in respective vintages for each estimation. A typical approach in the existing literature is to form one time series of real-time data and repeatedly truncate it to obtain recursive esti-

[13] See, for example, Orphanides and van Norden (2002) and Ince and Papell (2013).

[14] See part 2.4 for two exceptions with selective year-on-year transformation.

[15] When the chain does not converge, results are not reported.

[16] See http://www.dynare.org/; version 4.4.1 was used.

mates. The approach of this article is different: it uses the whole vintage of data for each successive recursive estimation. The data source for one "real-time" macroeconomic variable is therefore not one time series, but a matrix of data with separate vintages. Although computationally more demanding, this approach mimics more closely the data actually available to economic agents at any point in time.

For real-time estimates, the trend estimates are based only on the data that were actually available. For quasi real-time estimates, all data are fully revised but values for future periods are not known. For final estimates, all data are fully revised and future values are available for computation of the trend. In order to capture the evolution of the estimates, final data series are truncated to match the period of estimation.

In order to be able to conclude whether potential differences between real-time, quasi real-time and final estimates are statistically significant, Section 3 reports the lowest significance level at which the (major) mode is out of the Highest Posterior Density bands for the two most different estimates.

2.4 Estimated Variants

The study offers several different model specifications, data treatments and estimation procedures to show the robustness of the results. The setting introduced in previous sections of part 2 is called *baseline*.

CPI not s.a. stands for the model variant with the Consumer Price Index not seasonally adjusted; *GDP HP* denotes a variant with the growth rate of Gross Domestic Product detrended by a Hodrick-Prescott filter; *rolling* uses estimation (of the baseline model) in a moving window of fixed length 30; *YOY* stands for a variant with monetary authorities that care about year-on-year (rather than quarter-on-quarter) changes and *YOY+GDP HP* denotes a variant with monetary authorities that care about year-on-year changes and GDP is detrended by HP filter.

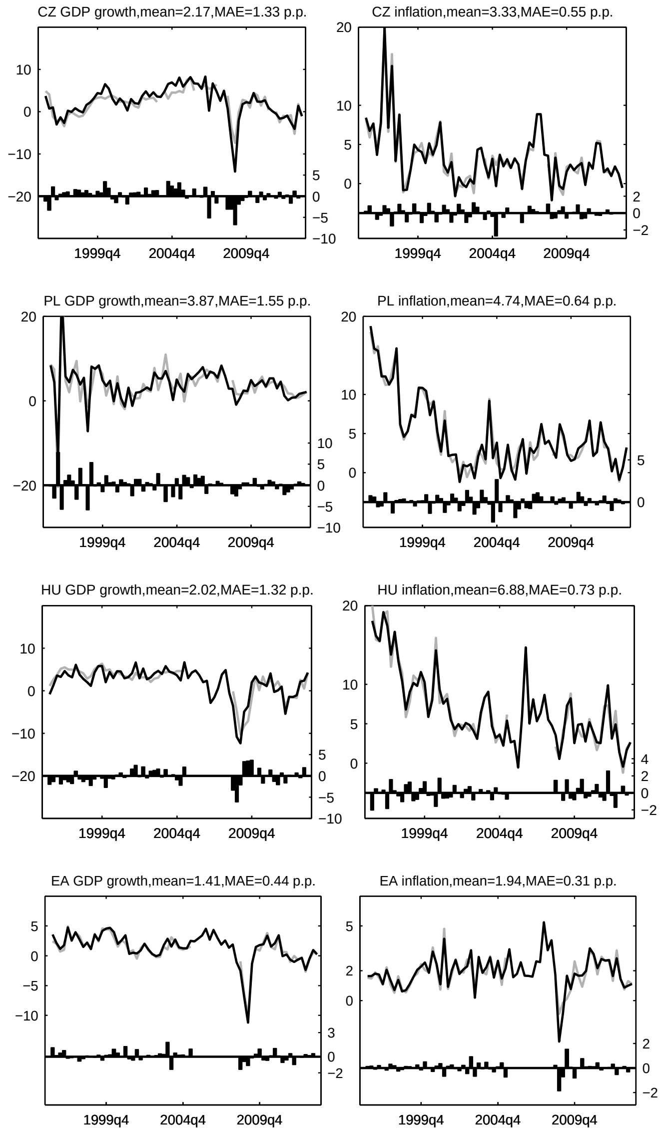
3. Recursive Analysis of Real-Time Data

Figure 1 displays real-time and final data for the Czech, Polish, Hungarian and euro area economies. Note that in order to see the magnitude and regularity of data revisions, the data are not detrended here. Also, in order to form one "real-time" time series for each macroeconomic variable, only the most recent data point per vintage is considered. This treatment corresponds to the use of "real-time" data in the literature and is therefore directly comparable. However, note that these depicted series do not enter the estimation; separate vintages do.

In all of the countries in the study GDP growth suffers from revision much more than inflation. Also, the depicted data for the euro area are markedly less subject to data revision than data of the remaining countries.

Close inspection of the GDP graphs reveals some similarities and some specific aspects among countries. In all of the countries, the severity of the economic crisis of 2009 was underestimated, i.e. it was milder according to real-time data than it actually was after revisions. The economic crisis was also the period with the largest revisions on the sample for the Czech and Hungarian economies. Polish revisions of GDP growth were biggest prior to 1999 and there are no apparent differences in revisions in the euro area.

Figure 1 Real-Time and Final Data



Note: The *black* line denotes final data (left axis), the *gray* line denotes real-time data (left axis), *errorbar* depicts data revision (right axis), *mean* denotes the average of final data over the sample, and *MAE* denotes the mean absolute error in data revision.

As for the inflation graphs depicted in the right-hand panels of *Figure 1*, revisions for all of the countries except for the euro area are autocorrelated in lags 2 and 4. This autocorrelation seems to be an artifact of seasonal adjustment of the underlying CPI series. However, since the series of revisions do not enter the estimation (separate data vintages do), this problem remains only with the statistics in *Table 2.*

Table 2 Summary Statistics of Data Revisions

	Mean	*Min*	*Max*	*St. Dev.*	*RMSE*	*N/S*	*Corr*	*p-val*	*AR(1)*	*Rev+*
					Output growth					
Czech Rep.	0.13	-6.77	3.46	1.80	1.80	0.49	0.88	0.68	0.32	0.58
Poland	0.04	-5.86	7.72	2.17	2.18	0.52	0.91	0.83	-0.40	0.57
Hungary	-0.19	-6.14	3.63	1.73	1.74	0.49	0.88	0.49	0.36	0.40
EA12	0.14	-1.59	2.26	0.71	0.73	0.29	0.96	0.15	0.07	0.63
					Inflation					
Czech Rep.	-0.02	-2.71	1.20	0.73	0.73	0.20	0.98	0.73	0.00	0.54
Poland	0.03	-2.36	2.66	0.83	0.83	0.19	0.98	0.60	-0.16	0.54
Hungary	0.01	-2.02	2.56	0.93	0.93	0.20	0.98	0.94	-0.31	0.45
EA12	0.04	-1.86	1.54	0.44	0.45	0.40	0.91	0.29	-0.14	0.59

Notes: *N/S* denotes the noise-to-signal ratio defined as the standard deviation of the revisions divided by the standard deviation of the final value of the variable; *Corr* is the correlation of final and real-time data; *p-val* is a *p*-value for a test that the mean revision is zero using autocorrelation and hetero-scedasticity-consistent standard errors. *AR(1)* denotes an autocorrelation coefficient of the first order (missing data are estimated in an iterative fashion using default order state-space models); *Rev+* denotes the frequency at which final data is greater than real-time data, i.e. final revision is positive.

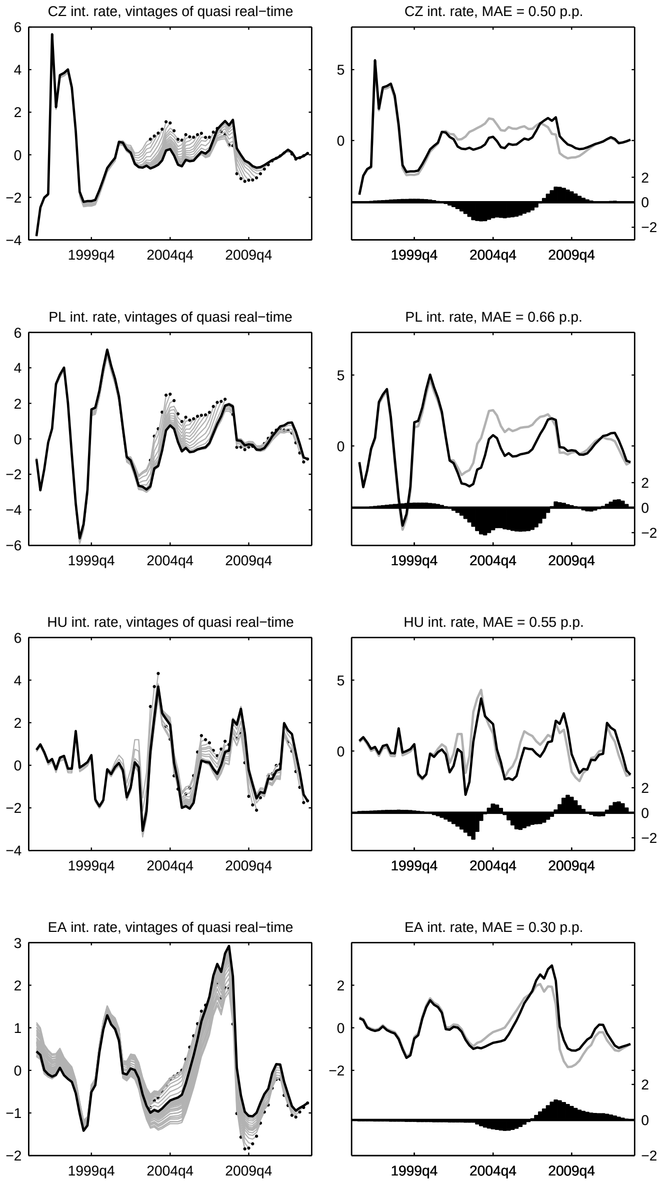
Apart from the autocorrelation, there does not seem to be any other regularity in inflation revisions.

Table 2 summarizes the statistics of data revision. The revisions are unbiased for all of the countries and both variables since the lowest p-value (0.15) is greater than conventional levels. All correlations are high, which means that the final and real-time data have very similar dynamics. Revisions show a very low autocorrelation coefficient (of order 1), which indicates that period-to-period revisions are not systematic. Euro area GDP growth suffers from modest underestimation of growth, as the real-time estimate is lower than the final data 63% of the time. The noise-to-signal ratio for GDP growth confirms that the relative magnitude of revisions is lowest for the euro area, and the remaining countries exhibit an approximately 70% higher value of this indicator. The noise-to-signal ratios for inflation are rather unexpected, with the value of the ratio being twice that of the remaining central European countries. Note that this result also probably stems from the seasonal adjustment of the CPI series.[17]

Figure 2 presents a graphical illustration of trend revisions for the Czech, Polish, Hungarian and euro area interest rates. The left-hand panels illustrate the evolution of vintages of quasi real-time data. Each thin gray line displays the evolution of quasi real-time data in a particular vintage. Considering only the *last* values (denoted as black dots) of each quasi real-time vintage yields the curve of quasi real-time data depicted in the right-hand panels (in gray). The difference between final and quasi real-time data is trend revision. Since the trend estimation changes gradually, trend revision is highly autocorrelated. Note that the interest rates depicted in first three rows of the panels (Czech, Polish and Hungarian interest rates) are detrended by a Hodrick-Prescott filter, whereas the euro area interest rate is detrended by a linear time trend.

[17] Summary statistics of data revisions for the Czech economy are largely in line with Rusnák (2013), with the exception of p-values, where the author in Table 1 (p. 250) reports p-values even greater than one.

Figure 2 Final and Quasi Real-Time Data



Notes: The left-hand panels display final detrended data (black) and all vintages of quasi real-time data (gray). The right-hand panels display final detrended data (black) and quasi real-time data (gray), while both use the left axis. The difference between final detrended data and quasi real-time data is trend revision depicted with the errorbar that uses the right axis. *MAE* denotes the mean absolute error in trend revision.

Table 3 offers summary statistics for trend revisions. As has already been mentioned, all trend revisions are highly (first-order) autocorrelated. Over/under-

Table 3 Summary Statistics of Trend Revisions (final–quasi real-time data)

	Mean	*Min*	*Max*	*St. Dev.*	*RMSE*	*N/S*	*Corr*	*p-val*	*AR(1)*	*Rev+*
					Interest rate					
Czech Rep. (HP)	-0.15	-1.49	1.21	0.68	0.70	0.44	0.91	0.39	0.98	0.57
Poland (HP)	-0.39	-2.16	0.63	0.86	0.95	0.43	0.91	0.08	0.98	0.50
Hungary (HP)	-0.14	-2.11	1.39	0.72	0.74	0.55	0.85	0.42	0.93	0.51
EA12 (linear)	0.06	-0.58	1.12	0.41	0.41	0.40	0.92	0.56	0.98	0.36

Notes: *Corr* is the correlation of final and quasi real-time data, *Rev+* denotes the frequency at which final data is greater than quasi real-time data. *HP* denotes the Hodrick-Prescott filter as a detrending method; *linear* denotes the linear time trend as a detrending method. For other notes, see *Table 2*.

Table 4 Summary of Model Parameters Relating to Domestic and Foreign Monetary Policy

Par.	*Description*
ψ_1	weight on inflation in domestic monetary rule
ψ_2	weight on output growth in domestic monetary rule
ψ_3	weight on nominal depreciation in domestic monetary rule
ψ_1^*	weight on inflation in foreign monetary rule
ψ_2^*	weight on output growth in foreign monetary rule
ρ_r	AR1 persistence in domestic monetary rule
ρ_r^*	AR1 persistence in foreign monetary rule

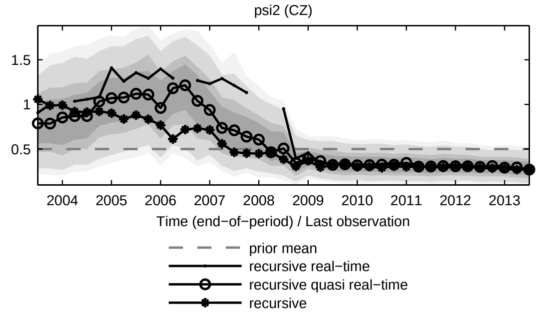
estimation of the final trend is roughly balanced for all economies except for the euro area, where quasi real-time data underestimate the final trend in 64% of periods. *p-val* for Poland is another result that stands out in *Table 3*. It is 0.08, which makes it the only biased revision in the whole study, i.e. the mean of the revision -0.39 is statistically different from zero at the 0.1 significance level.

4. Recursive Estimates of the Preferences of Monetary Authorities

4.1 Czech Republic

Figure 3 displays the evolution of recursive, recursive real-time and recursive quasi real-time estimates for the weight on output growth ψ_2 in the Czech Taylor rule. Probability bands are drawn around quasi real-time estimates and this depiction is therefore convenient for identification of deviations from quasi real-time estimates. The evolution of the estimates in time is different. Recursive estimates (on final data) gradually decrease, while recursive real-time estimates remain at parameter values over 1 and the estimate falls below 0.5 in 2008q3. Quasi real-time estimates are between the other two, which suggests that data revision and trend revision shift the estimate in the same direction. The fact that only real-time estimates do not gradually decrease indicates that the information of the change was not in the data

Figure 3 Recursive Estimates of the Weight on Output Growth ψ_2 in Domestic Taylor Rule, Czech Economy



Note: The depicted estimates are posterior modes with 95%, 90%, 68% and 50% Highest Posterior Density Intervals (HPDI) for recursive quasi real-time estimates.

earlier than 2008q3. The other two estimates contain the information due to subsequent revisions. The period 2008q3 matches the onset of the current economic crisis, which is—considering real-time conditions—unpredicted. Following the results for real-time data, the central bank did not change its reaction to output growth until 2008q3, when a drastic change occurred. On the other hand, following the results for quasi real-time and final data, the central bank gradually lowered its preference parameter towards output growth. This result suggests that conducting ex-post analysis of monetary policy on revised data indeed generates misleading results.

Table 5 presents significance values for various model alternatives in order to determine whether the difference between recursive real-time estimates and the remaining estimates is statistically significant.

Following the discussion related to parameter ψ_2, we can see that the real-time vs. quasi real-time significance value is 0.06, thus being significant at the 10% level. Since the difference between real-time and quasi real-time estimates is only caused by data revision, the interpretation is that the nature of real-time data causes significantly different estimates of the preferences of the Czech central bank to output growth. However, the robustness of this result is limited. While it is robust to seasonal adjustment of CPI, the year-on-year specification in Taylor rules and the moving-window estimation, it is clearly not robust to the detrending method of GDP growth.

There are two more selectively significant results in *Table 5*. The weight on output growth ψ_2^* in the foreign Taylor rule is significantly different in *Trend revision*, which indicates that detrending is the reason. Moreover, the result is valid only for model variants where the central bank uses the year-on-year specification. A rather different story is behind the significant result for the smoothing term in the domestic Taylor rule ρ_r. The difference is only significant for the moving-window estimation, but since the results are under *Data revision*, the reason is the nature of real-time data itself, not the detrending issue. This significance occurs due to a lag in shifting between two regimes in the parameter ρ_r.

Table 5 Significance Values for Estimates on the Czech and Euro Area Economies

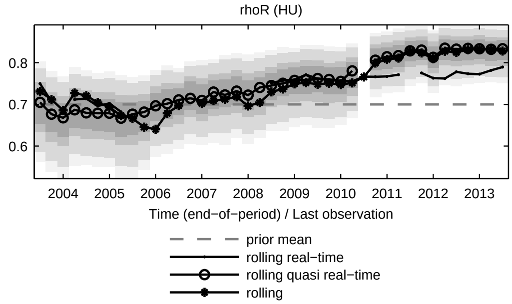
	Baseline	CPI not s.a.	GDP HP	Rolling	YOY	YOY+GDP HP
	ψ_1 weight on inflation in domestic monetary rule					
Trend revision	0.86	0.89	0.77	0.23	0.82	0.84
Total revision	0.61	0.62	0.54	0.16	0.79	0.84
Data revision	0.56	0.62	0.55	0.36	0.77	0.84
	ψ_2 weight on output growth in domestic monetary rule					
Trend revision	0.05**	0.32	0.73	0.06*	0.22	0.59
Total revision	0.01***	0.02**	0.73	0.01***	0.01***	0.38
Data revision	0.06*	0.06*	0.73	0.10*	0.02**	0.44
	ψ_3 weight on nominal depreciation in domestic monetary rule					
Trend revision	0.75	0.79	0.81	0.71	0.80	0.89
Total revision	0.44	0.68	0.63	0.45	0.65	0.83
Data revision	0.58	0.70	0.63	0.42	0.70	0.85
	ψ_1^* weight on inflation in foreign monetary rule					
Trend revision	0.61	0.68	0.70	0.55	0.33	0.26
Total revision	0.48	0.64	0.53	0.56	0.27	0.06*
Data revision	0.59	0.77	0.63	0.50	0.49	0.25
	ψ_2^* weight on output growth in foreign monetary rule					
Trend revision	0.31	0.45	0.46	0.28	0.01***	0.03**
Total revision	0.47	0.42	0.23	0.35	0.04**	0.01***
Data revision	0.45	0.54	0.35	0.42	0.53	0.08*
	ρ_r AR1 persistence in domestic monetary rule					
Trend revision	0.49	0.61	0.68	0.59	0.63	0.76
Total revision	0.29	0.46	0.54	0.01***	0.57	0.67
Data revision	0.60	0.61	0.71	0.01***	0.42	0.59
	ρ_r^* AR1 persistence in foreign monetary rule					
Trend revision	0.58	0.62	0.53	0.51	0.15	0.34
Total revision	0.27	0.61	0.32	0.31	0.07*	0.30
Data revision	0.41	0.77	0.51	0.36	0.57	0.31

Notes: The numbers in the table are the lowest levels of significance at which the posterior mode is out of the Highest Posterior Density interval bands for the most different estimates. *Trend revision* relates to computation of significance values on recursive versus quasi real-time results; *Total revision* relates to recursive versus real-time results' and *Data revision* relates to quasi real-time versus real-time results. Values lower than or equal to 0.1 are denoted with a star, those lower than or equal to 0.05 with two stars, and those lower than or equal to 0.01 with three stars (due to the computational procedure, values in the table are rounded up).

4.2 Hungary

This subsection investigates the case of the small open economy of Hungary in interaction with the euro area. *Figure 4* depicts the evolution of the persistence in the domestic monetary rule ρ_r. *Figure 4* demonstrates a different course of rolling

Figure 4 Recursive Estimates of the AR1 Persistence in the Domestic Monetary Rule ρ_r, Hungarian Economy



Note: The depicted estimates are posterior modes with 95%, 90%, 68% and 50% Highest Posterior Density Intervals (HPDI) for recursive quasi real-time estimates.

the real-time estimate and two remaining rolling estimates. This observation suggests that real-time data do not favor the rise in the smoothing parameter as the revised data do. The central bank does not have sufficient grounds in real time to change its policy, but according to the revised data, it should have changed the policy.[18]

Table 6 indicates statistically significant results for parameter ψ_2 in a variety of estimated models and for parameters ψ_2^*, ρ_r and ρ_r^* only in the rolling estimation. Observing the results more closely, the significance values in the case of parameter ψ_2 seems to stem from the uninformativeness and bimodality of respective distributions, which indicates either a switch between parameter regimes or an artifact of poor data quality. As for the remaining three parameters with statistically significant results only for rolling estimates, the explanation for domestic and foreign smoothing parameters is the same and it is a preference of lower interest rate smoothing when real-time data are considered. Note that the distributions are rather informative and unimodal in these two cases; the results are therefore convincing. On the other hand, the results for parameter ψ_2^* stem from the bimodality of respective distributions. The different modes are major at different times and switching between the two creates significant results.

4.3 Poland

This section presents the results for a model with the domestic economy of Poland. *Figure 5* displays the course of estimates of parameter ψ_2. There is an apparent switch in the weight on output from the values close to prior distribution to the values three times larger. However, this regime switch occurs in different periods, which generates statistically significant results across all estimated model variants.

Similarly to the previous countries, *Table 7* presents significance values of various estimates. The case of ψ_2 has already been discussed and there is no other economically interesting statistically significant result for the Polish economy.[19]

[18] Note that the discussed result is significant only for the moving-window estimation, since in the recursive estimates the newly added information in the data is not strong enough to dilute the predominant information from the beginning of the sample.

[19] The results for ψ_2^* are consequences of the bimodality of distributions.

Table 6 Significance Values for Estimates on the Hungarian and Euro Area Economies

	Baseline	CPI not s.a.	GDP HP	Rolling	YOY	YOY+ +GDP HP
	ψ_1 weight on inflation in domestic monetary rule					
Trend revision	0.85	0.56	0.78	0.13	0.64	0.60
Total revision	0.42	0.68	0.49	0.37	0.79	0.47
Data revision	0.50	0.54	0.55	0.33	0.65	0.19
	ψ_2 weight on output growth in domestic monetary rule					
Trend revision	0.01***	0.31	0.13	0.01***	0.01***	0.01***
Total revision	0.70	0.78	0.51	0.30	0.37	0.01***
Data revision	0.03**	0.40	0.67	0.04**	0.01***	0.01***
	ψ_3 weight on nominal depreciation in domestic monetary rule					
Trend revision	0.67	0.71	0.72	0.67	0.79	0.85
Total revision	0.55	0.77	0.58	0.47	0.80	0.81
Data revision	0.61	0.79	0.61	0.54	0.80	0.82
	ψ_1^* weight on inflation in foreign monetary rule					
Trend revision	0.66	0.62	0.65	0.65	0.25	0.44
Total revision	0.48	0.49	0.59	0.37	0.19	0.35
Data revision	0.62	0.71	0.64	0.25	0.63	0.76
	ψ_2^* weight on output growth in foreign monetary rule					
Trend revision	0.19	0.20	0.35	0.02**	0.08*	0.03**
Total revision	0.16	0.19	0.11	0.01***	0.17	0.02**
Data revision	0.49	0.35	0.12	0.02**	0.48	0.07*
	ρ_r AR1 persistence in domestic monetary rule					
Trend revision	0.44	0.58	0.50	0.32	0.47	0.20
Total revision	0.63	0.68	0.58	0.07*	0.62	0.29
Data revision	0.35	0.76	0.50	0.04**	0.49	0.09*
	ρ_r^* AR1 persistence in foreign monetary rule					
Trend revision	0.42	0.67	0.41	0.34	0.15	0.32
Total revision	0.51	0.54	0.22	0.03**	0.21	0.29
Data revision	0.56	0.68	0.46	0.04**	0.72	0.22

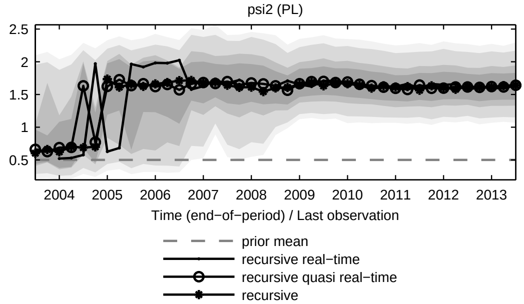
Note: See *Table 5* below.

Table 7 Significance Values for Estimates on the Polish and Euro Area Economies

	Baseline	CPI not s.a.	GDP HP	Rolling	YOY	YOY+ +GDP HP
	ψ_1 weight on inflation in domestic monetary rule					
Trend revision	0.52	0.41	0.59	0.62	0.72	0.71
Total revision	0.37	0.31	0.36	0.18	0.76	0.73
Data revision	0.43	0.32	0.15	0.20	0.72	0.73
	ψ_2 weight on output growth in domestic monetary rule					
Trend revision	0.74	0.68	0.72	0.20	0.11	0.18
Total revision	0.01***	0.03**	0.01***	0.04**	0.09*	0.02**
Data revision	0.01***	0.05**	0.01***	0.03**	0.09*	0.02**
	ψ_3 weight on nominal depreciation in domestic monetary rule					
Trend revision	0.70	0.66	0.70	0.65	0.75	0.87
Total revision	0.68	0.57	0.60	0.44	0.53	0.77
Data revision	0.68	0.58	0.65	0.51	0.54	0.83
	ψ_1^* weight on inflation in foreign monetary rule					
Trend revision	0.73	0.67	0.48	0.70	0.17	0.42
Total revision	0.65	0.69	0.46	0.53	0.32	0.43
Data revision	0.79	0.70	0.62	0.49	0.50	0.79
	ψ_2^* weight on output growth in foreign monetary rule					
Trend revision	0.37	0.30	0.51	0.47	0.03**	0.27
Total revision	0.39	0.24	0.40	0.45	0.02**	0.24
Data revision	0.40	0.27	0.69	0.44	0.01***	0.58
	ρ_r AR1 persistence in domestic monetary rule					
Trend revision	0.69	0.75	0.60	0.61	0.44	0.63
Total revision	0.62	0.64	0.49	0.23	0.35	0.30
Data revision	0.61	0.71	0.53	0.16	0.50	0.28
	ρ_r^* AR1 persistence in foreign monetary rule					
Trend revision	0.60	0.64	0.45	0.52	0.22	0.21
Total revision	0.33	0.35	0.22	0.33	0.21	0.07*
Data revision	0.51	0.44	0.56	0.50	0.77	0.09*

Note: See *Table 5* below.

Figure 5 Recursive Estimates of the Weight on Output Growth ψ_2 in Domestic Taylor Rule, Polish Economy



Notes: The depicted estimates are posterior modes with 95%, 90%, 68% and 50% Highest Posterior Density Intervals (HPDI) for recursive quasi real-time estimates.

5. Conclusion

The goal of this paper is to investigate the differences between parameter estimates of monetary policy reaction functions using real-time data and those using revised data. Because such analysis uses real-time datasets, this paper also offers an analysis of real-time macroeconomic data, data revision and trend revision in the Czech Republic, Poland, Hungary and the euro area.

Data revisions of GDP growth and inflation are unbiased and not autocorrelated in all countries. Inflation is usually measured accurately in real time. Its noise-to-signal ratio ranges from 0.19 in Poland to a surprising 0.40 in the euro area. GDP growth is generally subject to greater data revision, with a noise-to-signal ratio ranging from 0.29 in the euro area to 0.52 in Poland.

Trend revisions are calculated with a linear time trend in the euro area and using the Hodrick-Prescott filter in the remaining countries. As was expected, trend revisions are highly autocorrelated and unbiased; the only exception to this is in Poland, where trend revision is biased. The noise-to-signal ratios are similar in value and range from 0.40 in the euro area to 0.55 in Hungary.

In its main analysis, this paper has revealed many statistically significant differences between parameter estimates of monetary policy reaction functions calculated using real-time data and those calculated using revised data. However, only a few of these results are robust enough to conclude that monetary authorities' preferences are different when considering real-time data as opposed to final data. The difference between estimates for the preference for output growth is the most common statistically significant result across countries. In other words, the monetary policy reaction to changes in output growth is statistically significantly different in strength when based only on the data available at the time of the monetary policy decision rather than on revised data. This result is in line with expectations, since the statistics for output growth data revision indicate that it is much more pronounced than inflation revision. Although several more differences were ascertained, only a few of them are adequately robust to different model specifications.

The results achieved raise several questions. Since the statistical offices in the Visegrád Four countries are rather young, would the results of such an analysis for developed countries in Western Europe be different? Also, if there are differences in policy parameters, are there also differences in other model parameters (possibly even deep parameters)? These questions could be a possible focus of further research.

REFERENCES

Aruoba SB (2008): Data Revisions Are Not Well Behaved. *Journal of Money, Credit and Banking*, 40:319–340.

Aurelio MM (2005): Do we really know how inflation targeters set interest rates? *FRB Kansas City Working Papers*, no. 05-02.

Belke A, Klose J (2011): Does the ECB Rely on a Taylor Rule During the Financial Crisis? Comparing Ex-post and Real Time Data with Real Time Forecasts. *Economic Analysis & Policy*, 41:147–172.

Brooks SP, Gelman A (1998): General Methods for Monitoring Convergence of Iterative Simulations. (American Statistical Association, Institute of Mathematical Statistics, and Interface Foundation of America) *Journal of Computational and Graphical Statistics*, 7:434–455.

Čapek, J (2015): Estimating DSGE model parameters in a small open economy: Do real-time data matter? *Národohospodářský obzor-Review of Economic Perspectives*, forthcoming. (DOI: 10.1515/revecp-2015-0001)

Casares M, Vázquez J (2012): Data Revisions in the Estimation of DSGE models. *University of the Basque Country, Dpt. of Foundations of Economic Analysis II (DFAEII)*, Series/Report no. 2012.06 (http://hdl.handle.net/10810/8759).

Cogley T, Sbordone AM (2005): A search for a structural Phillips curve. FRB of New York, Research Paper Staff Report, no. 203.

Croushore D (2011): Frontiers of Real-Time Data Analysis. *Journal of Economic Literature*, 49:72–100.

Cusinato R, Minella A, Silva Pôrto (Jr.) S da (2013): Output gap in Brazil: A real-time data analysis. *Empirical Economics*, 44:1113–1127.

Gerberding C, Seitz F, Worms A (2005): How the Bundesbank really conducted monetary policy. *North American Journal of Economics and Finance*, 16:277–292.

Gerdesmeier D, Roffia B (2005): The relevance of real-time data in estimating reaction functions for the euro area. *North American Journal of Economics and Finance*. 16(3):293–307.

Horváth R (2009): The time-varying policy neutral rate in real-time: A predictor for future inflation? *Economic Modelling*, 26:71–81.

Ince O, Papell D (2013): The (un)reliability of real-time output gap estimates with revised data. *Economic Modelling*, 33:713–721.

Lubik T, Schorfheide F (2006): A Bayesian Look at the New Open Economy Macroeconomics. *NBER Macroeconomics Annual 2005* (vol. 20):313–382.

Maravall A, Pierce D (1986): The transmission of data noise into policy noise in US monetary control. *Econometrica*, 54:961–979.

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

Neri S, Ropele T (2012): Imperfect Information, Real-Time Data and Monetary Policy in the Euro Area. The *Economic Journal*, 122(561):651–674.

Orphanides A (2001): Monetary Policy Rules Based on Real-Time Data. *American Economic Review*, 91:964–985.

Orphanides A, Norden S van (2002): The Unreliability of Output-Gap Estimates in Real Time. *Review of Economics and Statistics*, 84:569–583.

Pivetta F, Reis R (2007): The persistence of inflation in the United States. *Journal of Economic Dynamics and Control*, 31:1326–1358.

Rusnák M (2013): Revisions to the Czech National Accounts: Properties and Predictability. *Finance a úvěr-Czech Journal of Economics and Finance*, 63:244–261.

Schorfheide F (2008): Comment on "How Structural Are Structural Parameters?" *NBER Macroeconomics Annual 2007* (vol. 22):149–163.

Vázquez J, María-Dolores R, Londoño J (2010): *Data Revisions and the Monetary Policy Rule: An analysis based on an extension of the basic New Keynesian model*. The University of the Basque Country. Mimeo, available at:
http://www.ehu.es/jesusvazquez/note_v12.pdf or http://www.ehu.es/jesusvazquez/VML_OBES.pdf.

APPENDIX

1. Model Description

This article uses a New Keynesian (NK) Dynamic Stochastic General Equilibrium (DSGE) model which is derived from microeconomic behavior of particular economic agents. These include domestic and foreign households, domestic and foreign producers, domestic importers and domestic and foreign monetary authority. The model is in small open economy (SOE) setting, therefore presuming two countries – a small open economy influenced by a big closed economy. The small open economy is the home (Czech) economy, the big large economy is the foreign (euro area with 12 countries) economy. Most of the model assumptions are adopted from Lubik and Schorfheide (2006).

The economy is populated by a continuum of households that consume, work in order to earn money for consumption, and enter financial market in order to bridge the time gap between pay-day and consumption.

The production part of the economy consists of a continuum of monopolistically competitive firms which produce a differentiated product. Each producer enters a perfectly competitive labor market and uses labor input to produce goods according to production with labor-augmenting home-specific stationary technology. Technology is identical for all producers and it evolves exogenously over time. Each period, fraction $1-\theta_H$ of domestic firms sets their prices optimally, and fraction θ_H of firms does not change prices. The optimized price-setting decisions of firms results in a New-Keynesian Phillips curve.

Deviations from purchasing power parity (PPP) can occur in the short run due to existence of monopolistically competitive importers. Importers buy from foreign producers for prices set by foreign producers in a foreign currency. This means that law of one price holds at the border but not necessarily in the domestic economy, because importers set prices in domestic currency for domestic consumers with a possible mark-up. The fact that the law of one price does not have to hold can be also interpreted by an incomplete pass-through from exchange rate movements to prices of imports for domestic consumers.

Similarly to producers, importers operate under the Calvo-style price-setting with $1-\theta_F$ importers who are able to re-optimize their prices. The solution of this optimization problem results in a Phillips curve for imported inflation which can be derived by analogy to producers' sector.

The model presumes complete markets for securities traded internationally as there is a perfect risk-sharing between households in domestic and foreign economy. In another words, stochastic discount factors for domestic and foreign economy must be equal.

2. Log-linearized model form

This section summarizes log-linearized equations that are used as model equations. Starting with households, the system contains equations for evolution of marginal utility of income (1), the law of motion of habit stock (2), Euler equation (3), and the definition of inflation from domestic and imported inflation (4):

$$-\lambda_t = \frac{\tau}{1-h\beta}c_t - \frac{h\beta}{1-h\beta}E_t(\tau c_{t+1} + z_{t+1}) \tag{1}$$

$$c_t = \frac{1}{1-h}(c_t - hc_{t-1} + hz_t) \tag{2}$$

$$-\lambda_t = -E_t\lambda_{t+1} - (r_t - E_t\pi_{t+1}) + E_t z_{t+1} \tag{3}$$

$$\pi_t = (1-\alpha)\pi_{H,t} + \alpha\pi_{F,t}, \tag{4}$$

Behavior of producers yield the New-Keynesian Phillips Curve (5) with marginal cost evolution described by (6)

$$\pi_{H,t} = \frac{1-\theta_H}{\theta_H}(1-\beta\theta_H)mc_{H,t} + \beta E_t\pi_{H,t+1} \tag{5}$$

$$mc_{H,t} = -\alpha q_t - \lambda_t - a_t \tag{6}$$

Importers' optimization is analogous to that of producers and for the log-linearized system is utilized importers' Phillips curve (7)

$$\pi_{F,t} = \frac{1-\theta_F}{\theta_F}(1-\beta\theta_F)\psi_{F,t} + \beta E_t \pi_{F,t+1} \tag{7}$$

There are also some simple definitions, namely definition of the depreciation rate of nominal exchange rate (8), differenced definition of terms of trade (9) and combined definition of real exchange rate and LOP gap (10).

$$\Delta e_t = \Delta s_t + \pi_t - \pi_t^*, \tag{8}$$

$$q_t = q_{t-1} + \pi_{H,t} - \pi_{F,t} \tag{9}$$

$$s_t = \psi_{F,t} - (1-\alpha) q_t \tag{10}$$

Equilibria equations includes equation regarding international risk-sharing (11), UIP condition (12) and log-linearized market clearing equation (13).

$$\lambda_t = \lambda_t^* - s_t. \tag{11}$$

$$r_t - r_t^* = E_t \Delta e_{t+1} \tag{12}$$

$$y_{H,t} = (1-\alpha)c_t + \alpha c_t^* + \alpha\eta(s_t - q_t) + g_{H,t} \tag{13}$$

Foreign economy is modeled structurally so that there exist foreign households and producers that also show optimizing behavior. However, since foreign economy is big and closed, its agents are not influenced in their optimization behavior by home economy activities.

Following equations are introduced in analogy to home case: A result of foreign representative household's optimizing behavior (14) and (15) analogous to (1) and (2), foreign producers' Phillips curve (16) analogous to (7) and rather collapsed version of market clearing (17) analogous to (13).

$$-\lambda_t^* = \frac{\tau}{1-h\beta} c_t^* - \frac{h\beta}{1-h\beta} E_t(\tau c_{t+1}^* + z_{t+1}) \tag{14}$$

$$c_t^* = \frac{1}{1-h}(c_t^* - hc_{t-1}^* + hz_t) \tag{15}$$

$$\pi_t^* = \frac{1-\theta^*}{\theta^*}(1-\beta\theta^*)(-\lambda_t^* - a_t^*) + \beta E_t \pi_{t+1}^* \tag{16}$$

$$y_t^* = c_t^* + g_t^* \tag{17}$$

The model is closed by specifying monetary policy. Towards this end, standard Taylor-type rule is used. This formulation of monetary policy assumes that central banks respond to deviations of inflation from steady state, growth rate of output from steady state growth rate γ and possibly to deviations of nominal exchange rate depreciation from steady state. Home and foreign monetary rules are therefore

$$r_t = \rho_r r_{t-1} + (1-\rho_r)[\psi_1 \pi_t + \psi_2(\Delta y_{H,t} + z_t) + \psi_3 \Delta e_t] + \varepsilon_{r,t} \tag{18}$$

$$r_t^* = \rho_r^* r_{t-1}^* + (1-\rho_r^*)[\psi_1^* \pi_t^* + \psi_2^*(\Delta y_t^* + z_t)] + \varepsilon_{r,t}^*, \tag{19}$$

where r_t is nominal interest rate, which is supposed to be monetary authority's tool, ρ_r is backward-looking parameter, ψ s are weights that monetary policy places on different economic variables it reacts to, and $\varepsilon_{r,t}$ is direct innovation to the rule that captures non-systematic part of monetary policy. Analogous explanations hold for foreign economy.

The model is supplemented with AR(1) processes describing evolution of government expenditures (acting as a demand or market clearing shock) g_t, country-specific technology shock to production function (acting as a supply shock) and the evolution of z_t, which is growth rate of world-wide non-stationary technology shock.

$$a_t = \rho_a a_{t-1} + \varepsilon_{a,t} \quad a_t^* = \rho_a^* a_{t-1}^* + \varepsilon_{a,t}^*$$

$$g_{H,t} = \rho_g g_{H,t-1} + \varepsilon_{g_H,t} \quad g_t^* = \rho_g^* g_{t-1}^* + \varepsilon_{g,t}^*$$

$$z_t = \rho_z z_{t-1} + \varepsilon_{z,t}$$

3. Summary of model variables, shocks and parameters

Table 1 Summary of model variables'

Variable	Loglinearized	Description
A_t	a_t	home-specific stationary technology shock
C_t	c_t	consumption relative to the level of technology
$C_{H,t}$	$c_{H,t}$	domestic consumption of domestic goods (relative to the level of technology)
$C_{H,t}^*$	$c_{H,t}^*$	foreign consumption of domestic goods (relative to the level of technology) = exports
$G_{H,t}$	$g_{H,t}$	domestic government expenditures
	g_t^*	foreign government expenditures
$MC_{H,t}$	$mc_{H,t}$	real marginal cost
$P_{H,t}$	$p_{H,t}$	domestic goods price index
$P_{F,t}$	$p_{F,t}$	foreign goods price index
P_t	p_t	price index
R_t	r_t	nominal interest rate (as growth coefficient)
S_t	s_t	real exchange rate
$Y_{H,t}$	$y_{H,t}$	domestic output
	y_t^*	foreign output
Z_t	z_t	growth coefficient of a world-wide technology shock
C_t	c_t	effective consumption relative to the level of technology
E_t	e_t	nominal exchange rate (direct quotation)
Q_t	q_t	terms of trade
Λ_t	λ_t	marginal utility of real income (adjusted for the level of technology)
Π_t	π_t	inflation
$\Psi_{F,t}$	$\psi_{F,t}$	law of one price gap

Note: Variables with direct transcript from domestic to foreign variable just with adding a star (*) are not listed.

Table 2 Summary of model shocks and innovations

Innovation	Enters as	Description
$\varepsilon_{r,t}$	directly	domestic monetary shock
$\varepsilon^*_{r,t}$	directly	foreign monetary shock
$\varepsilon_{a,t}$	AR1	domestic supply shock
$\varepsilon^*_{a,t}$	AR1	foreign supply shock
$\varepsilon_{g,t}$	AR1	domestic demand shock
$\varepsilon^*_{g,t}$	AR1	foreign demand shock
$\varepsilon_{z,t}$	AR1	world-wide technology shock

Notes: There is another misspecification innovation to the system. It is not listed here because the misspecification innovation is not a part of conceptual model. Final model for estimation therefore has 7 innovations listed here plus one more added due to model misspecification.

OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 82, 3 (2020) 0305–9049
doi: 10.1111/obes.12351

We just estimated twenty million fiscal multipliers*

Jan Čapek† and Jesús Crespo Cuaresma†,‡,§,¶,††

†*Faculty of Economics and Administration, Masaryk University, Lipová 41a, 602 00 Brno, Czech Republic (e-mail: jan.capek@econ.muni.cz)*
‡*Department of Economics, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria (e-mail: jcrespo@wu.ac.at)*
§*International Institute of Applied System Analysis (IIASA),*
¶*Wittgenstein Center for Demography and Global Human Capital (IIASA, VID/OEAW, WU),*
††*Austrian Institute of Economic Research (WIFO)*

Abstract

We analyse the role played by data and specification choices as determinants of the size of the fiscal multipliers obtained using structural vector autoregressive models. The results, based on over twenty million fiscal multipliers estimated for European countries, indicate that many seemingly harmless modelling choices have a significant effect on the size and precision of fiscal multiplier estimates. In addition to the structural shock identification strategy, these modelling choices include the definition of spending and taxes, the national accounts system employed, the use of particular interest rates or inflation measures, or whether data are smoothed prior to estimation. The cumulative effects of such arguably innocuous methodological choices can lead to a change in the spending multipliers of as much as 0.4 points.

I. Introduction

The estimation of fiscal multipliers (the ratio of the change in output to an exogenous change in government spending or taxes) is a central element for the evaluation of the macroeconomic effects of fiscal policy. Fiscal multipliers can be communicated and compared easily across different countries and time periods and the precision of their estimation contributes significantly to the quality of GDP growth predictions (Blanchard and Leigh, 2013). Since the work of Fatás and Mihov (2001) and the seminal contribution by Blanchard and Perotti (2002), empirical estimates of fiscal multipliers tend to rely on vector autoregressive

JEL Classification numbers: E62, C32.
*The authors thank the editor, Jonathan Temple, an anonymous referee and participants in research seminars at the Norwegian School of Economics, the University of Innsbruck, the University of Sankt Gallen, the Paris Lodron University of Salzburg, the Bank of Lithuania and ifo Munich, as well as participants in the Austrian Economic Association Meeting 2019 in Graz, the XXI Applied Economics Meeting in Alcalá de Henares and the Slovak Economic Association Meeting 2017 in Košice for helpful comments on an earlier draft of the paper. Financial support from the Czech Science Foundation, Grant 17-14263S, is gratefully acknowledged. Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the programme 'Projects of Large Research, Development, and Innovations Infrastructures'. Open Access funding enabled and organized by Projekt DEAL.

(VAR) models, with the current literature still demonstrating a widespread interest in the computation of such measures and the use of credible identification techniques to ensure the exogeneity of fiscal shocks in the framework of the estimation method. While the long time series available for the US allow for the use of narrative methods to identify exogenous shocks (Ramey, 2011) or the assessment of different regimes (Auerbach and Gorodnichenko, 2012), estimates based on shorter time series for European, Latin American, or African countries still rely on less sophisticated methods (Estevão and Samaké, 2013; Muir and Weber, 2013; Petrevski, Bogoev and Tevdovski, 2015).[1] Existing fiscal multiplier estimates (even using the same broad methodology, country, and time period) are notoriously heterogeneous. Some reasons for the differences across estimates have already been addressed in the literature, which has emphasized the role of institutional settings or the asymmetry of fiscal multipliers in different business cycle phases.

Our contribution aims to assess how the size and precision of fiscal multipliers obtained using structural VAR (SVAR) models depend on the different methodological choices that need to be made when specifying them. Rather than working on the results from the existing empirical literature on fiscal multipliers, we obtain the multiplier estimates ourselves, changing the data source and model settings in order to explore the determinants of the size and precision of the estimated multipliers. Using data for European countries, we estimate SVAR models that mimic different settings used in the empirical literature with respect to the particular specification of the model, data transformations and identification strategies. Making use of the estimated SVAR models, we obtain fiscal multipliers and assess how the size and precision of the multipliers depend on the particular characteristics of the modelling framework. Admittedly, the shorter time series available for European countries as compared to the US constrains the choice of modelling tools, but the literature which deals with the estimation of fiscal multipliers has used modelling tools such as those entertained here for countries with even shorter time series. A systematic investigation of the role of methodological choices on the size of fiscal multipliers for SVAR models appears thus justified for countries for which relatively short spans of data are available.[2]

Gechert (2015) and Rusnák (2011) present meta-analyses of the literature on fiscal multipliers that share some common ground with the research question posed in this piece. These contributions assess, among other aspects, the influence of the identification strategy for structural shocks, the effect of the number of variables in the VAR, the horizon at which the multiplier is reported, and the effect of sample size. However, a systematic analysis of the role played by data composition, data transformations, the methodology of fiscal data collection or the specific formulation of the reduced-form VAR model is absent in the existing literature, either in meta-analysis pieces like Gechert (2015) and Rusnák (2011), or in more systematic empirical approaches like Caldara and Kamps (2008). There are several reason for the missing evidence. From the meta-analysis perspective, there are so many possible combinations of these characteristics that there are simply not enough studies yet to have been able to cover the variability needed to identify their effects on the

[1] See e.g. Ramey (2016) for a modern review of the methods used for the identification of exogenous fiscal shocks.

[2] Canova and Pappa (2007), Estevão and Samaké (2013), Muir and Weber (2013), Perotti (2004), and Petrevski *et al.* (2015) are examples of studies reporting fiscal multipliers estimated on time series of 24–56 observations.

estimates of fiscal multipliers. In addition, from the perspective of the practitioner, some of these characteristics are often considered innocuous and do not tend to be reported in the published pieces.

Our results indicate that many seemingly inconsequential choices affect the value of the estimated multipliers as well as the precision with which they are estimated. For example, spending multipliers obtained using HICP to deflate nominal variables (instead of a GDP deflator) and following the European System of National and Regional Accounts (ESA) 95 rather than ESA 2010 tend to be significantly larger (by 0.122 and 0.119, respectively). The results demonstrate that data composition for government spending and government revenue play a role as well, leading to changes in the estimated multiplier by as much as 0.126 for spending in a group of western European countries and 0.189 for tax cut multipliers in a group of eastern European countries. We show that the way data are transformed prior to estimation also affects the size of the multiplier estimates, as well as the choice of identification strategy and the number of variables in the VAR model. Furthermore, the effect of some of these modelling choices appears different in western versus eastern European economies and in spending versus tax multipliers. The inclusion of data corresponding to the financial crisis period also has an effect on fiscal multiplier estimates, with the evidence presented supporting the existence of larger spending multipliers since the beginning of the current decade. In eastern European countries, this increase can be as much as 0.3, whereas the results for western European countries show an increase of 0.2.

Apparently unimportant methodological choices can lead to sizeable differences in multiplier estimates. Changing the source of the data, the deflator and the definition of government revenues and spending, for instance, leads to spending multiplier estimates that differ by 0.4 on average, irrespective of the identification scheme used to extract structural shocks. An implication of our analysis is that, when structural VARs are used to estimate fiscal multipliers, it is important for researchers to document their choices in detail, even for aspects of the research design that may seem innocuous.

The rest of the paper is organized as follows. Section II presents the methodology of the analysis in detail, section III reports the results for the determinants of differences in estimates of the fiscal multipliers and section IV analyses the determinants of differences in their precision. Section V concludes.

II. Estimating fiscal multipliers: The SVAR framework

Ever since the work of Blanchard and Perotti (2002), methodological frameworks that build upon SVAR specifications have become the workhorse for the estimation of fiscal multipliers. Abstracting from further deterministic terms, the estimation of the fiscal multiplier is based on the following reduced-form VAR model,

$$A(L)Y_t = u_t, \tag{1}$$

where Y_t is a K-dimensional vector containing output, fiscal variables and other covariates, $A(L) \equiv I_K - \sum_{j=1}^{p} A_j L^j$ denotes the autoregressive lag polynomial, where $A_j, j = 1,\ldots,p$ are $K \times K$ matrices and u_t is a vector of potentially correlated error terms with a variance-covariance matrix given by $\Sigma_u \equiv E(u_t u_t')$. In order to obtain the fiscal multiplier, we need to recover structural uncorrelated shocks ε_t. Pre-multiplying equation (1) with a convenient matrix A_0 results in the structural form of the VAR model,

$$B(L)Y_t = B\varepsilon_t, \tag{2}$$

where $B(L) = A_0A(L)$ and

$$A_0u_t = B\varepsilon_t \tag{3}$$

describes the relation between the reduced-form errors u_t and structural disturbances ε_t. With a proper choice of A_0 and B, ε_t has a diagonal covariance matrix Σ_ε and the structural shocks are uncorrelated with one another.

Various identification methods can be used to retrieve the structural shocks in ε_t. The method pioneered by Blanchard and Perotti (2002) relies on exact restrictions through a recursive identification scheme based on lags in the implementation of fiscal policy, while more recent methods (Rubio-Ramírez, Waggoner and Zha, 2010) use sign restrictions that constrain the direction of the response of variables to particular shocks. Once the structural shocks have been identified, government spending multipliers and tax cut multipliers can be computed. In line with recent literature (e.g. Ilzetzki, Mendoza and Végh, 2013; Gechert and Rannenberg, 2014; Caggiano *et al.*, 2015), we concentrate on discounted cumulative multipliers, defined as

$$m^s = \frac{\sum_{t=0}^{T}(1+i)^{-t}\Delta y_t}{\sum_{t=0}^{T}(1+i)^{-t}\Delta g_t}, \tag{4}$$

where i is the (average) interest rate, which we set to 1% per quarter for our computation,[3] y_t is output at time t, g_t denotes government expenditures at time t, Δ denotes the deviation from the respective baseline, and T is the horizon at which the multiplier is computed. Unless otherwise stated, the multipliers are reported for $T = 4$ in the context of data at quarterly frequency.[4] The superscript on m denotes the type of multiplier, m^s being the spending multiplier. Tax cut multipliers m^τ are calculated similarly, only with an increase in (net) taxes $\Delta\tau_t$ in the denominator of equation (4) and a switched sign in the reaction of output, $-\Delta y_t$, in the numerator.

As compared to log-level models, first-differenced VAR specifications are rarely used in the literature on the estimation of fiscal multipliers after the contribution by Blanchard and Perotti (2002) and are not included in our analysis. Blanchard and Perotti (2002) report significant differences between the log-level and first difference settings and aim at accounting for time-varying drift terms by subtracting a changing mean, constructed as the geometric average of past first differences, with a decay parameter equal to 2.5% per quarter. Due to the lack of direct comparability between the standard SVAR models in log levels and the VAR models in first differences after accounting for this particular adjustment, we decided to exclude them from the exercise.

[3] The interest rate corresponds to 4% *per annum* and means that the corresponding discount factor in the quarterly frequency is 0.99. While we concentrate on discounted cumulative multipliers in our analysis, results for different definitions of the fiscal multiplier for selected countries do not lead to qualitatively different conclusions.

[4] The results for horizons below $T = 4$ are qualitatively similar to those found for the one year horizon, although the effects of data and methodology tend to be weaker, a conclusion that is expected from a theoretical point of view and confirms the results in Gechert (2015).

Fiscal multipliers estimated in SVAR frameworks are the outcome of numerous data, modelling, and methodological choices. These choices can be separated into several categories: (i) the group of macroeconomic variables included in the SVAR model, (ii) the definition of the government spending and tax variables, as well as other macroeconomic covariates, (iii) the existence of data preprocessing related to smoothing of certain variables, (iv) the specification of the VAR model in terms of the inclusion of deterministic terms and the choice of lag length, and (v) the identification strategy for structural shocks. Below we describe the various data transformation and modelling choices used in the existing literature, which will be addressed in our empirical analysis.

Macroeconomic variables in the VAR model

The most used specifications in the empirical literature on the estimation of fiscal multipliers are VAR models with three variables (government expenditures, government revenues, and output), following the model put forward by Blanchard and Perotti (2002), and VAR models with five variables (the former three plus inflation and interest rate) following for instance the work of Perotti (2004). Although some other papers have enriched these basic settings with additional variables, we stick to these variable choices when assessing the effect of covariate choices on fiscal multipliers.

Definition and source of fiscal and other macroeconomic variables

Prior to the estimation of the model, the variables measuring government spending and/or revenues need to be defined based on their expected effect on output. Some contributions in the literature of fiscal multipliers adjust government spending and/or revenue for components that are not under direct control of the government. This adjustment mainly concerns automatic stabilizers such as social transfers but may also involve other components, like interest payments and subsidies. Crespo Cuaresma, Eller and Mehrotra (2011) and Muir and Weber (2013) offer a comprehensive treatment of the construction of fiscal variables for use in SVAR models.

Existing studies based on European countries also differ in the source of the fiscal data. Recent studies tend to use variables based on the European System of Accounts 2010 (ESA 2010), whereas older papers follow the ESA 95 methodology. Similarly, inflation is calculated employing the GDP deflator in some studies, while others compute it based on changes in the harmonized index of consumer prices (HICP). In addition, one finds inflation definitions based on year-on-year changes as well as on quarter-on-quarter rates of change. The maturity used for the interest rate also differs across studies, as does the source employed to retrieve the interest rate data.

Data preprocessing

The standard data source for the macroeconomic variables used in studies about fiscal multipliers in European economies, Eurostat, does not publish seasonally adjusted quarterly government data and only provides nominal values. Authors using these figures to obtain fiscal multipliers typically use seasonal adjustment procedures based on the TRAMO/

SEATS or X11 methods prior to the analysis. However, some studies also apply data smoothing with moving averages for seasonal adjustment (Klyviene and Karmelavičius, 2012) or for reasons related to the potential existence of outliers (Crespo Cuaresma *et al.*, 2011). Depending on the study, the published nominal data are deflated using a GDP deflator or a consumer price index.

Specification of the VAR model: deterministic terms and lag length

The specific form of the model given by equation (1) which is actually estimated varies across studies when it comes to the deterministic terms and lag length. While some models use deterministic linear time trends in addition to the intercept, others stick to a basic specification with the intercept term only. Furthermore, some studies add dummy variables that control for specific time periods of non-systematic behaviour like military buildup periods or for selling Universal Mobile Telecommunications System licenses. Due to the large number of estimated models, we use an automated approach to outlier detection to assign dummies. In particular, the time series of government spending and taxes are checked for outliers using seven different tests.[5] If five or more tests identify an outlier, a dummy that identifies it is added as a deterministic term to equation (1) when specifying it. In our analysis, since the frequency of the data is quarterly, the lag length of the VAR model is allowed to be one to four lags.

Identification strategy for structural shocks

The bulk of the literature on the estimation of fiscal responses based on SVAR models relies on three identification strategies to retrieve structural shocks: (i) recursive identification based on the Cholesky decomposition of the variance–covariance matrix of the reduced-form VAR shocks Σ_u, (ii) imposing restrictions on the A_0 and B matrices in equation (3) based on the elasticities of government purchases and taxes to output, in the spirit of Blanchard and Perotti (2002) (BP) and (iii) identification based on sign restrictions.

In shock identification designs based on recursive schemes, the order in which the variables enter the VAR model is the only aspect that matters to identify the shocks. The shock ordered first is assumed not to react contemporaneously to any other shocks in the system. The second shock reacts only to the first shock, while the last shock reacts contemporaneously to all shocks in the system. For a standard 3-variable VAR model, equation (3) takes the form

$$\begin{bmatrix} 1 & 0 & 0 \\ -\alpha_{yg} & 1 & 0 \\ -\alpha_{\tau g} & -\alpha_{\tau y} & 1 \end{bmatrix} \begin{bmatrix} u_t^g \\ u_t^y \\ u_t^\tau \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^g \\ \varepsilon_t^y \\ \varepsilon_t^\tau \end{bmatrix}, \tag{5}$$

where g denotes government expenditures, y output, and τ taxes. Therefore, for the case of recursive identification, $B = I$ and A_0 is a lower triangular matrix. Consequently, A_0^{-1} is also lower triangular, which implies that the Cholesky decomposition of the variance-

[5] The tests are based on the adjusted boxplot (Brys, Hubert and Rousseeuw, 2005), Grubbs' procedure (Grubbs, 1969), the moving window filtering algorithm (Brownlees and Gallo, 2006), the generalized ESD procedure, the modified Z-score method, and the interquartile range test (see Iglewicz and Hoaglin, 1993, for the last three methods).

covariance matrix Σ_u can be used for identification. Solving equation (3) for u_t, substituting to $\Sigma_u = E(u_t u_t')$, and setting $B = I$ results in

$$\Sigma_u = A_0^{-1} \Sigma_\varepsilon \left(A_0^{-1}\right)'. \tag{6}$$

The Cholesky decomposition of the variance-covariance matrix of the reduced-form residuals $\Sigma_u = PP'$ yields a lower triangular matrix P. If Σ_ε is not normalized, its Cholesky decomposition $\Sigma_\varepsilon = DD'$ provides the diagonal matrix D with the standard deviations of the structural shocks on the main diagonal. Following these two decompositions, $P = A_0^{-1} D$, which implies that A_0^{-1} is known once we account for (possible) non-unit standard deviations of the structural shocks stored in D.

The structural identification approach introduced in Blanchard and Perotti (2002) has been extremely influential in the modern literature on fiscal multipliers. It relies on institutional information about tax and transfer systems and about the timing of tax collections in order to identify the structural shocks ε_t. Sticking to the example of a 3-variable VAR, equation (3) takes the form

$$\begin{bmatrix} 1 & 0 & 0 \\ -\alpha_{yg} & 1 & -\alpha_{y\tau} \\ 0 & -1.85 & 1 \end{bmatrix} \begin{bmatrix} u_t^g \\ u_t^y \\ u_t^\tau \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \beta_{\tau g} & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^g \\ \varepsilon_t^y \\ \varepsilon_t^\tau \end{bmatrix}, \tag{7}$$

where the specific output elasticity of government revenue ($\alpha_{\tau y} = 1.85$) is adopted from Perotti (2004). In a 5-variable setting that includes inflation and the interest rate as additional variables, other elasticity values need to be fixed in order for the system (7) to be identified. Several variations of elasticity values found in Caldara and Kamps (2008) and Crespo Cuaresma *et al.* (2011) are used in the empirical analysis presented below. Generally, in the Blanchard and Perotti (2002) approach, A_0 is not lower triangular and B is not an identity matrix. In the typical setting, the concentrated log-likelihood corresponding to the VAR model can be maximized with respect to the free parameters in A_0 and B, yielding the estimates of these matrices.[6]

The sign restriction approach imposes conditions directly on the shape of the impulse response functions corresponding to the VAR model. Mountford and Uhlig (2009) and Caldara and Kamps (2008) propose restrictions that imply that business cycle shocks are identified by the positive reaction of both taxes and output, tax cut shocks are identified by the negative reaction of taxes and spending shocks by the positive reaction of spending. All of these restrictions are assumed to hold for four quarters. While one strand of literature follows the penalty function approach introduced in Uhlig (2005) and Mountford and Uhlig (2009), recent approaches employ an algorithm based on rotation matrices (see e.g. Canova and Pappa, 2007; Rubio-Ramírez *et al.*, 2010; Arias, Rubio-Ramirez and Waggoner, 2018). The algorithm used in our implementation of this identification strategy makes use of the so-called QR-factorization and relies on 300 solutions that fulfil the required sign restrictions.

[6]Alternatively, some authors use a two-step procedure, starting with the estimation of cyclically adjusted taxes and government expenditures.

III. Fiscal multipliers: Methodological determinants

Using all possible combinations of the methodological choices described above, we estimate SVAR models for all the EU-28 economies as well as for Switzerland, Norway and Iceland. The data, with quarterly frequency, are sourced from Eurostat and typically span the period 1999–2014 (subject to availability). For each model, we simulate 300 multipliers based on the distribution of the estimate and work with the median multiplier m_{median} as well as with the range between 16th and 84th percentiles $m_{16-84pr}$, which will serve as a measure of uncertainty.[7] The total number of estimated fiscal multipliers is therefore 26,373,098 for each one of the horizons evaluated.

We concentrate on analysing the fiscal multipliers obtained from models that (i) are stable, (ii) are among the best models according to information criteria, and (iii) are among the models least burdened by residual autocorrelation. An estimated model is considered stable if the maximum eigenvalue modulus of the VAR is below unity. Model selection criteria are computed for all estimated models and residual autocorrelation is tested using the Ljung–Box Q test. We order all our models by selection criteria using the Schwarz, Akaike and Hannan–Quinn criteria as well as Ljung–Box statistics and concentrate exclusively on the 10% best models according to this ordering. In particular, we record for each model the share of Q-tests which do not imply a rejection of the null hypothesis of autocorrelation for all variables at lags 4, 8 and 20 and the number of times the model is chosen as a best model within the class of comparable VAR specifications using the three selection criteria mentioned above. We select the top 10% models in these two dimensions.

By concentrating on a selected group of specifications in the baseline setting, we favour economic interpretation over the completeness of the set of all possible multipliers obtained by combining modelling options. Such a selection appears in line with the typical workflow for estimating multipliers in empirical studies. We also evaluate the importance of verification and model selection measures by relaxing the requirements (i)–(iii) and thus increasing the number of multipliers used for analysis.[8] The results of the baseline regressions are not significantly affected by estimating them with these expanded samples.

Table 1 shows the descriptive statistics of the median multipliers, as well as of the 16th–84th percentile range for the selected models (2,540,877 of them). The vast majority of the estimated multipliers have sensible values. The spending multipliers m^s seem generally higher in absolute value than the tax cut multipliers and less precisely estimated. The minimum number of observations used to estimate them is 27, while the most common number of observations is 43.

In order to quantify the effect of methodological choices on the multiplier values and dispersion of the estimates, we employ a meta-regression (Stanley and Jarrell, 2005, eq. 3):

$$m = \alpha + \beta_c D_c + \beta_m D_m + v, \tag{8}$$

[7] In sign restriction identification schemes, the 300 solutions are the actual draws. Other identification approaches rely on bootstrapping to compute the 300 draws.

[8] These alternative settings expand the number of observations of our baseline regression models ($N = 2,540,877$), to $N = 8,688,247$; 14,221,717; 22,972,983; and 25,015,940, depending on the set of conditions that the multipliers are assumed to fulfil. The online Appendix S1 presents the results for the regression based on the broad set of 22,972,983 multipliers.

TABLE 1

Descriptive statistics of multiplier medians and percentiles in the subgroup of 'best' models, $N = 2,540,877$

	Minimum	*5-th p.*	*16-th p.*	*Median*	*84-th p.*	*95-th p.*	*Maximum*
m^s_{median}	−115.53	−3.82	−1.67	0.07	1.97	4.61	112.21
m^τ_{median}	−72.14	−2.63	−1.31	−0.33	0.21	0.91	118.67
$m^s_{16-84pr}$	0.05	0.92	1.60	4.06	11.61	24.72	740.41
$m^\tau_{16-84pr}$	0.02	0.23	0.42	1.33	4.23	9.02	458.78
Observations	27	32	34	43	58	69	136

where m is a vector containing all multipliers (or alternatively, the dispersion measure), D_c is a matrix whose columns are dummies identifying the different countries, D_m is a matrix that collects dummies related to data transformations, modelling details and structural identification procedures, and v is a vector error term. The meta-regression model given by equation (8) is estimated using weighted least squares (WLS) with weights given by the inverse of the variance of the estimates for models where the dependent variable is m^s_{median} or m^τ_{median} and with the standard least squares method for meta-regressions of multiplier ranges $m^s_{16-84pr}$ or $m^\tau_{16-84pr}$. The results of the estimations are reported in Tables 3–8. Since the main aim of our study is to quantify the role of methodological choices as a determinant of differences in the size of the estimated fiscal multiplier, we do not report the coefficient estimates for the country fixed effects β_c in equation (8) in the tables.[9]

The results are reported for the full set of countries as well as for two subgroups of economies, with the aim of investigating possible differences in the relationship between modelling choices and multiplier size within the core countries that joined the European Union prior to 2004 as compared to the eastern European economies that have formed part of the EU since 2004 and used to be centrally planned economies (see Appendix A for the identity of the countries in each group). In the spirit of Ilzetzki *et al.* (2013), we try to construct both groups in a way that ensures a higher degree of homogeneity in economic structure within the country groups than when assessing the full group of European economies for which data are available.

Since the predictors are only dummies, the coefficients have the simple interpretation of a change in the multiplier for deviations from the baseline setting. In the specification used, the baseline setting is chosen on the basis of corresponding to the most common case in the existing literature. Table 2 lists the baseline setting and various alternative settings investigated.

Table 3 presents the results for the effects of variable definitions, data source, VAR specification and identification based on the median of spending multiplier m^s_{median} at horizon $T = 4$.[10] In this setting, we entertain fiscal multipliers based on a single choice of inflation and interest rates (the benchmark one) for the 5-variable VAR models. We assess the potential differences in fiscal multipliers based on the different choices of interest rate

[9] The estimates of the country fixed effects are available in the online Appendix S1.

[10] We only present in our tables coefficient estimates for selected regressors, the online Appendix S1 contains the results for the full set of estimates.

TABLE 2

Baseline and alternative settings for regression models

Baseline specification	*Alternative specification/s*
Nominal variables deflated by GDP deflator	Nominal variables deflated by HICP
European System of Accounts (ESA) 2010	Older ESA 95
Revenues definition: total revenues less interest payments, transfers, and social contributions	Several different revenues definitions
Spending definition: total spending less transfers and social contributions	Several different spending definitions
No smoothing of data	Fiscal data (and GDP) smoothed using MA(3) or MA(5)
Identification of a 3-variable VAR with Cholesky ordering	Identification of 3- and 5-variable VARs with Cholesky, sign restrictions, and BP with various elasticities
Outliers in fiscal time series detected and shift/jump dummies added	Possible outliers in the fiscal time series ignored
Constant but no trend in the VAR	Constant + time trend in the VAR
VAR with 4 lags	VAR with 1, 2, or 3 lags
Full time sample	Time sample ends in 2008 or 2010
Inflation rate based on GDP deflator (quarter-on-quarter, annualized)	Deflator inflation computed year-on-year and HICP inflation computed as both q-o-q and y-o-y
Interest rate: Maastricht criterion bond yields (long term)	3-month and 6-month interbank rates

and inflation measures in 5-variable VARs in an additional regression model whose results are presented in Table 4.

We start by discussing the results that appear significant and robust to the choice of country groups. Data source and methodological choices have significant effects on the size of the estimated multipliers which can be very important in magnitude. If the nominal variables are not deflated with a GDP deflator but with the HICP index, the estimated spending multiplier increases on average by 0.122. If the European System of Accounts (ESA) 95 is used, this leads to a median value of the multiplier that is higher on average by 0.119. The definition of revenues and spending used to calculate the multipliers also appears to affect the size of the multiplier. The baseline for these data composition choices (see Table 2) is similar: for both series, we subtract transfers and social contributions. In the case of revenues, we also subtract interest payments. If the researcher instead follows the definition of revenues in Crespo Cuaresma *et al.* (2011) or defines spending as total spending less interest payments, the value of the multiplier is on average higher by 0.112 or 0.041, respectively. The smoothing of fiscal data with a moving average filter, in addition, leads on average to a significant but small decrease in the estimated multiplier.

Turning to the effects of the structural shock identification strategies, here the results show strong variation with respect to the choice of country groups. The sign restrictions approach for both 3-variable and 5-variable VAR and the Blanchard and Perotti (2002) approach lead to very different results for a group of western economies as compared to eastern European countries. Also, the 5-variable approach, which includes the interest rate and inflation, generally leads to higher multiplier values than the 3-variable approach,

TABLE 3

Determinants of spending multiplier m^s_{median}: Regression results

Predictor	*Country subgroup*		
	All	*West*	*East*
(a) Variable definitions, data source & transformations			
Nominal variables deflated by HICP	0.122***	0.010***	0.107***
	(0.0025)	(0.0034)	(0.0040)
ESA 95 used	0.119***	0.092***	0.083***
	(0.0024)	(0.0033)	(0.0040)
Revenues following Crespo Cuaresma *et al.* (2011)	0.112***	0.126***	0.065***
	(0.0039)	(0.0052)	(0.0067)
Total spending less interest payments	0.041***	0.079***	0.108***
	(0.0034)	(0.0044)	(0.0060)
Fiscal data smoothed with moving average of length 5	−0.045***	−0.027***	−0.028***
	(0.0041)	(0.0056)	(0.0070)
(b) Structural identification			
5-variable VAR identified with Cholesky decomposition	0.113***	0.046***	0.147***
	(0.0041)	(0.0050)	(0.0080)
5-variable VAR identified with sign restrictions	0.320***	−0.061***	0.836***
	(0.0106)	(0.0132)	(0.0182)
5-variable VAR identified with BP (elasticities from Caldara and Kamps, 2008)	−0.058***	−0.130***	0.518***
	(0.0129)	(0.0136)	(0.0349)
5-variable VAR identified with BP (elasticities from Crespo Cuaresma *et al.*, 2011)	−0.176***	−0.309***	0.471***
	(0.0160)	(0.0169)	(0.0431)
(c) VAR specification and sample			
Constant + time trend in the VAR	−0.123***	−0.174***	0.062***
	(0.0025)	(0.0033)	(0.0043)
VAR with 1 lag	−0.103***	−0.133***	−0.061***
	(0.0063)	(0.0083)	(0.0114)
VAR with 2 lags	−0.094***	−0.160***	−0.047***
	(0.0057)	(0.0074)	(0.0106)
Sample ends in 2008	−0.105***	0.039***	−0.302***
	(0.0032)	(0.0042)	(0.0059)
Sample ends in 2010	−0.146***	−0.218***	−0.178***
	(0.0036)	(0.0047)	(0.0069)
Observations	420,986	218,791	132,054
Number of regressors in model	61	45	39
R^2	0.47	0.30	0.46

Notes: Estimates correspond to the specification in equation (8). Dependent variable is the estimated multiplier. All covariates are dummy variables, baseline specification given in Table 2. 'All': all countries in the sample, 'West': western European countries, 'East': eastern European countries. ***, **, * denotes significance at 1%, 5%, 10% level, respectively. Standard errors in parentheses. Estimation by WLS with inverse variance as weight. Country fixed effects in all specifications. Parameter estimates reported if significant in at least one of the country group samples.

although this result depends on the choice of calibrated elasticities. Identifying shocks by means of Cholesky ordering using the 5-variable specification instead of the 3-variable specification, for instance, leads to an average increase of 0.113 in the estimated multiplier. The results also show that using fewer lags than four in the VAR specification leads to a decrease in the estimated multiplier. The results for estimates based on data prior to the

TABLE 4

Determinants of spending multiplier m^s_{median}, selected results for VAR models based on five variables

Predictor	*Country subgroup*		
	All	*West*	*East*
Variable definitions and data source: 5-variable VAR			
Deflator inflation, year-on-year	0.051***	−0.022***	0.111***
	(0.0014)	(0.0021)	(0.0018)
HICP inflation, year-on-year	0.007***	−0.011***	0.061***
	(0.0016)	(0.0023)	(0.0020)
HICP inflation, quarter-on-quarter, annualized	0.049***	0.024***	0.082***
	(0.0012)	(0.0016)	(0.0018)
3-month interbank rate	−0.246***	0.014***	−0.494***
	(0.0019)	(0.0027)	(0.0024)
6-month interbank rate	−0.259***	−0.012***	−0.466***
	(0.0019)	(0.0027)	(0.0024)
Observations	2,318,268	1,137,774	990,406
Number of regressors in model	60	48	41
R^2	0.41	0.30	0.63

Notes: Estimates correspond to the specification in equation (8). Dependent variable is the estimated multiplier. All covariates are dummy variables, baseline specification given in Table 2. 'All': all countries in the sample, 'West': western European countries, 'East': eastern European countries. ***, **, * denotes significance at 1%, 5%, 10% level, respectively. Standard errors in parentheses. Estimation by WLS with inverse variance as weight. Country fixed effects in all specifications. Only parameter estimates for the dummies corresponding to the 5-variable VAR reported.

crisis years indicate that spending multipliers have become on average larger in the second decade of the 21st century, lending support to the hypothesis that fiscal multipliers are larger in recessions than in expansions, and were particularly large in the aftermath of the financial crisis (see e.g. Auerbach and Gorodnichenko, 2012; Gechert, Hallett and Rannenberg, 2016).

Table 4 presents the results for alternative choices of inflation and interest rate variables. Since these two variables only enter VAR specifications which contain five variables, we restrict our sample to fiscal multipliers obtained in these specifications. A standard set of predictors similar to those in the specifications reported in Table 3 was used, but we only report the estimates corresponding to the choice of data on inflation and interest rates. A robust but quantitatively small increase in the size of the spending multiplier when HICP (instead of the GDP deflator) is used to calculate inflation can be observed in our exercise, with important differences across subgroups of countries. Using interbank rates in the multiplier estimation tends to decrease the estimate of the spending multiplier by almost 0.5 in eastern European countries, while the effect for western Europe is clearly smaller in absolute value and its direction depends on the maturity of the interest rate.

Although some of the values of the effects found in Tables 3 and 4 and discussed above may seem small, the joint effect of different modelling choices can lead to sizeable cumulative effects. To illustrate this, we can define two sets of sensible methodological choices that differ only in what may appear to be 'innocuous' methodological choices and report the difference in the estimate of the spending multiplier.[11] For example, starting from the baseline specification (see Table 2) we define a scenario where the econometrician uses data based on ESA 2010, defines revenue as total revenues less interest payments, transfers, and social contributions, defines spending as total spending less transfers and social contributions, and deflates nominal variables by the GDP deflator. Compared to a scenario with ESA 95, revenue defined as in Crespo Cuaresma *et al.* (2011), spending defined as total spending less interest payments, and nominal variables deflated by HICP, the spending multiplier at the 4-quarter ahead horizon would be larger on average by 0.394.

Table 5 shows the estimation results for tax cut multipliers in the same structure as in Table 3. The absolute value of the parameter estimates for tax cut multipliers is generally smaller than that of their spending counterparts, which is in line with the smaller variability found in tax cut multipliers (see Table 1). The data composition definitions play a major role as determinants of differences in the size of tax cut multipliers. Switching from the baseline revenue composition to the one introduced in Crespo Cuaresma *et al.* (2011), or from the baseline spending composition to total spending less interest payments, both increase the estimate of the tax cut multiplier. Note that the same qualitative results (which are robust across country groups) are obtained for the spending multiplier. Smoothing the fiscal data decreases the estimate of the tax cut multiplier on average by 0.134. In general, higher tax cut multipliers are obtained if specifications based on VAR models with five variables are used. Specifically, the multiplier increases by 0.160 after adding inflation and the interest rate to the baseline setting. The results for the parameter estimates attached to the dummies that identify subsample stability reveal varied results when different spans of time are considered in the sample. If the estimation period ends before the onset of the Great Recession, the tax cut multipliers tend to be higher (by 0.132), while if the time period ends close to the trough of the recession, the multipliers tend to be lower (by −0.098). This contrasts with the results obtained for the spending multiplier, which imply lower fiscal multipliers when using data prior to the crisis. Table 6 shows that, unlike in the case of spending multipliers in Table 4, the effects of changing the method of inflation calculation or the interest rate used do not affect the tax cut multiplier substantially, with small effects found for all methodological differences studied.

The results in Tables 3–6 unveil magnitudes of the effect of methodological changes which differ strongly across multiplier type. There are also several results that are common for both spending and tax cut multipliers and also robust to country group selection. As an example of the size of such effects, the use of the revenue definition adopted from Crespo Cuaresma *et al.* (2011) and total spending less interest payments increases both the spending and tax cut multiplier. On the other hand, smoothing fiscal data decreases both the spending multiplier (by 0.045) and tax cut multiplier (by 0.134). While using the Cholesky identification strategy, introducing inflation and the interest rate to the original three variables in the VAR increases the spending multiplier by 0.113 and the tax cut multiplier by 0.160.

[11] In order to illustrate only robust results across European economies, we do not employ choices that lead to a significant change in the multipliers in only a subset of countries.

TABLE 5

Determinants of tax cut multiplier m^{τ}_{median}: Regression results

Predictor	*Country subgroup*		
	All	*West*	*East*
(a) Variable definitions, data source & transformations			
Nominal variables deflated by HICP	−0.024***	−0.037***	0.005**
	(0.0011)	(0.0013)	(0.0023)
ESA 95 used	0.005***	0.016***	−0.037***
	(0.0011)	(0.0013)	(0.0024)
Revenues following Crespo Cuaresma *et al.* (2011)	0.058***	0.044***	0.189***
	(0.0022)	(0.0025)	(0.0056)
Total spending less interest payments	0.033***	0.028***	0.055***
	(0.0014)	(0.0017)	(0.0031)
Fiscal data is smoothed with moving average of length 5	−0.134***	−0.142***	−0.103***
	(0.0017)	(0.0020)	(0.0041)
(b) Structural identification			
5-variable VAR identified with Cholesky decomposition	0.160***	0.158***	0.268***
	(0.0017)	(0.0020)	(0.0039)
5-variable VAR identified with sign restrictions	0.007	0.050***	−0.028**
	(0.0048)	(0.0055)	(0.0117)
5-variable VAR identified with BP (elasticities from Caldara and Kamps, 2008)	0.040***	0.061***	0.051**
	(0.0066)	(0.0071)	(0.0225)
5-variable VAR identified with BP (elasticities from Crespo Cuaresma *et al.*, 2011)	0.165***	0.166***	0.253***
	(0.0021)	(0.0024)	(0.0046)
(c) VAR specification and sample			
Constant + time trend in the VAR	−0.012***	−0.025***	0.001
	(0.0011)	(0.0013)	(0.0023)
VAR with 1 lag	0.024***	0.016***	0.069***
	(0.0034)	(0.0042)	(0.0071)
VAR with 2 lags	0.008***	−0.002	0.103***
	(0.0031)	(0.0039)	(0.0065)
Time sample ends in 2008	0.132***	0.132***	0.366***
	(0.0015)	(0.0017)	(0.0042)
Time sample ends in 2010	−0.098***	−0.082***	−0.031***
	(0.0014)	(0.0017)	(0.0030)
Observations	420,986	218,791	132,054
Number of regressors in model	61	45	39
R^2	0.62	0.53	0.69

Notes: Estimates correspond to the specification in equation (8). Dependent variable is the estimated multiplier. All covariates are dummy variables, baseline specification given in Table 2. 'All': all countries in the sample, 'West': western European countries, 'East': eastern European countries. ***, **, * denotes significance at 1%, 5%, 10% level, respectively. Standard errors in parentheses. Estimation by WLS with inverse variance as weight. Country fixed effects in all specifications. Parameter estimates reported if significant in at least one of the country group samples.

IV. The determinants of multiplier precision

Data, modelling, and methodological choices do not only affect the point estimates of the multipliers, but also their precision. Some of the methodological choices lead to a more precise estimate of the multiplier, whereas others increase the dispersion of multiplier

TABLE 6

Determinants of tax cut multiplier m^{τ}_{median}, selected results for VAR models based on five variables

Predictor	*Country subgroup*		
	All	*West*	*East*
Variable definitions and data source: 5-variable VAR			
Deflator inflation, year-on-year	−0.016	−0.015***	−0.013***
	(0.0004)	(0.0005)	(0.0007)
HICP inflation, year-on-year	−0.020***	−0.029***	−0.012***
	(0.0005)	(0.0006)	(0.0007)
HICP inflation, quarter-on-quarter, annualized	−0.019	−0.024***	−0.005***
	(0.0004)	(0.0005)	(0.0006)
3-month interbank rate	0.038***	0.019***	0.047***
	(0.0006)	(0.0008)	(0.0009)
6-month interbank rate	0.040***	0.015***	0.052***
	(0.0006)	(0.0008)	(0.0009)
Observations	2,318,268	1,137,774	990,406
Number of regressors in model	60	48	41
R^2	0.58	0.52	0.67

Notes: Estimates correspond to the specification in equation (8). Dependent variable is the estimated multiplier. All covariates are dummy variables, baseline specification given in Table 2. ‘All’: all countries in the sample, ‘West’: western European countries, ‘East’: eastern European countries. ***, **, * denotes significance at 1%, 5%, 10% level, respectively. Standard errors in parentheses. Estimation by WLS with inverse variance as weight. Country fixed effects in all specifications. Only parameter estimates for the dummies corresponding to the 5-variable VAR reported.

estimates around their median. Table 7 reports the estimation results of a regression model such as the one in equation (8) addressing the determinants of the spending multiplier dispersion (18th–84th percentile range) at horizon $T=4$.[12] The choice of whether to deflate nominal variables with a GDP deflator or HICP plays a significant role when it comes to the precision of multiplier estimates. Using HICP reduces the dispersion of the estimate of the spending multiplier, giving an estimate with higher precision. The effect is much more pronounced for the eastern European country group. A similar effect is also found for the methodological choice of ESA 95, however, this effect does not appear to exist for Western EU countries.

As for the effect of the definitions of fiscal variables, spending variables that follow Muir and Weber (2013) and Crespo Cuaresma *et al.* (2011) increase the dispersion of both spending and tax cut multiplier estimates. The results for the data smoothing choice delivers mixed results, except for the case where only fiscal time series are smoothed, which increases the dispersion of the estimates of spending multiplier. Identification strategies affect the dispersion significantly: sign restriction estimates increase the dispersion considerably, as does the Blanchard and Perotti (2002) approach applied to a 5-variable VAR. Our results indicate that including a time trend in the formulation of the VAR increases the precision of the spending multiplier estimate. As for subsample stability, the results for

[12] The results for the dispersion of the tax cut multiplier can be found in the online Appendix S1.

TABLE 7

Determinants of spending multiplier ranges $m^{s}_{16-84pr}$: Regression results

Predictor	*Country subgroup*		
	All	*West*	*East*
(a) Variable definitions, data source & transformations			
Nominal variables deflated by HICP	−0.555***	−0.228***	−2.030***
	(0.0256)	(0.0274)	(0.0611)
ESA 95 used	−0.969***	0.007	−2.537***
	(0.0260)	(0.0276)	(0.0614)
Revenues following Crespo Cuaresma *et al.* (2011)	−0.748***	−0.749***	−0.113
	(0.0418)	(0.0431)	(0.1068)
Total spending less interest payments	−1.299***	−1.352***	−1.484***
	(0.0395)	(0.0418)	(0.0962)
Fiscal data is smoothed with moving average of length 5	0.847***	0.727***	0.854***
	(0.0362)	(0.0382)	(0.0909)
(b) Structural identification			
5-variable VAR identified with Cholesky decomposition	0.281***	0.457***	0.055
	(0.0603)	(0.0641)	(0.1497)
5-variable VAR identified with sign restrictions	4.676***	4.855***	4.661***
	(0.0605)	(0.0642)	(0.1502)
5-variable VAR identified with BP (elasticities from Caldara and Kamps, 2008)	7.612***	6.989***	9.841***
	(0.0603)	(0.0641)	(0.1497)
5-variable VAR identified with BP (elasticities from Crespo Cuaresma *et al.*, 2011)	10.585***	8.821***	14.930***
	(0.0603)	(0.0641)	(0.1497)
(c) VAR specification and sample			
Constant + time trend in the VAR	−0.875***	−0.523***	−1.581***
	(0.0260)	(0.0280)	(0.0625)
VAR with 1 lag	−0.697***	−0.987***	−0.640***
	(0.0740)	(0.0797)	(0.1791)
VAR with 2 lags	−0.867***	−1.322***	−0.373**
	(0.0677)	(0.0720)	(0.1652)
Time sample ends in 2008	−1.595***	−1.402***	−0.992***
	(0.0350)	(0.0359)	(0.0969)
Time sample ends in 2010	0.432***	0.800***	0.615***
	(0.0320)	(0.0352)	(0.0776)
Observations	420,986	218,791	132,054
Number of regressors in model	61	45	39
R^2	0.27	0.28	0.30

Notes: Dependent variable is the dispersion (16th–84th percentile range) of the estimated multipliers. All covariates are dummy variables, baseline specification given in Table 2. 'All': all countries in the sample, 'West': western European countries, 'East': eastern European countries. ***, **, * denotes significance at 1%, 5%, 10% level, respectively. Standard errors in parentheses. Country fixed effects in all specifications. Parameter estimates reported if significant in at least one of the country group samples.

the spending multiplier indicate that postcrisis estimates are associated with less precisely estimated multipliers. On the other hand, the time sample that ends during the Great Recession tends to produce estimates which are characterized by lower dispersion. Tax cut multipliers (see the online Appendix S1) tend to provide similar results for the full sample, although the estimates for eastern European countries differ across multiplier types.

TABLE 8

Determinants of spending multiplier ranges $m^{s}_{16-84pr}$, selected results for VAR models based on five variables

Predictor	*Country subgroup*		
	All	*West*	*East*
Variable definitions and data source: 5-variable VAR			
Deflator inflation, year-on-year	−0.070***	−0.211***	−0.253
	(0.0202)	(0.0241)	(0.0354)
HICP inflation, year-on-year	0.637***	0.694***	0.522***
	(0.0219)	(0.0265)	(0.0382)
HICP inflation, quarter-on-quarter, annualized	0.305***	0.197***	0.413***
	(0.0172)	(0.0193)	(0.0324)
3-month interbank rate	−0.699***	−0.392***	−0.939***
	(0.0286)	(0.0369)	(0.0468)
6-month interbank rate	−0.802***	−0.444***	−0.883***
	(0.0279)	(0.0367)	(0.0444)
Observations	2,318,268	1,137,774	990,406
Number of regressors in model	60	48	41
R^2	0.26	0.25	0.28

Notes: Dependent variable is the dispersion (16th–84th percentile range) of the estimated multipliers. All covariates are dummy variables, baseline specification given in Table 2. 'All': all countries in the sample, 'West': western European countries, 'East': eastern European countries. ***, **, * denotes significance at 1%, 5%, 10% level, respectively. Standard errors in parentheses. Country fixed effects in all specifications. Only parameter estimates for the dummies corresponding to the 5-variable VAR reported.

The results in Table 8 indicate that using HICP inflation instead of GDP deflator inflation increases the dispersion of spending multipliers. Similarly, using long-term bond yields instead of interbank rates increases the dispersion of spending (and partially also tax cut) multipliers.

V. Conclusions

This paper addresses how (sometimes seemingly unimportant) data, modelling, and methodological choices can affect the estimates of fiscal multipliers obtained from SVAR models. Both spending and tax cut multipliers are sensitive to specific choices regarding the composition of government spending and revenues. The particular definition of government revenues or spending, as well as specific ways of treating the data prior to estimation, can be very influential for both spending and tax cut multipliers.

The spending multiplier is sensitive to different, seemingly innocuous, modelling and methodological choices. In particular, using HICP to deflate nominal variables (rather than a GDP deflator) and using data based on ESA 95 (instead of ESA 2010), for instance, increases the estimate of the spending multiplier by 0.122 and 0.119, respectively. We also find that the identification strategy used to isolate structural shocks matters in some cases. In cases that a causal ordering based on Cholesky decompositions or sign restriction identification are used to identify fiscal shocks in VAR models that contain inflation and the interest rate, the value of the spending multiplier tends to be larger (by 0.113 and 0.320, respectively). This qualitative result holds also for the tax cut multiplier in the case of Cholesky-based identification, which is also strongly affected by the particular values of the elasticities used when implementing the Blanchard and Perotti (2002) approach. Data choices and identification strategies are also found to have important effects on the precision of multiplier estimates. The results also point to significant heterogeneity across country groupings when comparing western European economies to their eastern European counterparts, as well as when comparing multipliers estimated with data which include the global financial crisis to those that do not. The most pronounced difference between the results for eastern and western European countries are obtained for spending multipliers estimated with models that contain inflation and the interest rate. Investigating the variation in identification strategies for such models in eastern European countries, the change in spending multiplier reaches 0.836, whereas for the western European country group, the change with respect to the baseline is negative and as low as -0.309 for these specifications.

Our analysis provides ample evidence of important quantitative effects of modelling choices on fiscal multiplier estimates. Given the central role that fiscal multipliers play in the design and evaluation of macroeconomic policy, the results of our study call for a rigorous assessment of specification uncertainty when multipliers based on estimates from SVAR models are used. Further research on how to address such uncertainty, for example, using model averaging techniques, appears necessary to advance our knowledge of the effect of fiscal shocks on the real economy.

Appendix A. Countries in full sample and country groupings

Sample	Country codes	Country names
All countries	AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MT, NL, NO, PL, PT, RO, SE, SI, SK	Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Italy, Ireland, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovenia, Slovakia, Spain, Sweden, Switzerland, United Kingdom
Western EU	AT, BE, DE, DK, ES, FI, FR, GB, GR, IE, IT, NL, PT, SE, SI	Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Ireland, Netherlands, Portugal, Slovenia, Spain, Sweden, United Kingdom
Eastern EU	BG, CZ, EE, HU, LT, LV, PL, RO, SK	Bulgaria, Czechia, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia

Final Manuscript Received: October 2019

References

Arias, J., Rubio-Ramírez, J. F. and Waggoner, D. F. (2018). ‘Inference based on SVARs identified with sign and zero restrictions: theory and applications’, *Econometrica*, Vol. 86, pp. 685–720.

Auerbach, A. J. and Gorodnichenko, Y. (2012). 'Measuring the output responses to fiscal policy', *American Economic Journal: Economic Policy*, Vol. 4, pp. 1–27.

Blanchard, O. J. and Leigh, D. (2013). 'Growth forecast errors and fiscal multipliers', *American Economic Review*, Vol. 103, pp. 117–120.

Blanchard, O. and Perotti, R. (2002). 'An empirical characterization of the dynamic effects of changes in government spending and taxes on output', *Quarterly Journal of Economics*, Vol. 117, pp. 1329–1368.

Brownlees, C. T. and Gallo, G. M. (2006). 'Financial econometric analysis at ultra-high frequency: data handling concerns', *Computational Statistics and Data Analysis*, Vol. 51, pp. 2232–2245.

Brys, G., Hubert, M. and Rousseeuw, P. (2005). 'A robustification of independent component analysis', *Journal of Chemometrics: A Journal of the Chemometrics Society*, Vol. 19, pp. 364–375.

Caggiano, G., Castelnuovo, E., Colombo, V. and Nodari, G. (2015). 'Estimating fiscal multipliers: news from a non-linear world', *The Economic Journal*, Vol. 125, pp. 746–776.

Caldara, D. and Kamps, C. (2008). *What are the Effects of Fiscal Shocks? A VAR-Based Comparative Analysis*, Technical report, ECB Working Paper No. 877.

Canova, F. and Pappa, E. (2007). 'Price differentials in monetary unions: the role of fiscal shocks', *The Economic Journal*, Vol. 117, pp. 713–737.

Crespo Cuaresma, J., Eller, M. and Mehrotra, A. (2011). 'The economic transmission of fiscal policy shocks from western to eastern Europe', *BOFIT Discussion Papers*, Vol. 2011, pp. 3.

Estevão, M. M. M. and Samaké, I. (2013). *The Economic Effects of Fiscal Consolidation with Debt Feedback*, No. 13-136, International Monetary Fund.

Fatás, A. and Mihov, I. (2001). *The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence*, Technical report, CEPR Discussion Paper No. 2760.

Gechert, S. (2015). 'What fiscal policy is most effective? A meta-regression analysis', *Oxford Economic Papers*, Vol. 67, 553–580.

Gechert, S. and Rannenberg, A. (2014). *Are Fiscal Multipliers Regime-Dependent? A Meta Regression Analysis*, Technical report September.

Gechert, S., Hallett, A. H. and Rannenberg, A. (2016). 'Fiscal multipliers in downturns and the effects of euro area consolidation', *Applied Economics Letters*, Vol. 23, pp. 1138–1140.

Grubbs, F. E. (1969). 'Procedures for detecting outlying observations in samples', *Technometrics*, Vol. 11, pp. 1–21.

Iglewicz, B. and Hoaglin, D. (1993). 'How to detect and handle outliers', ASQC basic references in quality control, Mykytka E. F. (ed.), *Vol. 16*, Milwaukee, ASQC Quality Press, pp. 87.

Ilzetzki, E., Mendoza, E. G. and Végh, C. A. (2013). 'How big (small?) are fiscal multipliers?', *Journal of Monetary Economics*, Vol. 60, pp. 239–254.

Klyviene, V. and Karmelavičius, J. (2012). 'Svar analysis of the impacts of corporate taxation on the macroeconomy of lithuania', *Ekonomika/Economics*, Vol. 91, pp. 107–124.

Mountford, A. and Uhlig, H. (2009). 'What are the effects of fiscal policy shocks?', *Journal of Applied Econometrics*, Vol. 24, pp. 960–992.

Muir, D. and Weber, A. (2013). *Fiscal Multipliers in Bulgaria: Low But Still Relevant*, Technical report, IMF Working Paper No. 13/49.

Perotti, R. (2004). *Estimating the Effects of Fiscal Policy in OECD Countries*, Technical report December, IGIER Working Paper No. 276.

Petrevski, G., Bogoev, J. and Tevdovski, D. (2015). 'Fiscal and monetary policy effects in three south eastern european economies', *Empirical Economics*, Vol. 50, pp. 415–441.

Ramey, V. A. (2011). 'Identifying government spending shocks: It's all in the timing', *Quarterly Journal of Economics*, Vol. 126, pp . 1–50.

Ramey, V. A. (2016). 'Macroeconomic shocks and their propagation', in J. B. Taylor and H. Uhlig (eds.), *Handbook of Macroeconomics, Vol. 2*, Amsterdam: Elsevier, pp. 71–162.

Rubio-Ramírez, J. F., Waggoner, D. F. and Zha, T. (2010). 'Structural vector autoregressions: Theory of identification and algorithms for inference', *The Review of Economic Studies*, Vol. 77, pp. 665–696.

Rusnák, M. (2011). *Why Do Government Spending Multipliers Differ? A Meta-Analysis*, Mimeo. Prague: CERGE-EI.

Stanley, T. D. and Jarrell, S. B. (2005). 'Meta-regression analysis: A quantitative method of literature surveys', *Journal of Economic Surveys*, Vol. 19, pp. 299–308.
Uhlig, H. (2005). 'What are the effects of monetary policy on output? Results from an agnostic identification procedure', *Journal of Monetary Economics*, Vol. 52, pp. 381–419.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix S1. Additional regression results.

We just estimated twenty million fiscal multipliers
Online Appendix

Jan Čapek[1] and Jesús Crespo Cuaresma[1,2,3,4,5]

[1]Masaryk University (MU)
[2]Vienna University of Economics and Business (WU)
[3]International Institute of Applied System Analysis (IIASA)
[4]Wittgenstein Center for Demography and Global Human Capital (IIASA,VID/OEAW,WU)
[5]Austrian Institute of Economic Research (WIFO)

Tables 1–3 correspond to Tables 3, 5, and 7 in the article, with added rows for country dummies and other regressors. Table 4 does not have a corresponding table in the article. Tables 5 and 6 report regression results for similar regressions to the ones in Tables 1 and 2, but for a much broader set of multipliers due to a more relaxed verification conditions. For the baseline specification, which reflects the intercept estimate, see Table 2 in the article. In the estimation of full set of countries and western European countries, Austria is the baseline country. In the estimation of eastern European and Baltic countries, Czechia is the baseline country.

Predictor	All	West	East
Intercept	−0.096*** (0.0124)	−0.055*** (0.0134)	−0.444*** (0.0156)
Belgium	−0.440*** (0.0118)	−0.418*** (0.0115)	
Bulgaria	−0.023* (0.0124)		0.307*** (0.0082)
Croatia	1.225*** (0.0246)		
Cyprus	−0.087*** (0.0118)		
Czechia	−0.358*** (0.0108)		
Denmark	0.144*** (0.0147)	0.208*** (0.0140)	
Estonia	0.448*** (0.0195)		0.790*** (0.0163)
Germany	−0.218*** (0.0140)	−0.138*** (0.0135)	
Greece	1.017*** (0.0113)	0.944*** (0.0110)	
Finland	0.900*** (0.0255)	0.926*** (0.0242)	
France	0.179*** (0.0133)	0.194*** (0.0128)	
Hungary	0.714*** (0.0121)		0.979*** (0.0075)
Iceland	0.312*** (0.0255)		
Ireland	0.071*** (0.0135)	0.089*** (0.0129)	
Italy	0.635*** (0.0119)	0.616*** (0.0115)	
Latvia	1.509*** (0.0158)		1.815*** (0.0123)
Lithuania	2.541*** (0.0243)		2.826*** (0.0214)
Luxembourg	−1.079*** (0.0158)		
Malta	0.372*** (0.0112)		
Netherlands	−0.420*** (0.0127)	−0.426*** (0.0122)	
Norway	−2.106*** (0.0116)		
Poland	0.139*** (0.0122)		0.386*** (0.0079)
Portugal	0.399*** (0.0110)	0.391*** (0.0107)	
Romania	1.045*** (0.0120)		1.261*** (0.0074)
Spain	−0.698*** (0.0141)	−0.626*** (0.0136)	
Sweden	0.544*** (0.0130)	0.550*** (0.0125)	
Switzerland	−1.045*** (0.0847)		

Continued on next page

Table 1 – *Continued from previous page*

Predictor	All	West	East
Slovakia	0.305*** (0.0112)		0.712*** (0.0061)
Slovenia	0.668*** (0.0180)	0.666*** (0.0172)	
United Kingdom	0.022** (0.0107)	−0.005 (0.0104)	
Nominal variables deflated by HICP	0.122*** (0.0025)	0.010*** (0.0034)	0.107*** (0.0040)
ESA 95 used	0.119*** (0.0024)	0.092*** (0.0033)	0.083*** (0.0040)
Revenues following Crespo Cuaresma et al. (2011)	0.112*** (0.0039)	0.126*** (0.0052)	0.065*** (0.0067)
Revenues following Muir and Weber (2013)	0.021*** (0.0038)	0.096*** (0.0049)	−0.144*** (0.0066)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY (subsidies, payable)	0.030*** (0.0041)	0.069*** (0.0053)	−0.018** (0.0070)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY and social contributions	0.008** (0.0038)	0.024*** (0.0049)	−0.025*** (0.0067)
Spending following Crespo Cuaresma et al. (2011)	−0.035*** (0.0046)	0.118*** (0.0061)	0.026*** (0.0076)
Spending following Muir and Weber (2013)	0.025*** (0.0042)	0.138*** (0.0054)	−0.010 (0.0073)
Spending following Crespo Cuaresma et al. (2011) including D62PAY (Social benefits other than social transfers in kind, payable)	−0.075*** (0.0040)	−0.091*** (0.0052)	0.073*** (0.0071)
Total spending less interest payments	0.041*** (0.0034)	0.079*** (0.0044)	0.108*** (0.0060)
Fiscal data is smoothed with moving average of length 3	−0.019*** (0.0036)	0.030*** (0.0049)	−0.022*** (0.0059)
Fiscal data is smoothed with moving average of length 5	−0.045*** (0.0041)	−0.027*** (0.0056)	−0.028*** (0.0070)
Fiscal data and GDP is smoothed with moving average of length 3	−0.035*** (0.0047)	−0.057*** (0.0066)	0.062*** (0.0079)
Fiscal data and GDP is smoothed with moving average of length 5	−0.041*** (0.0049)	−0.120*** (0.0072)	0.148*** (0.0078)
3-variable VAR identified with sign restrictions	−0.080*** (0.0055)	0.183*** (0.0084)	−0.290*** (0.0080)
3-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	0.003 (0.0036)	0.031*** (0.0051)	−0.061*** (0.0060)
5-variable VAR identified with Cholesky decomposition	0.113*** (0.0041)	0.046*** (0.0050)	0.147*** (0.0080)
5-variable VAR identified with sign restrictions	0.320*** (0.0106)	−0.061*** (0.0132)	0.836*** (0.0182)
5-variable VAR identified with BP with elasticities from Caldara and Kamps (2008)	−0.058*** (0.0129)	−0.130*** (0.0136)	0.518*** (0.0349)
5-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	−0.176*** (0.0160)	−0.309*** (0.0169)	0.471*** (0.0431)

Continued on next page

Table 1 – *Continued from previous page*

Predictor	**All**	**West**	**East**
No dummies for possible outliers in the fiscal time series	0.004 (0.0038)	−0.034*** (0.0047)	0.078*** (0.0067)
Constant + time trend in the VAR	−0.123*** (0.0025)	−0.174*** (0.0033)	0.062*** (0.0043)
VAR with 1 lag	−0.103*** (0.0063)	−0.133*** (0.0083)	−0.061*** (0.0114)
VAR with 2 lags	−0.094*** (0.0057)	−0.160*** (0.0074)	−0.047*** (0.0106)
VAR with 3 lags	−0.086*** (0.0055)	−0.081*** (0.0072)	−0.232*** (0.0106)
Time sample ends in 2008, before the onset of the Great Recession	−0.105*** (0.0032)	0.039*** (0.0042)	−0.302*** (0.0059)
Time sample ends in 2010, typically in a trough of the Great Recession	−0.146*** (0.0036)	−0.218*** (0.0047)	−0.178*** (0.0069)
Observations	420,986	218,791	132,054
R^2	0.47	0.30	0.46

Table 1: Determinants of spending multiplier m^s_{median}: Regression results

Notes: ***, **, * denotes significance at 1, 5, 10% level, respectively. Standard errors in parentheses. Estimation by WLS with inverse variance as weight.

Predictor	**All**	**West**	**East**
Intercept	−0.156*** (0.0075)	−0.121*** (0.0079)	−0.788*** (0.0096)
Belgium	−0.183*** (0.0068)	−0.196*** (0.0067)	
Bulgaria	−0.435*** (0.0073)		−0.091*** (0.0044)
Croatia	−1.152*** (0.0095)		
Cyprus	−0.350*** (0.0070)		
Czechia	−0.304*** (0.0068)		
Denmark	−0.466*** (0.0071)	−0.491*** (0.0070)	
Estonia	−2.219*** (0.0104)		−1.717*** (0.0086)
Germany	−0.260*** (0.0074)	−0.278*** (0.0073)	
Greece	0.108*** (0.0078)	0.077*** (0.0077)	
Finland	−0.811*** (0.0074)	−0.846*** (0.0074)	
France	−0.202*** (0.0069)	−0.230*** (0.0068)	
Hungary	0.063*** (0.0106)		0.417*** (0.0086)
Iceland	0.272*** (0.0138)		
Ireland	−0.990*** (0.0080)	−1.029*** (0.0079)	
Italy	−0.105*** (0.0068)	−0.130*** (0.0068)	
Latvia	−1.682*** (0.0081)		−1.262*** (0.0055)
Lithuania	−1.560*** (0.0115)		−1.139*** (0.0098)
Luxembourg	−0.552*** (0.0156)		
Malta	−0.141*** (0.0072)		
Netherlands	−0.409*** (0.0068)	−0.423*** (0.0068)	
Norway	−1.367*** (0.0076)		
Poland	−0.335*** (0.0068)		0.007** (0.0033)
Portugal	−0.166*** (0.0069)	−0.200*** (0.0068)	
Romania	−0.981*** (0.0097)		−0.577*** (0.0076)
Spain	−0.292*** (0.0068)	−0.322*** (0.0068)	
Sweden	−0.092*** (0.0069)	−0.129*** (0.0068)	
Switzerland	−0.629*** (0.0905)		

Continued on next page

Table 2 – *Continued from previous page*

Predictor	All	West	East
Slovakia	0.123*** (0.0077)		0.368*** (0.0051)
Slovenia	−1.548*** (0.0101)	−1.599*** (0.0099)	
United Kingdom	−0.212*** (0.0068)	−0.235*** (0.0067)	
Nominal variables deflated by HICP	−0.024*** (0.0011)	−0.037*** (0.0013)	0.005** (0.0023)
ESA 95 used	0.005*** (0.0011)	0.016*** (0.0013)	−0.037*** (0.0024)
Revenues following Crespo Cuaresma et al. (2011)	0.058*** (0.0022)	0.044*** (0.0025)	0.189*** (0.0056)
Revenues following Muir and Weber (2013)	0.151*** (0.0021)	0.106*** (0.0024)	0.333*** (0.0053)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY (subsidies, payable)	0.083*** (0.0021)	0.086*** (0.0025)	0.185*** (0.0055)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY and social contributions	−0.064*** (0.0024)	0.001 (0.0028)	−0.229*** (0.0062)
Spending following Crespo Cuaresma et al. (2011)	−0.072*** (0.0019)	−0.073*** (0.0023)	−0.098*** (0.0039)
Spending following Muir and Weber (2013)	−0.062*** (0.0019)	−0.059*** (0.0022)	−0.088*** (0.0040)
Spending following Crespo Cuaresma et al. (2011) including D62PAY (Social benefits other than social transfers in kind, payable)	−0.011*** (0.0016)	−0.006*** (0.0019)	−0.030*** (0.0034)
Total spending less interest payments	0.033*** (0.0014)	0.028*** (0.0017)	0.055*** (0.0031)
Fiscal data is smoothed with moving average of length 3	−0.079*** (0.0015)	−0.067*** (0.0018)	−0.115*** (0.0034)
Fiscal data is smoothed with moving average of length 5	−0.134*** (0.0017)	−0.142*** (0.0020)	−0.103*** (0.0041)
Fiscal data and GDP is smoothed with moving average of length 3	0.012*** (0.0023)	−0.014*** (0.0031)	0.026*** (0.0043)
Fiscal data and GDP is smoothed with moving average of length 5	0.030*** (0.0023)	−0.002 (0.0029)	0.123*** (0.0049)
3-variable VAR identified with sign restrictions	0.176*** (0.0055)	0.199*** (0.0079)	0.153*** (0.0088)
3-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	−0.161*** (0.0021)	−0.126*** (0.0027)	−0.257*** (0.0044)
5-variable VAR identified with Cholesky decomposition	0.160*** (0.0017)	0.158*** (0.0020)	0.268*** (0.0039)
5-variable VAR identified with sign restrictions	0.007 (0.0048)	0.050*** (0.0055)	−0.028** (0.0117)
5-variable VAR identified with BP with elasticities from Caldara and Kamps (2008)	0.040*** (0.0066)	0.061*** (0.0071)	0.051** (0.0225)
5-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	0.165*** (0.0021)	0.166*** (0.0024)	0.253*** (0.0046)

Continued on next page

Table 2 – *Continued from previous page*

Predictor	**All**	**West**	**East**
No dummies for possible outliers in the fiscal time series	−0.014*** (0.0014)	−0.005*** (0.0016)	0.018*** (0.0028)
Constant + time trend in the VAR	−0.012*** (0.0011)	−0.025*** (0.0013)	0.001 (0.0023)
VAR with 1 lag	0.024*** (0.0034)	0.016*** (0.0042)	0.069*** (0.0071)
VAR with 2 lags	0.008*** (0.0031)	−0.002 (0.0039)	0.103*** (0.0065)
VAR with 3 lags	0.019*** (0.0031)	0.004 (0.0038)	0.083*** (0.0065)
Time sample ends in 2008, before the onset of the Great Recession	0.132*** (0.0015)	0.132*** (0.0017)	0.366*** (0.0042)
Time sample ends in 2010, typically in a trough of the Great Recession	−0.098*** (0.0014)	−0.082*** (0.0017)	−0.031*** (0.0030)
Observations	420,986	218,791	132,054
R^2	0.62	0.53	0.69

Table 2: Determinants of tax cut multiplier m^{τ}_{median}: Regression results

Notes: ***, **, * denotes significance at 1, 5, 10% level, respectively. Standard errors in parentheses. Estimation by WLS with inverse variance as weight.

Predictor	All	West	East
Intercept	4.319*** (0.1287)	4.055*** (0.1181)	3.746*** (0.2427)
Belgium	−1.178*** (0.1124)	−1.002*** (0.0880)	
Bulgaria	0.869*** (0.1260)		2.203*** (0.1321)
Croatia	−0.450*** (0.1584)		
Cyprus	−3.327*** (0.1243)		
Czechia	−0.493*** (0.1077)		
Denmark	−0.596*** (0.1165)	−0.477*** (0.0913)	
Estonia	3.944*** (0.1302)		5.759*** (0.1409)
Germany	0.318*** (0.1205)	0.316*** (0.0940)	
Greece	−2.528*** (0.1170)	−2.432*** (0.0915)	
Finland	6.191*** (0.1251)	6.204*** (0.0975)	
France	−1.060*** (0.1269)	−0.955*** (0.0991)	
Hungary	−2.112*** (0.1174)		−1.539*** (0.1129)
Iceland	−0.051 (0.2353)		
Ireland	1.196*** (0.1231)	1.366*** (0.0962)	
Italy	−0.567*** (0.1107)	−0.536*** (0.0871)	
Latvia	3.254*** (0.1220)		3.512*** (0.1254)
Lithuania	15.235*** (0.1224)		15.766*** (0.1234)
Luxembourg	3.187*** (0.1159)		
Malta	−1.050*** (0.1089)		
Netherlands	0.695*** (0.1108)	0.829*** (0.0870)	
Norway	1.160*** (0.1320)		
Poland	−0.981*** (0.1165)		−0.237** (0.1141)
Portugal	−3.046*** (0.1092)	−2.809*** (0.0863)	
Romania	−1.502*** (0.1190)		−0.728*** (0.1160)
Spain	−2.526*** (0.1463)	−2.162*** (0.1138)	
Sweden	−0.697*** (0.1181)	−0.560*** (0.0918)	
Switzerland	11.425*** (0.4542)		

Continued on next page

Table 3 – *Continued from previous page*

Predictor	**All**	**West**	**East**
Slovakia	0.758*** (0.1174)		1.297*** (0.1145)
Slovenia	4.081*** (0.1319)	4.077*** (0.1030)	
United Kingdom	−3.236*** (0.1168)	−3.047*** (0.0910)	
Nominal variables deflated by HICP	−0.555*** (0.0256)	−0.228*** (0.0274)	−2.030*** (0.0611)
ESA 95 used	−0.969*** (0.0260)	0.007 (0.0276)	−2.537*** (0.0614)
Revenues following Crespo Cuaresma et al. (2011)	−0.748*** (0.0418)	−0.749*** (0.0431)	−0.113 (0.1068)
Revenues following Muir and Weber (2013)	−1.013*** (0.0416)	−0.794*** (0.0428)	−0.699*** (0.1060)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY (subsidies, payable)	−0.789*** (0.0428)	−0.836*** (0.0450)	−0.334*** (0.1076)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY and social contributions	−0.303*** (0.0411)	−0.352*** (0.0426)	0.068 (0.1051)
Spending following Crespo Cuaresma et al. (2011)	2.286*** (0.0406)	2.370*** (0.0437)	1.975*** (0.0963)
Spending following Muir and Weber (2013)	1.638*** (0.0400)	1.453*** (0.0426)	1.428*** (0.0961)
Spending following Crespo Cuaresma et al. (2011) including D62PAY (Social benefits other than social transfers in kind, payable)	0.139*** (0.0404)	−0.096** (0.0433)	0.366*** (0.0968)
Total spending less interest payments	−1.299*** (0.0395)	−1.352*** (0.0418)	−1.484*** (0.0962)
Fiscal data is smoothed with moving average of length 3	0.258*** (0.0359)	0.449*** (0.0388)	−0.193** (0.0841)
Fiscal data is smoothed with moving average of length 5	0.847*** (0.0362)	0.727*** (0.0382)	0.854*** (0.0909)
Fiscal data and GDP is smoothed with moving average of length 3	−0.104* (0.0529)	−0.094 (0.0592)	−0.654*** (0.1273)
Fiscal data and GDP is smoothed with moving average of length 5	0.193*** (0.0589)	0.084 (0.0671)	−0.180 (0.1357)
3-variable VAR identified with sign restrictions	4.752*** (0.0571)	5.172*** (0.0654)	4.168*** (0.1309)
3-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	−0.024 (0.0562)	−0.013 (0.0648)	−0.108 (0.1296)
5-variable VAR identified with Cholesky decomposition	0.281*** (0.0603)	0.457*** (0.0641)	0.055 (0.1497)
5-variable VAR identified with sign restrictions	4.676*** (0.0605)	4.855*** (0.0642)	4.661*** (0.1502)
5-variable VAR identified with BP with elasticities from Caldara and Kamps (2008)	7.612*** (0.0603)	6.989*** (0.0641)	9.841*** (0.1497)
5-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	10.585*** (0.0603)	8.821*** (0.0641)	14.930*** (0.1497)

Continued on next page

Table 3 – *Continued from previous page*

Predictor	**All**	**West**	**East**
No dummies for possible outliers in the fiscal time series	0.028 (0.0373)	-0.156^{***} (0.0396)	0.387^{***} (0.0858)
Constant + time trend in the VAR	-0.875^{***} (0.0260)	-0.523^{***} (0.0280)	-1.581^{***} (0.0625)
VAR with 1 lag	-0.697^{***} (0.0740)	-0.987^{***} (0.0797)	-0.640^{***} (0.1791)
VAR with 2 lags	-0.867^{***} (0.0677)	-1.322^{***} (0.0720)	-0.373^{**} (0.1652)
VAR with 3 lags	-0.700^{***} (0.0680)	-1.140^{***} (0.0707)	0.064 (0.1679)
Time sample ends in 2008, before the onset of the Great Recession	-1.595^{***} (0.0350)	-1.402^{***} (0.0359)	-0.992^{***} (0.0969)
Time sample ends in 2010, typically in a trough of the Great Recession	0.432^{***} (0.0320)	0.800^{***} (0.0352)	0.615^{***} (0.0776)
Observations	420,986	218,791	132,054
R^2	0.27	0.28	0.30

Table 3: Determinants of spending multiplier ranges $m^s_{16-84pr}$: Regression results

Notes: ***, **, * denotes significance at 1, 5, 10% level, respectively. Standard errors in parentheses.

Predictor	All	West	East
Intercept	5.102*** (0.0833)	5.086*** (0.0651)	0.859*** (0.1093)
Belgium	−3.150*** (0.0727)	−3.000*** (0.0485)	
Bulgaria	−2.853*** (0.0815)		0.868*** (0.0595)
Croatia	−4.056*** (0.1025)		
Cyprus	−4.938*** (0.0804)		
Czechia	−3.493*** (0.0697)		
Denmark	−3.957*** (0.0754)	−3.862*** (0.0504)	
Estonia	−2.310*** (0.0843)		1.569*** (0.0635)
Germany	−2.334*** (0.0780)	−2.237*** (0.0518)	
Greece	−2.417*** (0.0757)	−2.247*** (0.0505)	
Finland	−4.355*** (0.0810)	−4.157*** (0.0538)	
France	−4.345*** (0.0821)	−4.344*** (0.0546)	
Hungary	−0.616*** (0.0760)		2.798*** (0.0509)
Iceland	−3.700*** (0.1523)		
Ireland	−3.552*** (0.0797)	−3.353*** (0.0530)	
Italy	−4.142*** (0.0716)	−4.050*** (0.0480)	
Latvia	−2.816*** (0.0790)		0.701*** (0.0565)
Lithuania	1.268*** (0.0792)		4.803*** (0.0556)
Luxembourg	5.509*** (0.0750)		
Malta	−3.323*** (0.0705)		
Netherlands	−3.818*** (0.0717)	−3.684*** (0.0480)	
Norway	−4.055*** (0.0854)		
Poland	−4.225*** (0.0754)		−0.710*** (0.0514)
Portugal	−4.535*** (0.0707)	−4.325*** (0.0476)	
Romania	−2.288*** (0.0770)		1.207*** (0.0523)
Spain	−4.668*** (0.0947)	−4.636*** (0.0628)	
Sweden	−4.219*** (0.0764)	−4.075*** (0.0506)	
Switzerland	4.688*** (0.2939)		

Continued on next page

Table 4 – *Continued from previous page*

Predictor	**All**	**West**	**East**
Slovakia	−1.189*** (0.0760)		2.250*** (0.0516)
Slovenia	−2.696*** (0.0853)	−2.499*** (0.0568)	
United Kingdom	−3.444*** (0.0756)	−3.361*** (0.0502)	
Nominal variables deflated by HICP	−0.443*** (0.0166)	−0.037** (0.0151)	−0.420*** (0.0275)
ESA 95 used	−0.368*** (0.0168)	−0.172*** (0.0152)	−0.597*** (0.0277)
Revenues following Crespo Cuaresma et al. (2011)	−1.826*** (0.0270)	−1.580*** (0.0238)	−1.650*** (0.0481)
Revenues following Muir and Weber (2013)	−2.411*** (0.0269)	−2.004*** (0.0236)	−2.158*** (0.0478)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY (subsidies, payable)	−2.031*** (0.0277)	−1.661*** (0.0248)	−1.870*** (0.0485)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY and social contributions	−0.634*** (0.0266)	−0.492*** (0.0235)	−0.473*** (0.0474)
Spending following Crespo Cuaresma et al. (2011)	1.006*** (0.0263)	0.872*** (0.0241)	1.022*** (0.0434)
Spending following Muir and Weber (2013)	0.947*** (0.0259)	0.820*** (0.0235)	0.943*** (0.0433)
Spending following Crespo Cuaresma et al. (2011) including D62PAY (Social benefits other than social transfers in kind, payable)	0.083*** (0.0262)	0.106*** (0.0239)	0.024 (0.0436)
Total spending less interest payments	−0.222*** (0.0256)	−0.185*** (0.0231)	−0.241*** (0.0434)
Fiscal data is smoothed with moving average of length 3	0.802*** (0.0233)	0.594*** (0.0214)	0.776*** (0.0379)
Fiscal data is smoothed with moving average of length 5	1.661*** (0.0234)	1.219*** (0.0211)	1.471*** (0.0410)
Fiscal data and GDP is smoothed with moving average of length 3	0.587*** (0.0342)	0.334*** (0.0326)	0.409*** (0.0574)
Fiscal data and GDP is smoothed with moving average of length 5	0.492*** (0.0381)	0.418*** (0.0370)	0.471*** (0.0611)
3-variable VAR identified with sign restrictions	3.002*** (0.0370)	2.534*** (0.0361)	3.780*** (0.0590)
3-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	0.861*** (0.0364)	0.563*** (0.0357)	0.069 (0.0584)
5-variable VAR identified with Cholesky decomposition	−0.315*** (0.0390)	−0.119*** (0.0354)	0.137** (0.0674)
5-variable VAR identified with sign restrictions	1.763*** (0.0391)	1.844*** (0.0354)	2.164*** (0.0677)
5-variable VAR identified with BP with elasticities from Caldara and Kamps (2008)	5.422*** (0.0390)	4.313*** (0.0354)	6.677*** (0.0674)
5-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	0.620*** (0.0390)	0.605*** (0.0354)	1.343*** (0.0674)

Continued on next page

Table 4 – *Continued from previous page*

Predictor	**All**	**West**	**East**
No dummies for possible outliers in the fiscal time series	0.394*** (0.0241)	0.231*** (0.0218)	0.549*** (0.0387)
Constant + time trend in the VAR	−0.370*** (0.0168)	−0.325*** (0.0154)	−0.398*** (0.0282)
VAR with 1 lag	0.088* (0.0479)	−0.052 (0.0440)	0.172** (0.0807)
VAR with 2 lags	−0.046 (0.0438)	−0.275*** (0.0397)	0.085 (0.0744)
VAR with 3 lags	−0.220*** (0.0440)	−0.205*** (0.0390)	−0.196*** (0.0756)
Time sample ends in 2008, before the onset of the Great Recession	−0.152*** (0.0226)	−0.246*** (0.0198)	0.770*** (0.0437)
Time sample ends in 2010, typically in a trough of the Great Recession	0.033 (0.0207)	0.139*** (0.0194)	0.161*** (0.0350)
Observations	420,986	218,791	132,054
R^2	0.23	0.23	0.23

Table 4: Determinants of tax cut multiplier ranges $m^{\tau}_{16-84pr}$: Regression results

Notes: ***, **, * denotes significance at 1, 5, 10% level, respectively. Standard errors in parentheses.

Predictor	All	West	East
Intercept	0.148*** (0.0036)	0.157*** (0.0039)	−0.253*** (0.0053)
Belgium	−0.587*** (0.0034)	−0.585*** (0.0032)	
Bulgaria	−0.413*** (0.0035)		0.187*** (0.0030)
Croatia	1.090*** (0.0126)		
Cyprus	−0.228*** (0.0032)		
Czechia	−0.535*** (0.0030)		
Denmark	−0.405*** (0.0047)	−0.379*** (0.0045)	
Estonia	0.472*** (0.0077)		1.031*** (0.0074)
Germany	−0.556*** (0.0041)	−0.541*** (0.0039)	
Greece	0.609*** (0.0032)	0.578*** (0.0030)	
Finland	0.256*** (0.0077)	0.247*** (0.0073)	
France	−0.104*** (0.0033)	−0.116*** (0.0032)	
Hungary	0.372*** (0.0035)		0.846*** (0.0029)
Iceland	0.202*** (0.0102)		
Ireland	−0.199*** (0.0044)	−0.192*** (0.0042)	
Italy	0.223*** (0.0042)	0.207*** (0.0040)	
Latvia	1.420*** (0.0047)		1.936*** (0.0043)
Lithuania	1.823*** (0.0088)		2.337*** (0.0086)
Luxembourg	−1.563*** (0.0057)		
Malta	0.326*** (0.0034)		
Netherlands	−0.914*** (0.0041)	−0.928*** (0.0039)	
Norway	−2.415*** (0.0035)		
Poland	−0.013*** (0.0035)		0.460*** (0.0030)
Portugal	−0.014*** (0.0032)	−0.030*** (0.0030)	
Romania	0.799*** (0.0037)		1.261*** (0.0032)
Spain	−0.959*** (0.0032)	−0.978*** (0.0030)	
Sweden	0.591*** (0.0038)	0.587*** (0.0036)	
Switzerland	−1.434*** (0.0169)		

Continued on next page

Table 5 – *Continued from previous page*

Predictor	All	West	East
Slovakia	-0.183^{***} (0.0034)		0.372^{***} (0.0027)
Slovenia	0.255^{***} (0.0059)	0.242^{***} (0.0056)	
United Kingdom	-0.292^{***} (0.0028)	-0.318^{***} (0.0027)	
Nominal variables deflated by HICP	0.072^{***} (0.0009)	0.038^{***} (0.0011)	0.044^{***} (0.0017)
ESA 95 used	0.065^{***} (0.0009)	0.057^{***} (0.0011)	0.036^{***} (0.0017)
Revenues following Crespo Cuaresma et al. (2011)	0.090^{***} (0.0014)	0.090^{***} (0.0017)	-0.006^{**} (0.0029)
Revenues following Muir and Weber (2013)	0.073^{***} (0.0014)	0.069^{***} (0.0017)	-0.011^{***} (0.0029)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY (subsidies, payable)	0.066^{***} (0.0014)	0.089^{***} (0.0017)	-0.062^{***} (0.0028)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY and social contributions	0.045^{***} (0.0014)	0.033^{***} (0.0017)	0.021^{***} (0.0028)
Spending following Crespo Cuaresma et al. (2011)	0.052^{***} (0.0016)	0.117^{***} (0.0020)	0.101^{***} (0.0031)
Spending following Muir and Weber (2013)	0.114^{***} (0.0015)	0.158^{***} (0.0018)	0.112^{***} (0.0030)
Spending following Crespo Cuaresma et al. (2011) including D62PAY (Social benefits other than social transfers in kind, payable)	-0.069^{***} (0.0014)	-0.121^{***} (0.0017)	0.109^{***} (0.0029)
Total spending less interest payments	0.059^{***} (0.0012)	0.063^{***} (0.0015)	0.132^{***} (0.0024)
Fiscal data is smoothed with moving average of length 3	-0.022^{***} (0.0014)	-0.012^{***} (0.0017)	-0.046^{***} (0.0028)
Fiscal data is smoothed with moving average of length 5	-0.086^{***} (0.0015)	-0.083^{***} (0.0019)	-0.087^{***} (0.0031)
Fiscal data and GDP is smoothed with moving average of length 3	-0.032^{***} (0.0013)	-0.023^{***} (0.0016)	-0.049^{***} (0.0025)
Fiscal data and GDP is smoothed with moving average of length 5	-0.053^{***} (0.0013)	-0.068^{***} (0.0016)	0.002 (0.0025)
3-variable VAR identified with sign restrictions	-0.075^{***} (0.0023)	-0.016^{***} (0.0033)	-0.240^{***} (0.0037)
3-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	0.020^{***} (0.0013)	0.049^{***} (0.0016)	-0.077^{***} (0.0024)
5-variable VAR identified with Cholesky decomposition	0.030^{***} (0.0015)	-0.010^{***} (0.0017)	0.092^{***} (0.0035)
5-variable VAR identified with sign restrictions	0.245^{***} (0.0040)	-0.108^{***} (0.0051)	0.649^{***} (0.0070)
5-variable VAR identified with BP with elasticities from Caldara and Kamps (2008)	-0.070^{***} (0.0046)	-0.121^{***} (0.0047)	0.418^{***} (0.0160)
5-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	-0.385^{***} (0.0049)	-0.484^{***} (0.0050)	0.513^{***} (0.0190)

Continued on next page

Table 5 – *Continued from previous page*

Predictor	**All**	**West**	**East**
No dummies for possible outliers in the fiscal time series	0.021*** (0.0014)	0.003* (0.0016)	−0.012*** (0.0029)
Constant + time trend in the VAR	−0.117*** (0.0009)	−0.142*** (0.0011)	−0.011*** (0.0017)
VAR with 1 lag	−0.095*** (0.0013)	−0.072*** (0.0015)	−0.123*** (0.0025)
VAR with 2 lags	0.025*** (0.0013)	0.068*** (0.0016)	−0.038*** (0.0025)
VAR with 3 lags	−0.003** (0.0013)	0.026*** (0.0016)	−0.050*** (0.0024)
Time sample ends in 2008, before the onset of the Great Recession	−0.075*** (0.0011)	0.001 (0.0013)	−0.154*** (0.0022)
Time sample ends in 2010, typically in a trough of the Great Recession	−0.258*** (0.0012)	−0.289*** (0.0015)	−0.272*** (0.0027)
Observations	3,173,390	1,920,535	845,035
R^2	0.44	0.31	0.39

Table 5: Determinants of spending multiplier m^s_{median}: Regression results

Notes: ***, **, * denotes significance at 1, 5, 10% level, respectively. Standard errors in parentheses. Estimation by WLS with inverse variance as weight.

Predictor	All	West	East
Intercept	−0.140*** (0.0019)	−0.108*** (0.0020)	−0.794*** (0.0031)
Belgium	−0.226*** (0.0018)	−0.238*** (0.0018)	
Bulgaria	−0.427*** (0.0020)		−0.037*** (0.0017)
Croatia	−1.157*** (0.0035)		
Cyprus	−0.384*** (0.0018)		
Czechia	−0.370*** (0.0020)		
Denmark	−0.504*** (0.0018)	−0.514*** (0.0018)	
Estonia	−2.238*** (0.0036)		−1.757*** (0.0034)
Germany	−0.279*** (0.0020)	−0.285*** (0.0020)	
Greece	−0.225*** (0.0021)	−0.236*** (0.0022)	
Finland	−0.813*** (0.0019)	−0.825*** (0.0019)	
France	−0.218*** (0.0017)	−0.232*** (0.0017)	
Hungary	0.035*** (0.0030)		0.443*** (0.0028)
Iceland	0.197*** (0.0057)		
Ireland	−1.094*** (0.0022)	−1.109*** (0.0022)	
Italy	−0.161*** (0.0019)	−0.174*** (0.0019)	
Latvia	−1.716*** (0.0024)		−1.254*** (0.0021)
Lithuania	−1.603*** (0.0033)		−1.191*** (0.0031)
Luxembourg	−0.672*** (0.0057)		
Malta	−0.220*** (0.0023)		
Netherlands	−0.484*** (0.0019)	−0.493*** (0.0019)	
Norway	−1.317*** (0.0023)		
Poland	−0.342*** (0.0018)		0.038*** (0.0014)
Portugal	−0.227*** (0.0019)	−0.239*** (0.0019)	
Romania	−1.121*** (0.0034)		−0.676*** (0.0031)
Spain	−0.320*** (0.0017)	−0.337*** (0.0017)	
Sweden	−0.140*** (0.0018)	−0.162*** (0.0018)	
Switzerland	−0.136*** (0.0129)		

Continued on next page

Table 6 – *Continued from previous page*

Predictor	**All**	**West**	**East**
Slovakia	0.077*** (0.0025)		0.434*** (0.0022)
Slovenia	−1.360*** (0.0030)	−1.377*** (0.0030)	
United Kingdom	−0.315*** (0.0017)	−0.331*** (0.0017)	
Nominal variables deflated by HICP	−0.041*** (0.0004)	−0.060*** (0.0004)	0.034*** (0.0009)
ESA 95 used	0.027*** (0.0004)	0.028*** (0.0004)	−0.009*** (0.0010)
Revenues following Crespo Cuaresma et al. (2011)	0.060*** (0.0007)	0.066*** (0.0007)	0.183*** (0.0023)
Revenues following Muir and Weber (2013)	0.144*** (0.0006)	0.135*** (0.0007)	0.310*** (0.0022)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY (subsidies, payable)	0.082*** (0.0006)	0.092*** (0.0007)	0.188*** (0.0022)
Revenues following Crespo Cuaresma et al. (2011) without subtracting D3PAY and social contributions	−0.040*** (0.0007)	−0.003*** (0.0008)	−0.182*** (0.0025)
Spending following Crespo Cuaresma et al. (2011)	−0.067*** (0.0006)	−0.072*** (0.0007)	−0.052*** (0.0016)
Spending following Muir and Weber (2013)	−0.066*** (0.0006)	−0.070*** (0.0007)	−0.055*** (0.0016)
Spending following Crespo Cuaresma et al. (2011) including D62PAY (Social benefits other than social transfers in kind, payable)	−0.011*** (0.0005)	−0.011*** (0.0006)	−0.005*** (0.0014)
Total spending less interest payments	0.024*** (0.0005)	0.017*** (0.0005)	0.060*** (0.0013)
Fiscal data is smoothed with moving average of length 3	−0.075*** (0.0006)	−0.071*** (0.0006)	−0.085*** (0.0015)
Fiscal data is smoothed with moving average of length 5	−0.128*** (0.0006)	−0.136*** (0.0007)	−0.082*** (0.0017)
Fiscal data and GDP is smoothed with moving average of length 3	0.011*** (0.0005)	0.006*** (0.0006)	0.005*** (0.0014)
Fiscal data and GDP is smoothed with moving average of length 5	0.003*** (0.0005)	0.000 (0.0006)	0.013*** (0.0014)
3-variable VAR identified with sign restrictions	0.095*** (0.0022)	0.046*** (0.0029)	0.169*** (0.0040)
3-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	−0.205*** (0.0006)	−0.208*** (0.0008)	−0.212*** (0.0016)
5-variable VAR identified with Cholesky decomposition	0.164*** (0.0005)	0.148*** (0.0006)	0.258*** (0.0014)
5-variable VAR identified with sign restrictions	0.013*** (0.0019)	0.017*** (0.0022)	0.026*** (0.0045)
5-variable VAR identified with BP with elasticities from Caldara and Kamps (2008)	0.035*** (0.0016)	0.027*** (0.0017)	−0.000 (0.0079)
5-variable VAR identified with BP with elasticities from Crespo Cuaresma et al. (2011)	0.165*** (0.0007)	0.149*** (0.0007)	0.243*** (0.0017)

Continued on next page

Table 6 – *Continued from previous page*

Predictor	**All**	**West**	**East**
No dummies for possible outliers in the fiscal time series	-0.006^{***} (0.0004)	-0.007^{***} (0.0005)	0.027^{***} (0.0011)
Constant + time trend in the VAR	-0.011^{***} (0.0003)	-0.017^{***} (0.0004)	0.002^{*} (0.0009)
VAR with 1 lag	0.022^{***} (0.0005)	0.029^{***} (0.0006)	0.007^{***} (0.0014)
VAR with 2 lags	0.044^{***} (0.0006)	0.040^{***} (0.0006)	0.091^{***} (0.0014)
VAR with 3 lags	0.026^{***} (0.0006)	0.022^{***} (0.0007)	0.060^{***} (0.0014)
Time sample ends in 2008, before the onset of the Great Recession	0.149^{***} (0.0005)	0.150^{***} (0.0005)	0.255^{***} (0.0016)
Time sample ends in 2010, typically in a trough of the Great Recession	-0.082^{***} (0.0004)	-0.076^{***} (0.0005)	-0.035^{***} (0.0012)
Observations	3,173,390	1,920,535	845,035
R^2	0.61	0.54	0.70

Table 6: Determinants of tax cut multiplier m^{τ}_{median}: Regression results

Notes: ***, **, * denotes significance at 1, 5, 10% level, respectively. Standard errors in parentheses. Estimation by WLS with inverse variance as weight.

References

Caldara D and Kamps C (2008) What are the effects of fiscal shocks? A VAR-based comparative analysis. Technical report, ECB Working Paper No. 877

Crespo Cuaresma J, Eller M and Mehrotra A (2011) The Economic transmission of fiscal policy shocks from Western to Eastern Europe. *BOFIT Discussion Papers* 2011(12), 3

Muir D and Weber A (2013) Fiscal Multipliers in Bulgaria: Low But Still Relevant. Technical report, IMF Working Paper No. 13/49

Oxford Economic Papers, 2021, 1–19
doi: 10.1093/oep/gpab027

OXFORD

Fiscal multipliers in a small open economy: the case of Austria

Jan Čapek[a], Jesús Crespo Cuaresma[a,b,c,d,e], Johannes Holler[f], and Philip Schuster[f]

[a]Department of Economics, Masaryk University (MU), Lipová 507/41a, 60200 Brno, Czech Republic
[b]Department of Economics, Vienna University of Economics and Business (WU), Welthandelsplatz 1, 1020 Vienna, Austria;
e-mail: jcrespo@wu.ac.at
[c]International Institute of Applied System Analysis (IIASA), Population and Just Societies Program, Schlossplatz 1, 2361 Laxenburg, Austria
[d]Wittgenstein Center, IIASA, OeAW, University of Vienna, Schlossplatz 1, 2361 Laxenburg, Austria
[e]Austrian Institute of Economic Research, WIFO, Arsenal Objekt 20, 1030 Vienna, Austria
[f]Office of the Austrian Fiscal Advisory Council, Otto-Wagner-Platz 3, 1090 Vienna, Austria

Abstract

We estimate fiscal multipliers for Austria in a framework of model uncertainty emanating from the choice of a particular econometric model. We present a comprehensive framework that allows to assess the effects of different multiplier definitions and choices related to the data, the model employed, and further technical choices associated with the specification of the model exert on fiscal multiplier estimates. The mean present-value government spending multiplier over all models entertained, based on around 3,000 estimates, is 0.68. Estimates of the peak spending multiplier for Austria tend to be larger than present-value spending multipliers, with a mean value of 0.85. The magnitude of the present-value tax multiplier is relatively high, with an average value across specifications of −1.12 and the mean peak tax multiplier is −0.54 for all specifications used.

JEL classifications: E62, C32

1. Introduction

The interest in assessing the macroeconomic effects of fiscal policy in industrialized countries has gained renewed momentum since the Great Recession. Given the limited scope of action of monetary policy in the context of very low nominal interest rates, fiscal policy re-emerged as a policy of choice and a large literature has concentrated on investigating how fiscal policy affects macroeconomic variables and gross domestic product (GDP) in particular.[1] A convenient way to communicate the effects of fiscal stimulus on the economy is the fiscal multiplier, measured as the dollar reaction of GDP as a result of a 1-dollar fiscal stimulus. Fiscal multipliers are easily comparable across time and countries and the precision of their estimation contributes significantly to the quality of GDP growth predictions (Blanchard and Leigh, 2013). The estimates of fiscal multipliers are infamously heterogeneous both across countries and methods used for their calculation, and may be very sensitive to arguably minor specification choices, as recently shown in Čapek and Crespo Cuaresma (2020).

There is little evidence on the size of fiscal multipliers for developed European small open economies.[2] Ravn and Spange (2012) enhance the Blanchard–Perotti methodology, based on structural vector autoregression (SVAR) models to estimate spending multipliers for Denmark and obtain a point estimate of approximately 0.6 after four quarters. Jemec *et al.* (2011) investigate Slovenian fiscal policy employing a standard SVAR approach and estimate an impact spending multiplier of 1.5, which diminishes in subsequent periods. Unfortunately, not all studies investigating the effects of fiscal stimuli report the results in the form of multipliers (e.g. Afonso and Sousa, 2011 for Portugal or Benetrix and Lane, 2009 for Ireland). In addition to estimates for single countries, evidence from panel studies also exists. Ilzetzki *et al.* (2013) report that the subgroups of countries corresponding to high-income, open, low-debt, and fixed exchange rate countries have average spending multipliers of 0.4, 0, 0.2, and 0.6, respectively. The empirical evidence can be supplemented making use of the work by Barrell *et al.* (2012), where a model-based consumption multiplier of 0.5 is reported for Austria. Breuss *et al.* (2009) provide an overview of fiscal multipliers derived by Austrian forecasting institutions from large-scale macroeconometric models (within the tradition of the *Cowles commission approach*). Spending multipliers over the first year after the fiscal shock are typically below unity, first-year wage and income tax multipliers are below 0.5. Recent papers by Koch *et al.* (2019) and Schuster (2019) complement the existing results by simulating fiscal multipliers for Austria using calibrated New-Keynesian general equilibrium models and derive multipliers of comparable magnitudes. However, to our knowledge, a pure empirical assessment of fiscal multipliers specifically for Austria, as a stereotypical small open economy within the group of industrialized countries, does not exist. In this contribution, we provide for the first time a rigorous analysis of fiscal multiplier estimates in a small open economy (Austria) incorporating the uncertainty related to specification choice in several dimensions including that related to the particular variables included in the model, shock identification strategies, data preparation, or the analytical structure of the model. Given the importance of economic openness to determine the size of the fiscal multiplier, such an exercise allows the results to be interpreted in the framework of theoretical models of fiscal policy effects in small open economy settings. Theoretical results of this literature predict lower domestic effects of fiscal policy through the leaking of fiscal shocks to imported goods, combined with a higher sensitivity to international economic policy spill-overs (see, e.g. Karras, 2014).

The main bulk of the existing literature on the macroeconomic effects of fiscal interventions can be categorized as either model-based or empirical. Model-based approaches

1 See, for example, Hebous (2011) or Ramey (2011a) for earlier surveys on the issue, or Ramey (2019) for a recent one.

2 See the extensive summary of existing multiplier estimates in Mineshima *et al.* (2014) or the data used for the broad meta-analysis in Gechert (2015)

typically employ calibratedDSGE models to study the effects of fiscal stimuli in an internally consistent theoretical framework. Kilponen *et al.* (2019), for instance, compare such estimates of fiscal multipliers across models and countries in Europe, while Barrell *et al.* (2012) focus on model-based fiscal multipliers in the context of fiscal consolidation. The advantage of the model-based approach lies in the ability to analyse counterfactual scenarios by simulating the dynamics of the model variables under different conditions. On the other hand, empirical approaches, mostly based on SVAR models, tend to be more data-driven and typically impose less stringent restrictions on the structure of the economic model. The availability of long time series for some countries allows for the use of modern identification methods, such as the narrative approach (Ramey, 2011b) to extract exogenous fiscal shocks or the assessment of different regimes (Auerbach and Gorodnichenko, 2012) where fiscal multipliers may differ. In cases where such long time series are not available, countries are often pooled and the empirical analysis is conducted on a panel setting (Beetsma and Giuliodori, 2011; Ilzetzki *et al.*, 2013), or fiscal multipliers for single economies with shorter time series are studied using SVAR models inspired by the seminal contribution by Blanchard and Perotti (2002).[3]

The estimates of fiscal multipliers tend to differ, sometimes strongly, from study to study (see the evidence presented in the meta-analyses provided by Gechert (2015)). These differences can be attributed to various identification strategies (Caldara and Kamps, 2017) as well as to other technical choices made in the analysis (Čapek and Crespo Cuaresma, 2020). Given the additional dimension of uncertainty on fiscal multiplier estimates implied by the particular methodological choices, even within the class of SVAR models, the approach of this study is to present a consistent framework which encompasses a wide range of reasonable settings and choices which are routinely used in the empirical literature on fiscal multipliers. The framework entertains thousands of multiplier estimates, each for a particular model specification.

Expanding the methodological setting in Čapek and Crespo Cuaresma (2020), in this contribution, we exploit the differences in out-of-sample predictive power of the models constructed for GDP in order to gain insights into the size of fiscal multipliers in Austria. Our analysis expands the method put forward in Čapek and Crespo Cuaresma (2020) in several respects. First of all, by concentrating on a single economy, we gain comparability in the multiplier estimates, which correspond to the responses to fiscal impulses within the same institutional and historical setting. Such a focus allows for a direct interpretation of the differences in multiplier estimates as being driven by specification choice, and not caused by variation in the structural characteristics of the group of economies being analysed, including the institutional setting in which fiscal policy is enacted. As compared with Čapek and Crespo Cuaresma (2020), our analysis is thus able to better measure the pure effect of model uncertainty (as opposed to the that caused by differing institutions and underlying factors across countries) in the inference of fiscal multipliers for the country under scrutiny. Furthermore, in the econometric setting proposed, we expand the set of econometric specifications and modelling choices in Čapek and Crespo Cuaresma (2020) by including more modern models based on factor-augmented vector autoregression (VAR) structures. The focus on a single small open economy allows us to link the results in a more direct manner to the methodological framework provided by economic theory, in

3 See, for example, Ramey (2016) for a review of the methods used for the identification of exogenous fiscal shocks.

particular, when interpreting the results of the analysis, and allows for the assessment of additional sources of model uncertainty as compared with Čapek and Crespo Cuaresma (2020). This is the case, for example, for the composition of government spending and tax aggregates, or for the calibration values of tax and spending elasticities required for several identification techniques. Therefore, the setting used in this contribution provides us with a powerful instrument to assess the variability of fiscal multiplier estimates across specifications, as well as across definitions of spending and tax aggregates.

Our results expose the uncertainty and heterogeneity that is inherent to empirical estimates of fiscal multipliers even for a given country. In addition to entertaining different SVAR specifications based on Blanchard and Perotti (2002) and Perotti (2004), we also estimate fiscal multipliers from structural Factor-Augmented VAR (FAVAR) models. These specifications provide a more adequate framework to account for fiscal foresight and omitted variable biases (Fragetta and Gasteiger, 2014). Furthermore, we also exploit the existing data on government spending and tax composition in Austria in order to obtain additional multiplier estimates. We compare the results for the two most widely used formulations in the literature—the present-value multiplier and the peak multiplier and deliver the first set of credible multiplier estimates for a representative European small open economy after accounting for model uncertainty.

The mean spending multiplier for Austria is estimated at 0.68 for the present-value multiplier and 0.85 for the peak multiplier. The present-value tax multiplier is –1.12 and its peak counterpart is –0.54. Comparing the multipliers to the existing literature, our estimates suggest a stronger reaction of GDP after the increase of government spending as compared with the results for relevant subgroups of countries reported in Ilzetzki *et al.* (2013). Our estimate of present-value multiplier specification is comparable to that of Denmark (see Ravn and Spange, 2012). As in the case of the study on the Slovenian economy, our results also suggest that peak spending multipliers tend to be higher than their present-value counterparts (see Jemec *et al.*, 2011). The multiplier estimates obtained using the subset of models with relatively superior predictive ability for GDP tend to be larger in magnitude. Our results also indicate that the models based on subcomponents of government spending and taxes that deliver the best predictive ability for GDP dynamics tend to include compensation of employees, intermediate consumption, and gross capital formation as part of government expenditures and taxes on production, imports, income, and wealth. On average, SVAR models of a smaller dimension and using the Cholesky decomposition as an identification device tends to result in relatively lower spending multipliers. On the other hand, using more variables for estimation and employing identification schemes that follow the Blanchard–Perotti or sign restriction approach deliver results with relatively higher values of spending multipliers. Similar patterns hold for peak tax multipliers, although the differences are smaller. We also find evidence corroborating a conclusion in Ramey (2019) that the specific definition of the multiplier used may lead to significantly different estimates. After carrying out several sensitivity checks, we find that peak multipliers for Austria tend to appear more stable than their present-value counterparts.

The article is organized as follows. Section 2 briefly presents the methodological setting used to estimate fiscal multipliers, based on SVAR and structural FAVAR models. Section 3 describes the different specification designs assessed for the estimation of fiscal multipliers in Austria. Section 4 presents the results of the analysis in detail and Section 5 concludes.

2. Estimating fiscal multipliers: SVAR and structural FAVAR models

We can nest the set of models used to estimate fiscal multipliers in the stacked form of a dynamic factor model, following Stock and Watson (2016). A set of q dynamic factors are stacked to yield r static factors in the vector F_t and, abstracting from further deterministic terms, a FAVAR structure would be given by the following equations:

$$\begin{pmatrix} \underset{n\times 1}{Y_t} \\ \underset{m\times 1}{X_t} \end{pmatrix} = \begin{pmatrix} \underset{n\times n}{\mathbf{I}} & \underset{n\times r}{0} \\ \underset{m\times n}{\mathbf{\Lambda}^Y} & \underset{m\times r}{\mathbf{\Lambda}^F} \end{pmatrix} \begin{pmatrix} \underset{n\times 1}{\tilde{F}_t} \\ \underset{r\times 1}{F_t} \end{pmatrix} + \begin{pmatrix} \underset{n\times 1}{0} \\ \underset{m\times 1}{e_t} \end{pmatrix} \tag{1}$$

$$\underset{(n+r)\times(n+r)}{\mathbf{\Phi}(L)} \begin{pmatrix} \underset{n\times 1}{\tilde{F}_t} \\ \underset{r\times 1}{F_t} \end{pmatrix} = \begin{pmatrix} \underset{(n+q)\times(n+q)}{\mathbf{I}} \\ \underset{(r-q)\times(n+q)}{0} \end{pmatrix} \underset{(n+q)\times 1}{\eta_t} \tag{2}$$

$$\underset{(n+q)\times(n+q)}{\mathbf{A}} \underset{(n+q)\times 1}{\eta_t} = \underset{(n+q)\times(n+q)}{\mathbf{B}} \underset{(n+q)\times 1}{\varepsilon_t} \tag{3}$$

where Equation (1) is the measurement equation, Equation (2) is the transition equation, and Equation (3) is the identification equation, while the (matrix) lag polynomial $\mathbf{\Phi}(L)$ is given by $\mathbf{\Phi}(L) = \mathbf{I} - \mathbf{\Phi}_1 L - \ldots - \mathbf{\Phi}_p L^p$ for matrices $\mathbf{\Phi}_l$, $l = 1, \ldots, p$. The variables in Y_t (output, fiscal variables, and other covariates) are assumed to be measured without error by the observed factors $\tilde{F}_t$. X_t contains m observed time series (not contained in Y_t) summarizing information about other macroeconomic and financial phenomena, as well as variables related to labour markets, production, and sectoral developments. Variables in X_t are assumed to depend on observed factors $\tilde{F}_t$, unobserved factors F_t and an idiosyncratic component e_t, with matrix $\mathbf{\Lambda}^{\mathrm{F}}$ comprising the corresponding factor loadings. Equation (3) specifies the relationship between reduced-form (η_t) and structural shocks (ε_t). If the number of unobserved factors r is set to zero, the model collapses to a standard SVAR model which can be utilized to implement the methods in Blanchard and Perotti (2002) or Perotti (2004). The unobserved factors of the model (F_t) are estimated as principal components and the identification of the model is reached once matrices $\mathbf{A}$ and $\mathbf{B}$ are chosen (see Stock and Watson, 2016).

Various identification methods can be used to retrieve the structural shocks in ε_t. The method pioneered by Blanchard and Perotti (2002) relies on exact restrictions imposed on the error terms of a VAR model which includes GDP, government expenditure and taxes through an identification scheme based on lags in the implementation of fiscal policy. More modern methods (Rubio-Ramírez *et al.*, 2010) use sign restrictions that constrain the direction of the response of variables to particular shocks. Once the structural shocks have been identified, government spending and tax multipliers can be computed. In line with recent literature (e.g. Mountford and Uhlig, 2009; Ilzetzki et al., 2013; Caggiano et al., 2015; Gechert and Rannenberg, 2018), we report present-value (or discounted cumulative) multipliers at lag T as follows:

$$\text{present} - \text{valuespendingmultiplier} = \frac{\sum_{t=0}^{T} (1+i)^{-t} y_t}{\sum_{t=0}^{T} (1+i)^{-t} g_t} \times \frac{1}{g/y}, \tag{4}$$

where y_t is the response of output at time t (in logs), g_t denotes the response of government

Table 1. Modelling choices for the estimation of fiscal multipliers

Dimension	Variants considered
Government data composition	Nine variants, see Table 2; ESA2010 codes and time series in Supplementary Appendix A
Deflating index	GDP deflator and HICP (not lagged and lagged by four quarters)
Model	VAR and FAVAR models with 3–5 vars. (factors ordered first or last)
Identification strategy	Cholesky ordering, Blanchard–Perotti, sign restrictions
Number of factors	1–2 (FAVARs only)
Deterministics and lags	Constant or linear trend, 1–4 lags

Source: Authors' calculations.

expenditures at time t (in logs), and g/y is the average share of government expenditures in GDP over the sample. The multiplier is discounted with the interest rate i, which is set to 4% per annum.[4] In the context of data at quarterly frequency, we report discounted cumulative multipliers for $T=4$. The tax multiplier is calculated analogously, after substituting government expenditures in Equation (4) with taxes.

If we concentrate on the non-cumulative reaction of GDP, such effects can be summarized using the so-called peak multipliers (see, e.g. Blanchard and Perotti, 2002; Ramey, 2011b; Fragetta and Gasteiger, 2014; Caggiano et al., 2015):

$$\text{peakspendingmultiplier} = \frac{\max_{t=0,\dots,H}\{y_t\}}{\max_{t=0,\dots,H}\{g_t\}} \times \frac{1}{g/y}, \tag{5}$$

In Equation (5), in order to respect the business cycle nature of the multipliers (and the known unreliability of results for longer horizons in these specifications), we restrict the horizon to a maximum of 2 years and set $H=8$.

3. Model specifications and data

3.1 Specification choices

As reported in Čapek and Crespo Cuaresma (2020), in the context of estimating multipliers using SVAR specifications, seemingly harmless modelling choices may have a significant effect on the size and precision of fiscal multiplier estimates. In addition to the structural shock identification strategy, these modelling choices include the definition of spending and taxes, the national accounts system employed, the use of particular interest rates or inflation measures in the model, or whether data are smoothed prior to estimation. On the sample of European countries, the cumulative effects of such arguably innocuous methodological choices can lead to large changes in the spending multipliers. We explicitly integrate such uncertainty into our estimated for Austria, entertaining the large number of models which can be obtained by combining such possible methodological choices.

4 The discounting does not play major role in case of moderate interest rates, while it becomes more important in environments of high interest rate, such as emerging economies. The selection of a 4% interest rate corresponds to a commonly used discount factor of 0.99 per period.

Table 1 lists all the methodological choices considered to construct models aimed at estimating fiscal multipliers for Austria. The set of possible variants are obtained by combining choices relating to (i) the data employed, (ii) the model used, and (iii) the particular details related to the specification of the model. As for the data choices, these mainly concern the composition of government spending and revenues, but can also differ in the choice of the price index used to deflate nominal variables (CPI versus GDP deflator). As a large part of government spending in Austria is linked to the lagged CPI (e.g. pension payments), we additionally consider lagged CPI (four-quarters lag) in our analysis. The basic modelling choices in terms of specification structure are related to the (i) use of a simple VAR model versus employing a specification that incorporates unobserved factors, that is, a FAVAR model, (ii) the selection of variables in the (FA)VAR model, and (iii) the choice of the identification strategy. Given a model specification, the technical choices relate to the number of deterministic terms in the (FA)VAR equation and the number of lags. For each model specification, we bootstrap 4,000 multipliers and use the median as our point estimate.[5] The main analysis includes 2,987 different specifications that can be obtained by combining the choices at hand, each yielding a (peak and present-value) spending and a tax median multiplier.

Table 2 presents the compositions of government spending and revenues used to obtain fiscal multipliers. Each choice consists of a specific composition of the government spending and government taxes aggregate. The *Baseline* setting ('Core/Tax Tiny') employs a very simple composition which contains just three components of spending (compensation of employees, intermediate consumption, and gross capital formation) and two components of revenues (taxes on production, imports, income, and wealth).[6] The following three combinations adjust the baseline setting by including also social contributions, benefits, and subsidies as part of the fiscal aggregate (as in Crespo Cuaresma *et al.*, 2011, for instance). To reflect the particularities of the Austrian economy, other compositional choices reflecting the importance of transfers in kind, household social contributions, subsidies, and transfers for the country have to be made. Deviating from the existing literature, so as to cover the specific case of Austria, we introduce three new data compositions (tag starting with 'corefix' in Table 2). The inclusion of social transfers in kind in this composition of government spending accounts for the fact that social transfers in kind amount to >8% of overall government spending in the country. Because of their use to finance large parts of the healthcare and social protection system, changes in the provision of social transfers in kind create important economic spill-overs (e.g. by substituting private expenditure for old-age and long-term care) that should be considered in the analysis. The particular revenue compositions used reflect the importance of household social contributions, subsidies, and transfers for overall disposable household income in Austria. Following Muir and Weber (2013), we also entertain models based on government spending aggregates that contain acquisitions of assets and a battery of adjustments regarding social contributions, subsidies, and transfers (including capital transfers). The spending and tax aggregate compositions

5 In sign restriction identification schemes, the 4,000 solutions are the actual draws. Other identification approaches rely on bootstrapping to compute the 4,000 draws. The bootstrap employed builds on resampling raw residuals (with replacement) and subsequent refitting of the model. Portmanteau tests for residual autocorrelation suggest that around two thirds of the estimated models do not exhibit significant residual autocorrelation at any sensible lag.

6 See Appendix A for the ESA2010 codes corresponding to each component.

Table 2. Government spending and revenues composition

Tag	Government spending composition	Government revenues composition
core/tax tiny (Baseline)	Compensation of employees, intermediate consumption, and gross capital formation	Taxes on production, imports, income, and wealth
core/tax small net soc.t.		Baseline adjusted for actual social contributions
core/net tax small		Baseline adjusted for social contributions and subsidies
corefix+soc.t.kind/tax mid	Baseline (gross fixed capital) + transfers in kind	Baseline + household social contributions
corefix+soc.t.kind/net tax mid		Baseline + household social contributions adjusted for subsidies
corefix+soc.t.kind/net tax large		Baseline + household social contributions adjusted for subsidies and transfers
core/net tax all	Baseline + acquisitions of assets	Baseline + household social contributions adjusted for subsidies and transfers (incl. capital transfers)

Source: Authors' classification.
Notes: There are nine sets of compositions of government spending and revenues. Starting from 'core/tax tiny', which is the *Baseline* composition (shaded in grey), the other composition sets add extra spending and/or revenue items. These are ordered from narrower to broader sets, comprising many spending and/or revenue items. The corresponding tag is constructed with abbreviations of spending composition separated from abbreviations of revenue composition using a slash '/'. The term 'core' refers to the *Baseline* spending composition, 'corefix' highlights the use of fixed capital formation. The abbreviations for taxes range from 'tiny', with only several items, to 'all', with a broad selection of revenue items. The last row of the table is the only representative of top-down composition approach starting from total spending and total revenue. For specific ESA codes for each composition set, see Supplementary Appendix A.

mentioned above follow a bottom-up approach and are created by adding together the particular variables measuring the parts of spending and revenues that are relevant for the estimation of the fiscal shock. The last compositional choice considered ('Top Down Spend./Top Down Rev.') takes a top-down approach by starting from the full aggregates of total spending and total revenues and subtracting the parts that are not relevant for the estimation of the fiscal shock.

The Cholesky identification strategy identifies a fiscal shock using a particular ordering based on the contemporaneous responses across shocks. The first and most exogenous variable is assumed to be government spending, followed by GDP, inflation (in VAR models with four and five variables), taxes, and the interest rate (in VAR models with five variables only). The Blanchard–Perotti identification scheme follows Blanchard and Perotti (2002) for VAR models with three variables and Perotti (2004) for specifications with more variables. The (aggregate) output and price elasticities of government revenue required to perform the shock identification exercise are calculated as weighted averages of elasticities of different net-tax components. The calculation follows Burriel *et al.* (2010), and uses elasticities of specific net tax and transfer components from Mourre *et al.* (2014) and Price *et al.* (2014). The output and price elasticities of government revenue computed for Austria are 1.66 and 0.78, respectively.[7] The price elasticity of spending is assumed to be –0.5 (in line with the literature, e.g. Crespo Cuaresma *et al.*, 2011; Perotti, 2004). Our implementation of sign restrictions identifies three shocks: the business cycle shock is identified by requiring the impulse responses of output and taxes to be positive for at least the four quarters following the shock. The tax shock is identified by a positive response of taxes for at least the four quarters following the shock (and the shock is required not to meet the identifying restrictions for the business cycle shock). For the identification of a government spending shock, the responses of government spending need to be positive for at least the four quarters following the shock (and the shock is required not to meet the identifying restrictions for the business cycle shock).

The identification strategies mentioned above are unable to explicitly address the issue of fiscal foresight. If a fiscal policy change is known before its (official) implementation and economic agents react accordingly, the reaction in the real economy may be apparent earlier. This timing mismatch is known as fiscal foresight and essentially amounts to a limited information problem (Fragetta and Gasteiger, 2014). Forni and Gambetti (2014) suggest to remedy the problem by extending the VAR model with principal components (as estimates of unobservable factors), which are calculated from a broad range of additional time series containing relevant information. We add one or two principal components to the VAR specification with three variables, making the model a proper FAVAR specification. We estimate the principal components with the aid of 26 additional time series that relate to macroeconomic dynamics, financial markets, and the labour market.[8]

3.2 Data

The main source of data is Eurostat, while some financial variables used for the estimation of the unobserved factors are sourced from the European Central Bank. We use 30 different time series to construct the various disaggregated variables for government spending and revenue required to estimate our models. For extended versions of the VAR model with four and five variables, we also use inflation and the interest rate. The data cover the period spanned from the first quarter of 2001 to the fourth quarter of 2018, yielding 72 quarterly observations. If available, seasonally adjusted variables are employed. If seasonally adjusted data are unavailable, we use the X-13 toolbox to remove seasonal patterns from those variables that contain a seasonal component.[9] All the time series for spending and tax categories, as well as GDP, are downloaded from the source in nominal terms and subsequently

7 See Section 4 for a sensitivity analysis exercise in which we vary both of these elasticities and Appendix C for detailed results. See Appendix D for details of the calculation of these aggregate elasticities.

8 See Appendix A for the list of the time series used to estimate the factors.

9 We employ the X-13 Toolbox for Seasonal Filtering by Yvan Lengwiler on Matlab file exchange. The default setting lets TRAMO select additive or multiplicative filtering and then decomposes the series into a trend, cycle, and seasonal component using X-11, with additive outliers allowed, as well as trading day dummies.

Table 3. Fiscal multiplier estimates

Multiplier type	min	16th p.	mean	median	84th. p	max
Spending multiplier (present value)	−1.81	0.63	0.94	0.99	1.22	2.43
Best 40%	−1.38	0.52	0.87	0.89	1.21	2.15
Tax multiplier (present value)	−2.30	−1.28	−0.76	−0.82	−0.23	1.92
Best 40%	−2.30	−1.23	−0.76	−0.84	−0.24	1.11
Spending multiplier (peak)	0.25	0.87	1.08	1.06	1.30	2.22
Best 40%	0.25	0.83	1.07	1.03	1.34	1.99
Tax multiplier (peak)	−2.17	−0.90	−0.58	−0.58	−0.19	−0.02
Best 40%	−2.17	−0.90	−0.59	−0.57	−0.22	−0.05

Source: Authors' calculations.

Notes: The descriptive statistics of the full set of results are based on 2,987 median multipliers estimates, whereas the group based on the 40% best-forecasting models consists of 1,196 multipliers. See also Fig. 1 for kernel densities.

deflated using the corresponding deflator (see Table 1).[10] The corresponding fiscal variables and GDP enter the (FA)VAR models in logs, while inflation and the interest rate are added to the VAR without further transformation (i.e. in percentage points). The methodological framework employed for the identification of fiscal shocks, which correspond to the standard specifications used in the modern literature on fiscal multipliers, implies that the variables in the VAR model are assumed to be stationary or trend-stationary (i.e. stationary around deterministic linear trend). All time series used to estimate the factors are transformed to reach stationarity prior to obtaining estimates of the factors.[11]

4. Fiscal multipliers in Austria: the role of forecasting performance and specification choices

The estimated fiscal multipliers for Austria are summarized in Table 3. We make use of out-of-sample predictive accuracy as a validation device of the models used in our exercise. We utilize the last four observations of our GDP series as an out-of-sample period and compute the mean absolute error (MAE) of one-step-ahead GDP predictions for all specifications used to obtain multiplier estimates, after estimating the models using a sample that excludes the out-of-sample observations. The results of this forecasting exercise allow us to refine the inference on Austrian expenditure and tax multipliers by concentrating on the estimates corresponding to the set of models with best predictive ability.

The mean present-value spending multiplier over all models is 0.68 and increases to 0.79 if we focus on the group of best models according to predictive ability (specifications corresponding to the 40% best models in terms of MAE). Generally, peak spending multipliers are larger than the present-value spending multipliers. The mean peak spending multiplier is 0.85 and reaches 0.90 in the group of models with best predictive power. As for the tax multipliers, the magnitude of present-value tax multiplier is quite high in

10 Revenue categories are not available in real terms. In order to investigate the effects of deflating with different price indices while keeping consistency, we choose to source all-time series in nominal terms and deflate them with the same deflator.

11 See Appendix A for the transformations carried out in each of the time series.

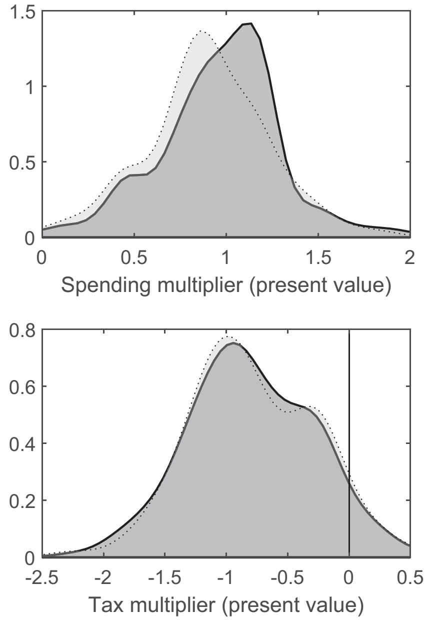


Fig. 1. Fiscal multiplier estimates: kernel densities. *Notes:* The dark density corresponds to the full set of results, the light density refers to the top 40% best models in terms of predictive ability. See also notes to Table 3.

absolute value at −1.12 and gets even larger when concentrating on the models with particularly good forecasting ability. The mean peak tax multiplier is −0.54 for the whole set of specifications entertained and −0.68 once we concentrate on the models with best forecasting performance. The smoothed densities of the estimated multipliers are presented in Fig. 1 for the full sample of fiscal multiplier estimates, as well as for the top 40% models in terms of out-of-sample predictive ability.

Across all specifications, focusing on the models with best predictive ability leads to larger multiplier estimates in absolute value. However, within certain types of specifications, sizeable differences can be found when zooming into the group of models which have a higher predictive power. The most pronounced differences between variants of the same type of specification are depicted in Fig. 2, which shows the empirical densities of peak tax multiplier for the full sample and for subsets based on predictive ability (best 20, 40, 60, and 80% models), split in three panels depending on the particular deflator used for nominal variables. The first panel shows that within the group of models that employ variables where the GDP deflator was used to transform nominal variables into their real counterparts, specifications with relatively good forecast performance tend to deliver tax multipliers of larger magnitude, with the mode of the distribution moving from approximately −0.4 to −0.9. A similar tendency is observed for models that use variables whereHICP was employed as a deflator, albeit in a less pronounced manner than for the GDP deflator.

For the cumulative spending multiplier, the effects of abstracting away from evaluating models with relatively poor forecasting performance are different in specifications when we use only a constant as a deterministic term in the (FA)VAR equation as compared with specifications in which we also add a time trend, with the results presented in Fig. 3.

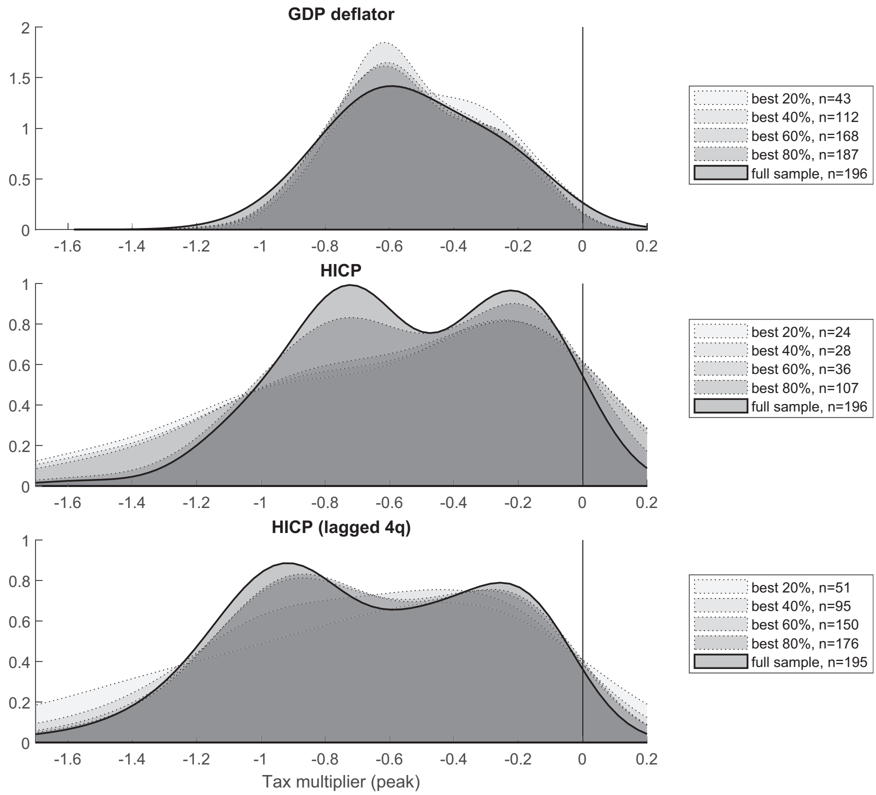


Fig. 2. Tax multiplier densities based on forecasting performance, split over deflating index. *Notes:* Each panel displays kernel densities calculated on subsets of multipliers according to the deflating index used in the (FA)VAR equation. The darkest density corresponds to the full set of results, the lighter ones correspond to subsets of models by predictive ability (best 20%, best 40%, best 60%, and best 80%).

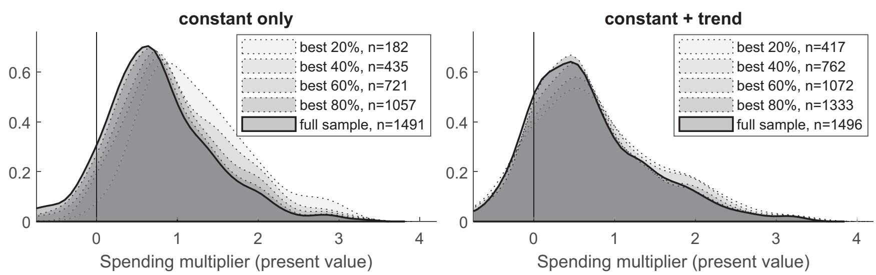


Fig. 3. Spending multiplier densities based on forecasting performance, split over the use of deterministic terms in the specification. *Note:* See notes to Fig. 2.

Cutting away 80% of the models based on their inferior forecasting performance leaves 417 (out of 1,496) models with a deterministic time trend, but only 182 (out of 1,491) models which feature only a constant. Such a result emphasizes the need to assess the non-stationary nature of the variables in the specification and cast doubts on the results based on models where macroeconomic variables are treated as stationary stochastic processes. In models with only a constant, focusing on the best predictive models shifts the whole distribution towards higher values of the spending multiplier (the mode of the distribution increases from approximately 0.6 to 0.9). For models with constant and trend, the picture is different: The distribution becomes flatter once we focus on multipliers obtained with models which have a particularly good forecasting performance, but the mode remains basically unchanged.

Table 4 summarizes the share of models with best forecasting performance in the full set of specifications by variable definition. The data composition which tends to improve forecasting performance for GDP data is the *Baseline* composition (tagged 'core/tax tiny'), which covers 17% of the models in the top 40% specifications by predictive ability. On the other side of the spectrum is a very similar data composition, which features *Baseline* revenues adjusted for actual social contributions ('core/tax small net soc.t.'), with a representation of 8.2% in the group of best forecasting models. As these two settings are very similar, we can identify the role played by particular components in terms of being responsible for differences in predictive ability across models. Models that include a tax variable that is adjusted for actual social contributions tend to have lower forecasting ability. If the researcher is interested in fiscal multipliers based on data compositions in models featuring good predictive ability, the *Baseline* ('core/tax tiny'), the 'corefix+soc.t.kind/tax mid', and the 'top down spend./top down rev.' variants appear particularly promising (see Table 2 for a description of data composition and Supplementary Appendix A for ESA codes).

Figure 4 shows multiplier estimates across different sets of government spending and revenue compositions. While most of the empirical densities obtained are relatively similar, three composition choices differ markedly from the rest. For the case of the spending multiplier (see top panels of Fig. 4), the composition including monetary social transfers ('core+m.soc.t./net tax small', inspired by Crespo Cuaresma *et al.*, 2011) leads to a distribution of multiplier estimates which has a similar mode as that of other data composition choices, but more mass around the mode. This indicates that adding monetary social transfers as part of spending composition leads to a higher precision for point estimates of the spending multiplier across models.

The sensitivity of spending multiplier estimates to the inclusion of monetary social transfers is a representative example of the importance of variable definitions and data

Table 4. Data composition and forecasting performance

Data composition	Count		Percentage	
	Total	Best 40%	Total	Best 40%
core/tax tiny	168	77	14.3	16.6
core/tax small net soc.t.	168	76	14.3	16.3
core/net tax small	168	57	14.3	12.3
corefix+soc.t.kind/tax mid	168	78	14.3	16.8
corefix+soc.t.kind/net tax mid	168	62	14.3	13.2
corefix+soc.t.kind/net tax large	168	45	14.3	9.7
core/net tax all	167	70	14.2	15.1
Total	1,175	465	100%	100%

Source: Authors' calculations.

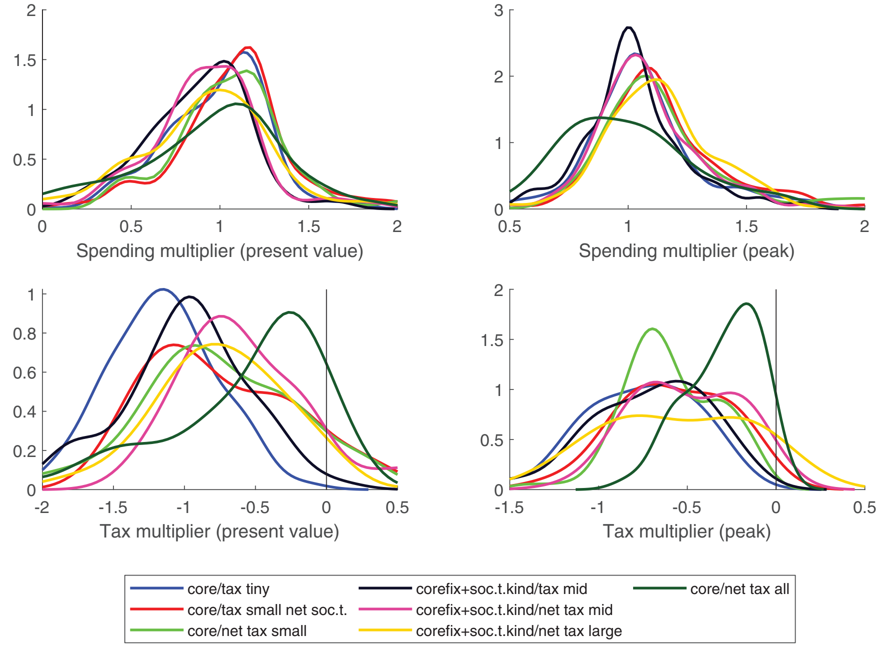


Fig. 4. Multiplier densities and data composition, based on all results. *Notes:* For the details of the data compositions, see Table 2. For the descriptive statistics and kernel densities based on the 40% best-forecasting models, see Supplementary Appendix E.

composition issues when it comes to fiscal multiplier estimates. In the case of Austria, changes of monetary social transfers (more than 20% of total expenditure) mainly reflect changes in pension payments. Despite the fact that pension payments are legally linked to the lagged national price index, VAR models tend to interpret many of the changes in monetary social transfers as exogenous impulses, which potentially decrease dispersion in the distribution of multiplier estimates. The second data composition set worth discussing is the only one constructed using a top-down approach, starting from total spending and total revenues, which are subsequently netted of subsidies and transfers ('top down spend./top down rev'). As Fig. 4 shows, the spending multipliers corresponding to models that include these variables tend to be more concentrated around a value of zero. In models that consider such a broad definition of government spending, changes in the variable are more likely to be interpreted as exogenous impulses. Besides the ignored endogenous reaction of monetary social transfers already discussed above, changes of interest payments (which are the part of government spending in this broad variable definition) should also be not treated as exogenous fiscal policy impulses, as governments have only limited power to influence interest payments in the short run. Our results further highlight that for tax multipliers, the choice of a particular group of fiscal variables in the model may have a larger effect on multiplier estimates than in the case of spending multipliers. The empirical distributions of multiplier estimates tend to be rather flat for certain cases, while a composition set including capital transfers ('core/net tax all', inspired by Muir and Weber, 2013), deliver more precise tax multiplier estimates (albeit relatively low in magnitude). The lower magnitude of tax multiplier is also due to misleading identification of exogenous shocks, especially for a revenue variable (net taxes) that includes capital transfers. In recent years, virtually all of the variations in capital transfers in Austria have been due to sizable banking support programmes, which arguably had only mild effects on GDP. This leads to more precise but lower magnitudes of (net)-tax multipliers once capital transfers are included, however, providing little information on how common taxes affect output.

Turning to the effects of using different econometric specifications, identification strategies, and number of variables (see Fig. 5), on average, models with three variables and a shock identification design based on the Cholesky decomposition tend to result in lower spending multiplier point estimates compared with models which employ more variables and different identification schemes. Whereas VAR models with three variables or models estimated with Cholesky ordering lead to median spending multipliers around 0.5, following more modern approaches can yield spending multiplier estimates with a median above

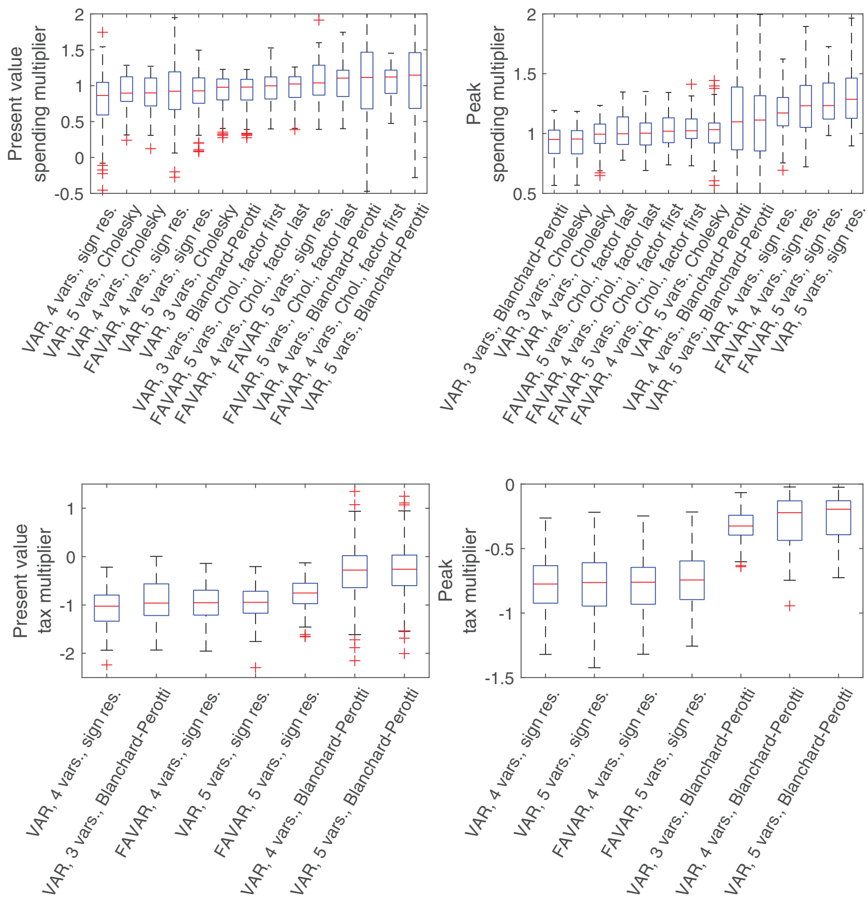


Fig. 5. Fiscal multipliers by model and identification strategy types. *Notes:* Boxplots are sorted by the median multiplier, the central (red) mark of the boxplot. The bottom and top edges of the box indicate the 25th and 75th percentiles.

unity. Similar patterns hold for peak tax multipliers, but the differences are smaller: models with fewer covariates and employing the Cholesky identification scheme tend to result in a median peak tax multiplier around –0.5, whereas the approach delivering the highest median magnitude (VAR model with five variables estimated with sign restrictions) reaches –0.7. As is evident in Fig. 5, based on peak responses, present-value tax multiplier estimates have a much larger spread than their counterparts.

Varying the output elasticity of taxes used to calibrate the identification schemes based on the Blanchard–Perotti method has negligible effect on spending multipliers, but a notable effect on tax multipliers, especially when calculated as present-value tax multiplier. The effect is larger in VAR models with three variables than in VARs with four or five variables. Increasing the output elasticity of taxes from its baseline setting of 1.66 to 2 reduces the average present-value tax multiplier by 0.3 in VARs with three variables, and by 0.1 in VARs with four and five variables. Varying the price elasticity of taxes, which is only present in VAR models with four and five variables, causes changes in the estimates in both spending and tax multipliers. Doubling the price elasticity of taxes from the baseline value of 0.78 to 1.5 increases both the present-value and the peak spending multiplier by approximately 0.3. The effect of the same change on tax multipliers is, however, very different if we focus on present-value or peak tax multiplier. In case of present-value tax multiplier, the change in the price elasticity pushes the multiplier towards unity, whereas the peak multiplier is largely unaffected.[12]

We assess subsample stability in the estimation of multipliers by means of discarding one (first or last) observation at a time and re-estimating the multipliers. We thus investigate the possible effects of influential observations at the beginning or the end of the sample on the multiplier estimates. The main result of the analysis is that peak multipliers are much more stable than present-value multipliers. In particular, present-value tax multipliers appear sensitive to discarding initial observations: discarding the observations corresponding to 2002 from the sample lowers the magnitude of the mean multiplier from −1.12 to −0.97, and the estimate goes down further to −0.75 if we eliminate the observations corresponding to 2003. Spending multipliers display some variability when changing the estimation sample. Present-value spending multiplier estimates get considerably lower once the first quarter of 2018 is considered in the recursive analysis (we observe a drop in mean present value spending multiplier from 0.96 to 0.64). Peak spending multipliers are subject to similar drop in the same time frame (from 1.02 to 0.85), but the values of the peak multiplier are generally higher than their present-value multiplier counterparts. The peak spending multiplier is rather robust to discarding observations from the beginning of the time frame, whereas the present-value multiplier drops once years 2002 and 2003 are removed from the sample (from 0.60 to 0.48 to 0.44). Detailed results on the subsample stability exercise can be found in Supplementary Appendix B. In addition, we also investigate the effects of adding different dummy variables to account for the potential effects of the financial crisis. The results indicate that (present-value) spending multiplier tends to have a higher magnitude once we control for the particularities of our model variables during the financial crisis. The peak spending multiplier and the tax multiplier are mostly unaffected by adding a crisis dummy. For detailed results, see Supplementary Appendix F.

12 See Appendix C for more detailed results on these robustness checks.

5. Conclusions

This article estimates fiscal multipliers for Austria, a stereotypical advanced small open economy, with a focus on the dimension of model uncertainty that emanates from the choice of a particular econometric model to obtain point estimates of the reaction of GDP to shocks in fiscal variables. We present a comprehensive framework that allows to assess the effects of different multiplier definitions and choices related to the data, the model employed, and further technical choices associated with the specification of the model exert on fiscal multiplier estimates.

The mean present-value spending multiplier over all models entertained is 0.68 and increases to 0.79 once we focus on the best models according to out-of-sample predictive ability. Generally, estimates of the peak spending multiplier for Austria tend to be larger than present-value spending multipliers. The mean peak spending multiplier is 0.85 and reaches 0.90 if calculated on the basis of the group of models with best predictive performance. As for the tax multipliers, the magnitude of the present-value tax multiplier is relatively high, with an average value across specifications of −1.12 and gets even larger in absolute value when concentrating on the best models in terms of predictive ability. The mean peak tax multiplier is −0.54 for all specifications used and −0.68 once we concentrate on the models with the best forecast performance.

For some multiplier definitions and modelling choices, major differences in estimates are found if we focus on the set of models with best predictive ability. Our results indicate that if the GDP deflator is used to deflate nominal variables, concentrating on best performing models leads to a larger peak tax multiplier in absolute value (the mode of the distribution shifts from approximately −0.4 to −0.9). Comparable results are found when we focus on forecasting performance and split models over different compositional definitions of government expenditures and taxes. The particular composition that delivers the highest percentage of models that predict well uses compensation of employees, intermediate consumption, and gross capital formation as part of government expenditures and taxes on production, imports, income, and wealth.

On average, multipliers obtained from models that require few variables and use Cholesky identification for the structural shocks tend to result in lower estimates of the spending multiplier. On the other hand, using more variables for estimation and employing identification schemes that follow the Blanchard–Perotti approach or sign restrictions deliver results with rather higher estimates of spending multipliers. Similar patterns hold for peak tax multipliers, but the differences are smaller.

Our analysis provides evidence that in a framework of model uncertainty in terms of the specification used to calculation of fiscal multipliers, concentrating on the subgroup of models that present good forecasting ability can deliver different results than assessing the full set of potential specifications. In line with conclusions in Ramey (2019), we find that the specific way used to obtain multipliers can make a big difference in terms of inference. Given the scarce evidence on multipliers in developed small open economies, the results we present for Austria have a value of their own for policymakers and fiscal authorities.

Supplementary material

Supplementary material is available on the OUP website. These are the data and replication files, as well as the Supplementary appendix.

Funding

This work was supported by the Czech Science Foundation (17-14263S to J.Č. and J.C.C.). Computational resources were supplied by the project 'e-Infrastruktura CZ' (e-INFRA LM2018140) provided within the program Projects of Large Research, Development and Innovations Infrastructures.

Acknowledgements

The authors would like to thank Rafael Domenech and two anonymous referees, as well as participants in the workshop *Fiscal Multipliers and Transfer Payments*, organized by the Austrian Fiscal Advisory Council, for helpful comments on an earlier version of this paper.

References

Afonso, A. and Sousa, R.M. (2011) The macroeconomic effects of fiscal policy in Portugal: a bayesian SVAR analysis, *Portuguese Economic Journal*, **10**, 61–82.

Auerbach, A.J. and Gorodnichenko, Y. (2012) Measuring the output responses to fiscal policy, *American Economic Journal: Economic Policy*, **4**, 1–27.

Barrell, R., Holland, D. and Hurst, I. (2012) Fiscal consolidation: Part 2. Fiscal multipliers and fiscal consolidations. OECD Working Papers, Organisation for Economic Co-Operation and Development (OECD), Paris, France.

Beetsma, R. and Giuliodori, M. (2011) The effects of government purchases shocks: review and estimates for the EU, *The Economic Journal*, **121**, F4–F32.

Benetrix, A.S. and Lane, P.R. (2009) The impact of fiscal shocks on the Irish economy, *The Economic and Social Review*, **40**, 407.

Blanchard, O. and Perotti, R. (2002) An empirical characterization of the dynamic effects of changes in government spending and taxes on output, *The Quarterly Journal of Economics*, **117**, 1329–68.

Blanchard, O.J. and Leigh, D. (2013) Growth forecast errors and fiscal multipliers, *American Economic Review*, **103**, 117–20.

Breuss, F., Kaniovski, S. and Schratzenstaller, M. (2009) Macro-economic effects of the fiscal stimulus measures in Austria, *Austrian Economic Quarterly*, **14**, 205–16.

Burriel, P., De Castro, F., Garrote, D., Gordo, E., Paredes, J. and Pérez, J.J. (2010) Fiscal policy shocks in the euro area and the US: an empirical assessment, *Fiscal Studies*, **31**, 251–85.

Caggiano, G., Castelnuovo, E., Colombo, V. and Nodari, G. (2015) Estimating fiscal multipliers: news from a non-linear world, *The Economic Journal*, **125**, 746–76.

Caldara, D. and Kamps, C. (2017) The analytics of SVARs: a unified framework to measure fiscal multipliers, *The Review of Economic Studies*, **84**, 1015–40.

Čapek, J. and Crespo Cuaresma, J. (2020) We just estimated twenty million fiscal multipliers, *Oxford Bulletin of Economics and Statistics*, **82**, 483–502.

Crespo Cuaresma, J., Eller, M. and Mehrotra, A. (2011) The Economic transmission of fiscal policy shocks from Western to Eastern Europe. BOFIT Discussion Papers 2011, 3, Suomen Pankki, Siirtymatalouksien Tutkimuslaitos, Helsinki, Finland.

Forni, M. and Gambetti, L. (2014) Sufficient information in structural VARs, *Journal of Monetary Economics*, **66**, 124–36.

Fragetta, M. and Gasteiger, E. (2014) Fiscal foresight, limited information and the effects of government spending shocks, *Oxford Bulletin of Economics and Statistics*, **76**, 667–92.

Gechert, S. (2015) What fiscal policy is most effective? A meta-regression analysis, *Oxford Economic Papers*, **67**, 553–80.

Gechert, S. and Rannenberg, A. (2018) Which fiscal multipliers are regime-dependent? A meta-regression analysis, *Journal of Economic Surveys*, **32**, 1160–82.

Hebous, S. (2011) The effects of discretionary fiscal policy on macroeconomic aggregates: a reappraisal, *Journal of Economic Surveys*, **25**, 674–707.

Ilzetzki, E., Mendoza, E.G. and Végh, C.A. (2013) How big (small?) are fiscal multipliers?, *Journal of Monetary Economics*, **60**, 239–54.

Jemec, N., Kastelec A.S. and Delakorda A. (2011) *How Do Fiscal Shocks Affect the Macroeconomic Dynamics in the Slovenian Economy*, Prikazi in Analize, Banka Slovenije, Ljubliana, Slovenia.

Karras, G. (2014) Fiscal policy spillovers through trade openness, *Journal of Economic Integration*, **29**, 563–81.

Kilponen, J., Pisani, M., Schmidt, S., Corbo, V., Hledik, T., Hollmayr, J., Hurtado, S., Júlio, P., Kulikov, D., Lemoine, M., *et al.* (2019) Comparing fiscal consolidation multipliers across models in Europe, *International Journal of Central Banking*, **15**, 285–320.

Koch, S., Monárová, Z. and Reiter, M. (2019) Macro-economic effects of the fiscal stimulus measures in Austria. Mimeo, Institute for Advanced Studies, Vienna, Austria.

Mineshima, A., Poplawski-Ribeiro, M. and Weber, A. (2014) Size of fiscal multipliers, in C. Cottarelli, P. Gerson, and A. Senhadji (eds) *Post-Crisis Fiscal Policy*, MIT Press, Massachusetts, USA, 315–372.

Mountford, A. and Uhlig, H. (2009) What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*, **24**, 960–92.

Mourre, G., Astarita, C., Princenare, S. (2014) Adjusting the budget balance for the business cycle: the EU methodology. Technical Report, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission, Brussels, Belgium.

Muir, D. and Weber, A. (2013) Fiscal multipliers in Bulgaria: low but still relevant. Technical Report, IMF Working Paper No. 13/49, International Monetary Fund, Washington DC, USA.

Perotti, R. (2004) Estimating the effects of fiscal policy in OECD countries. Technical Report, IGIER Working Paper No. 276, Bocconi University, Milano, Italy.

Price, R.W., Dang, T.T. and Guillemette, Y. (2014) New tax and expenditure elasticity estimates for EU budget surveillance. OECD Economics Department Working Papers No. 1174, Organisation for Economic Co-Operation and Development (OECD), Paris, France.

Ramey, V.A. (2011a) Can government purchases stimulate the economy?, *Journal of Economic Literature*, **49**, 673–85.

Ramey, V.A. (2011b) Identifying government spending shocks: it's all in the timing, *The Quarterly Journal of Economics*, **126**, 1–50.

Ramey, V.A. (2016) Macroeconomic shocks and their propagation, *Handbook of Macroeconomics*, **2**, 71–162.

Ramey, V.A. (2019) Ten years after the financial crisis: what have we learned from the renaissance in fiscal research?, *Journal of Economic Perspectives*, **33**, 89–114.

Ravn, S.H. and Spange, M. (2012) The effects of fiscal policy in a small open economy with a fixed exchange rate: the case of Denmark. Technical Report, Danmarks Nationalbank Working Papers, No. 80, Copenhagen, Denmark.

Rubio-Ramírez, J.F., Waggoner, D.F. and Zha, T. (2010) Structural vector autoregressions: theory of identification and algorithms for inference, *The Review of Economic Studies*, **77**, 665–96.

Schuster, P. (2019) On fiscal multipliers in New Keynesian small open economy models. Mimeo, Office of the Austrian Fiscal Advisory Council, Vienna, Austria.

Stock, J.H. and Watson, M.W. (2016) Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics, *Handbook of Macroeconomics*, **2**, 415–525.

Appendix to
"Fiscal multipliers in a small open economy: the case of Austria"
by J. Čapek, J. Crespo Cuaresma, J. Holler and P. Schuster

A Data

Figure A.1: Government spending, deflated by GDP deflator.

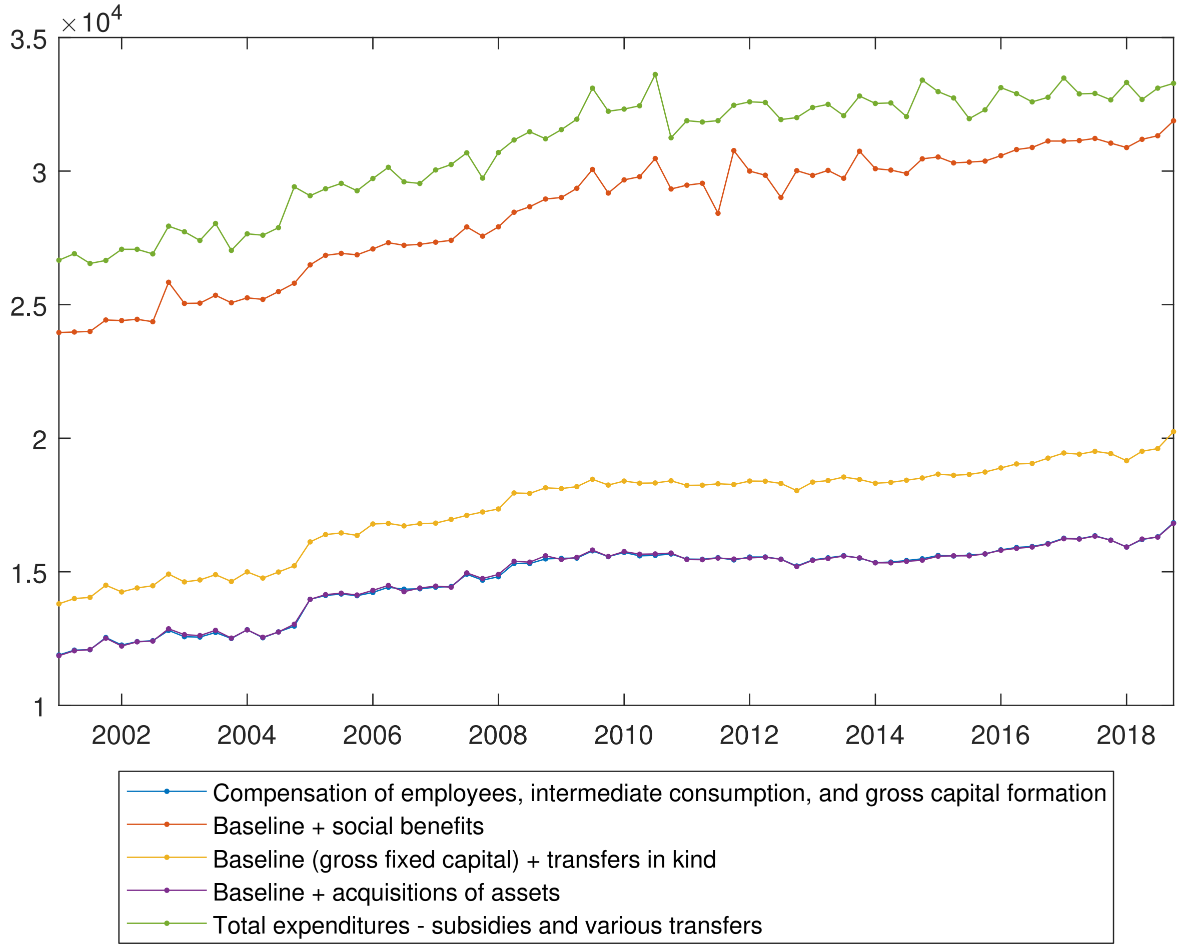


Figure A.2: Government revenue, deflated by GDP deflator.

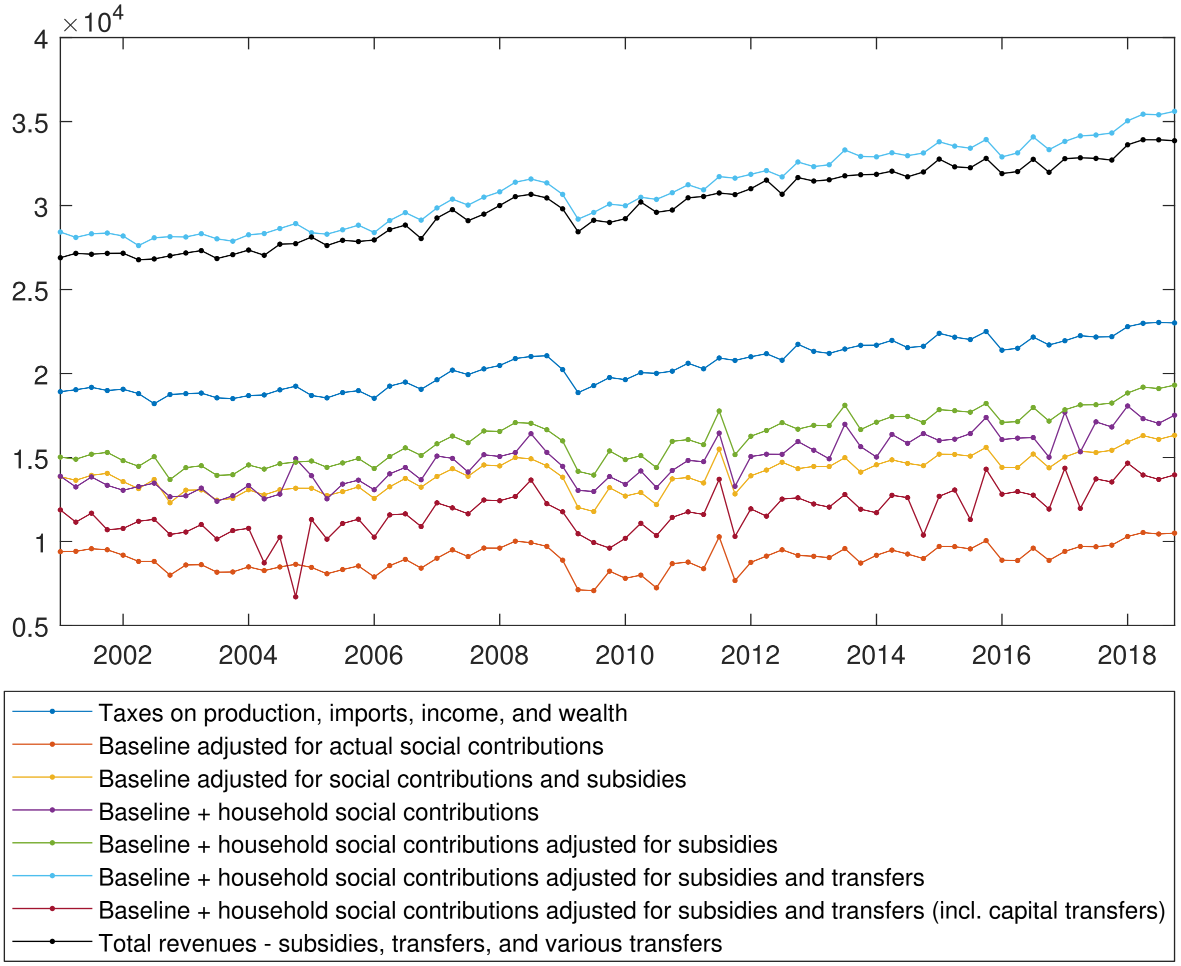


Table A.1: Time series employed for the computation of the factors

Source	Code	Series	Tr.
ECB	BSI,M,N,A,A25,A,1,U6,2250,Z01,E	Domestic credit for consumption (and other) to households (and other), currencies combined, stocks	5
ECB	BSI,Q,N,A,A20,A,1,U6,2000,EUR,E	Domestic loans from MFIs to non-MFIs, Euro	5
ECB	BSI,M,N,A,L60,X,4,Z5,0000,Z01,E	Capital and reserves, unspecified, flows	1
ECB	BSI,M,N,A,A20,A,4,U6,1000,Z01,E	Domestic loans to monetary financial institutions (MFIs), Euro, flows	1
Eurostat	ei_bsco_q/BS-HI-NY,SA,BAL	Home improvements over the next 12 months	2
Eurostat	ei_bsin_q_r2/BS-ICU-PC,SA	Current level of capacity utilization (percent)	5
Eurostat	ei_bsin_q_r2/BS-INO-BAL,SA	New orders in recent months	2
Eurostat	ei_bssi_m_r2/BS-CSMCI-BAL,SA	Consumer confidence indicator	2
Eurostat	ei_bssi_m_r2/BS-ESI-I,SA	Economic sentiment indicator	5
Eurostat	ei_isbr_m/RT12-CA,F_CC1,IS-IP	Production index	2
Eurostat	ert_eff_ic_q/REER_EA19_CPI,I10	Real effective exchange rate (deflator: consumer price index - 19 trading partners - euro area)	5
Eurostat	irt_lt_mcby_q/MCBY	EMU convergence criterion bond yields	2
Eurostat	lfsi_emp_q/THS_PER,T,ACT,Y15-64,SA	Employment - Active population age	5
Eurostat	lfsi_emp_q/ THS_PER,T,EMP_LFS,Y15-64,SA	Total employment (resident population concept - LFS)	5
Eurostat	lfsq_egais/THS,T,Y_GE15,EMP,OC8	Employed persons - Plant and machine operators and assemblers	5
Eurostat	lfsq_egais/THS,T,Y_GE15,EMP,OC5	Employed persons - Service and sales workers	5
Eurostat	lfsq_ewhuis/HR,T,TOTAL,EMP,OC8	Hours worked - Plant and machine operators and assemblers	5
Eurostat	namq_10_gdp/CLV10_MNAC,SCA,P51G	Gross fixed capital formation	5
Eurostat	namq_10_gdp/ CLV10_MNAC,SCA,P31_S14	Final consumption expenditure of households	5
Eurostat	namq_10_gdp/ CLV10_MNAC,SCA,P32_S13	Collective consumption expenditure of general government	5
Eurostat	namq_10_gdp/CLV10_MNAC,SCA,P6	Exports of goods and services	5
Eurostat	namq_10_gdp/CLV10_MNAC,SCA,P7	Imports of goods and services	5
Eurostat	namq_10_gdp/PD10_NAC,SCA,B1GQ	Price index (implicit deflator)	5
Eurostat	nasq_10_f_bs/MIO_NAC,S1,LIAB,F2	Liabilities - Currency and deposits	5
Eurostat	nasq_10_f_bs/MIO_NAC,S1,LIAB,F4	Liabilities - Loans	5
Eurostat	une_rt_q/SA,TOTAL,THS_PER,T	Unemployed	5

Note: 'Tr.' indicates the transformation applied to the series (1 = level, 2 = first difference, 3 = logarithm, 4 = second difference, 5 = first difference of logarithm, 6 = second difference of logarithm).

Table A.2: Government spending and revenue composition

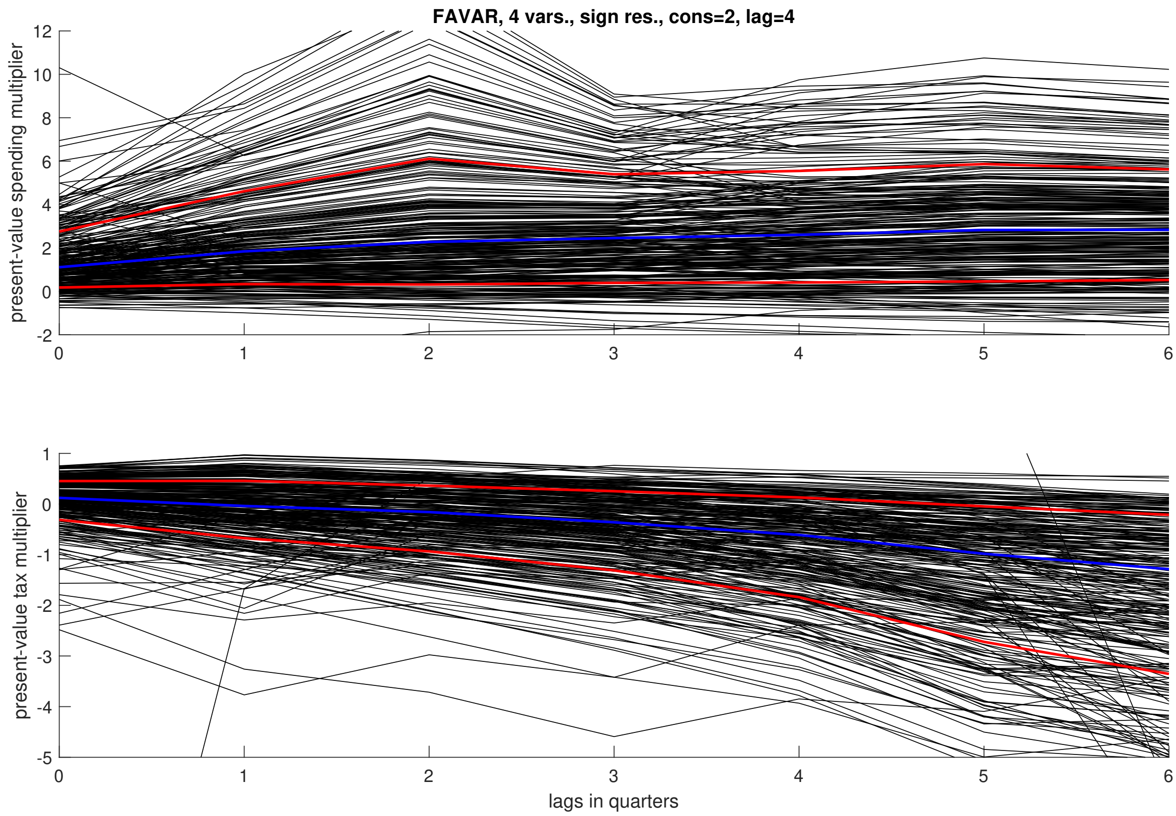
Tag	Gov't spending composition	Gov't revenues composition
core/tax tiny	D1PAY + P2 + P5	D2REC + D5REC
core/tax small net soc.t.	D1PAY + P2 + P5	D2REC + D5REC + D611REC - D62PAY - D632PAY
core/net tax small	D1PAY + P2 + P5	D2REC + D5REC + D61REC - D62PAY - D632PAY - D3PAY
corefix+soc.t.kind/tax mid	D1PAY + P2 + P51G + D632PAY	D2REC + D5REC + D611REC + D613REC + D91REC
corefix+soc.t.kind/net tax mid	D1PAY + P2 + P51G + D632PAY	D2REC + D5REC + D611REC + D613REC + D91REC - D3PAY - D62PAY
corefix+soc.t.kind/net tax large	D1PAY + P2 + P51G + D632PAY	D2REC + D5REC + D611REC + D613REC + D7REC + D91REC - D3PAY - D62PAY - D7PAY
core/net tax all	D1PAY + P2 + P5 + NP	D2REC + D5REC + D61REC + D7REC + D9REC - D62PAY - D632PAY - D3PAY - D7PAY - D9PAY

Note: Source of data is Eurostat, the codes follow ESA2010 system.

B Depiction of bootstraps (example)

An exemplary depiction of the results of the bootstrap for present-value spending and tax multiplier are below. The black lines are the bootstrapped values,[1] the blue line is the median multiplier and the red lines denote the 16th, and 84th percentiles, respectively.

Fiscal multiplier: bootstrap replications using sign restrictions



Note: FAVAR model with 4 variables (one factor) and 4 lags, constant and trend included, sign restriction identification. Government spending composition includes compensation of employees, intermediate consumption, and gross capital formation. Government revenues include taxes on production, imports, income, and wealth, adjusted for actual social contributions. Nominal data deflated by (4q)-lagged HICP.

[1] The full set of 4,000 bootstraps is thinned for readability: only each 15th bootstrap is drawn.

C Robustness check: The effect of the financial crisis

Figure C.1: Fiscal multiplier densities with financial crisis dummies

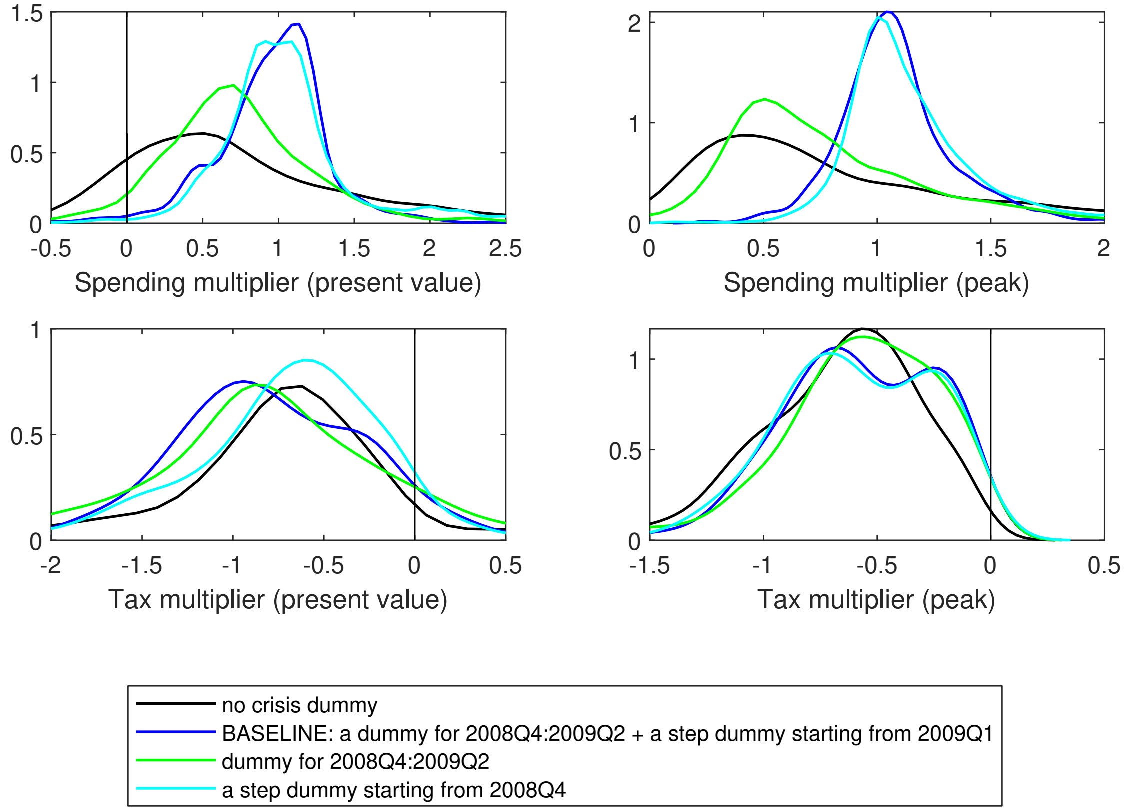


Table C.1: Fiscal multiplier estimates, no crisis dummy

Multiplier type	**min**	**16-th p.**	**mean**	**median**	**84-th. p**	**max**
Spending multiplier (present value)	-4.42	0.03	0.69	0.57	1.46	3.54
— best 40%	-1.37	0.12	0.79	0.65	1.49	3.40
Tax multiplier (present value)	-6.47	-1.95	-0.97	-0.71	-0.22	3.75
— best 40%	-6.47	-2.66	-1.28	-0.83	-0.48	2.95
Spending multiplier (peak)	-0.68	0.27	0.84	0.65	1.52	3.59
— best 40%	-0.68	0.34	0.87	0.69	1.57	3.48
Tax multiplier (peak)	-2.78	-1.04	-0.68	-0.63	-0.34	-0.03
— best 40%	-2.78	-1.16	-0.83	-0.76	-0.44	-0.10

Table C.2: Fiscal multiplier estimates, dummy for 2008Q4:2009Q2

Multiplier type	**min**	**16-th p.**	**mean**	**median**	**84-th. p**	**max**
Spending multiplier (present value)	-4.77	0.24	0.67	0.68	1.15	4.15
— best 40%	-4.77	0.48	0.77	0.76	1.24	2.83
Tax multiplier (present value)	-3.98	-1.43	-0.81	-0.82	-0.19	1.74
— best 40%	-3.98	-1.59	-0.95	-0.92	-0.33	1.31
Spending multiplier (peak)	-1.22	0.39	0.75	0.66	1.18	3.30
— best 40%	-1.22	0.48	0.79	0.70	1.24	2.39
Tax multiplier (peak)	-2.56	-0.88	-0.58	-0.55	-0.21	0.01
— best 40%	-2.56	-0.98	-0.64	-0.60	-0.23	0.00

Table C.3: Fiscal multiplier estimates, a step dummy starting from 2008Q4

Multiplier type	**min**	**16-th p.**	**mean**	**median**	**84-th. p**	**max**
Spending multiplier (present value)	-1.87	0.68	1.04	0.99	1.32	3.41
— best 40%	-0.41	0.71	1.01	0.98	1.28	3.16
Tax multiplier (present value)	-2.57	-1.19	-0.67	-0.63	-0.18	1.94
— best 40%	-2.57	-1.31	-0.75	-0.70	-0.28	1.44
Spending multiplier (peak)	0.09	0.92	1.18	1.08	1.42	3.57
— best 40%	0.22	0.96	1.16	1.07	1.37	3.36
Tax multiplier (peak)	-1.97	-0.91	-0.59	-0.60	-0.21	-0.00
— best 40%	-1.97	-0.97	-0.64	-0.65	-0.24	-0.00

Fiscal multipliers in a small open economy: the case of Austria*

Jan Čapek[1], Jesús Crespo Cuaresma[1,2,3,4,5], Johannes Holler[6], and Philip Schuster[6]

[1]Masaryk University (MU)
[2]Vienna University of Economics and Business (WU)
[3]International Institute of Applied System Analysis (IIASA)
[4]Wittgenstein Center for Demography and Global Human Capital (IIASA,VID/OEAW,WU)
[5]Austrian Institute of Economic Research (WIFO)
[6]Office of the Austrian Fiscal Advisory Council

Abstract

We estimate fiscal multipliers for Austria in a framework of model uncertainty emanating from the choice of a particular econometric model. We present a comprehensive framework which allows to assess the effects of different multiplier definitions and choices related to the data, the model employed, and further technical choices associated with the specification of the model exert on fiscal multiplier estimates. The mean present-value government spending multiplier over all models entertained, based on over one thousand estimates, is 0.94. Estimates of the peak spending multiplier tend to be larger than present-value spending multipliers, with a mean value of 1.08. The value of the mean present-value tax multiplier is -0.76 and the mean peak tax multiplier is -0.58 for all specifications used.

Keywords: Fiscal multiplier, structural VAR, predictive ability, small open economy, Austria

JEL codes: E62, C32

*Corresponding author: Jesús Crespo Cuaresma, Vienna University of Economics and Business. E-mail: jcrespo@wu.ac.at.
The authors would like to thank Rafael Domenech and two anonymous referees, as well as participants in the workshop *Fiscal Multipliers and Transfer Payments*, organized by the Austrian Fiscal Advisory Council, for helpful comments on an earlier version of this paper. Financial support from the Czech Science Foundation, Grant 17-14263S, is gratefully acknowledged. Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures".

1 Introduction

The interest in assessing the macroeconomic effects of fiscal policy in industrialized countries has gained renewed momentum since the Great Recession. Given the limited scope of action of monetary policy in the context of very low nominal interest rates, fiscal policy re-emerged as a policy of choice and a large literature has concentrated on investigating how fiscal policy affects macroeconomic variables and GDP in particular.[1] A convenient way to communicate the effects of fiscal stimulus on the economy is the fiscal multiplier, measured as the dollar reaction of GDP as a result of a one dollar fiscal stimulus. Fiscal multipliers are easily comparable across countries and over time, and the precision of their estimation contributes significantly to the quality of GDP growth predictions (Blanchard and Leigh, 2013). The estimates of fiscal multipliers are infamously heterogeneous both across countries and methods used for their calculation, and may be very sensitive to arguably minor specification choices, as recently shown in Čapek and Crespo Cuaresma (2020).

There is little evidence on the size of fiscal multipliers for developed European small open economies.[2] Ravn and Spange (2012) enhance the Blanchard-Perotti methodology based on structural vector autoregression (SVAR) models to estimate spending multipliers for Denmark and obtain a point estimate of approximately 0.6 after four quarters. Jemec et al. (2011) investigate Slovenian fiscal policy employing a standard SVAR approach and estimate an impact spending multiplier of 1.5, which diminishes in subsequent periods. Unfortunately, not all studies investigating the effects of fiscal stimuli report the results in the form of multipliers (e.g. Afonso and Sousa, 2011, for Portugal or Benetrix and Lane, 2009, for Ireland). In addition to estimates for single countries, evidence from panel studies also exists. Ilzetzki et al. (2013) report that the subgroups of countries corresponding to high income, open, low-debt and fixed exchange rate countries have average spending multipliers of 0.4, 0, 0.2, and 0.6, respectively. The empirical evidence can be supplemented making use of the work by Barrell et al. (2012), where a model-based consumption multiplier of 0.5 is reported for Austria. Breuss et al. (2009) provides an overview of fiscal multipliers derived by Austrian forecasting institutions from large-scale macroeconometric models (within the tradition of the *Cowles commission approach*). Spending multipliers over the first year after the fiscal shock are typically below unity, first year wage and income tax multipliers are below 0.5. Recent papers by Koch et al. (2019) and Schuster (2019) complement the existing results by simulating fiscal multipliers for Austria using calibrated New-Keynesian general equilibrium models and derive multipliers of comparable magnitudes. However, to our knowledge, a pure empirical assessment of fiscal multipliers specifically for Austria, as a stereotypical small open economy within the group of industrialized countries, does not exist. In this contribution, we provide for the first time a rigorous analysis of fiscal multiplier estimates in a small open economy (Austria) incorporating the uncertainty related to specification choice in several dimensions including that related to the particular variables included in the model, shock identification strategies, data preparation or the analytical structure of the model. Given the importance of economic openness to determine the size of the fiscal multiplier, such an exercise allows the results to be interpreted in the framework of theoretical models of fiscal policy effects in small open economy settings. Theoretical results of this literature predict lower domestic effects of fiscal policy through the leaking of fiscal shocks to imported goods, combined with a higher sensitivity to international economic policy spillovers (see Karras, 2014, for example).

The main bulk of the existing literature on the macroeconomic effects of fiscal interventions can be categorized as either model-based or empirical. Model-based approaches typically employ calibrated DSGE models to study the effects of fiscal stimuli in an internally-consistent theoretical framework. Kilponen et al. (2015), for instance, compare such estimates of fiscal multipliers across models and countries in

[1] See e.g. Hebous (2011) or Ramey (2011a) for earlier surveys on the issue, or Ramey (2019) for a recent contribution.

[2] See the extensive summary of existing multiplier estimates in Mineshima et al. (2014) or the data used for the broad meta-analysis in Gechert (2015).

Europe, while Barrell et al. (2012) focus on model-based fiscal multipliers in the context of fiscal consolidation. The advantage of the model-based approach lies in the ability to analyse counterfactual scenarios by simulating the dynamics of the model variables under different conditions. On the other hand, empirical approaches, mostly based on SVAR models, tend to be more data-driven and typically impose less stringent restrictions on the structure of the economic model. The availability of long time series for some countries allow for the use of modern identification methods such as the narrative approach (Ramey, 2011b) to extract exogenous fiscal shocks or the assessment of different regimes (Auerbach and Gorodnichenko, 2012) where fiscal multipliers may differ. In cases where such long time series are not available, countries are often pooled and the empirical analysis is conducted on a panel setting (Beetsma and Giuliodori, 2011; Ilzetzki et al., 2013), or fiscal multipliers for single economies with shorter time series are studied using SVAR models inspired by the seminal contribution by Blanchard and Perotti (2002).[3]

The estimates of fiscal multipliers tend to differ, sometimes strongly, from study to study (see the evidence presented in the meta-analysis provided by Gechert, 2015). These differences can be attributed to various identification strategies (Caldara and Kamps, 2017) as well as to other technical choices made in the analysis (Čapek and Crespo Cuaresma, 2020). Given the additional dimension of uncertainty on fiscal multiplier estimates implied by the particular methodological choices, even within the class of SVAR models, the approach of this study is to present a consistent framework which encompasses a wide range of reasonable settings and choices which are routinely used in the empirical literature on fiscal multipliers. The framework delivers over one thousand multiplier estimates, each for a particular model specification. We exploit the differences in out-of-sample predictive power of the models entertained for GDP in order to gain insights into the size of fiscal multipliers in Austria. Our analysis expands the methodological setting put forward in Čapek and Crespo Cuaresma (2020) in several respects. First of all, by concentrating on a single economy, we gain comparability in the multiplier estimates, which correspond to the responses to fiscal impulses within the same institutional and historical setting. Furthermore, we expand the set of econometric specifications and modelling choices in Čapek and Crespo Cuaresma (2020) by including new models based on factor-augmented VAR structures and using out-of-sample predictive ability as a model selection tool. The focus on a single small open economy allows us to link the results in a more direct manner to the methodological framework provided by economic theory, in particular when interpreting the results of the analysis, and allows for the assessment of additional sources of model uncertainty as compared to Čapek and Crespo Cuaresma (2020). This is the case, for example, for the composition of government spending and tax aggregates, or for the calculation of the values of tax and spending elasticities required for several identification techniques. In our analysis, we also contribute to the literature by identifying structural fiscal shocks in models where subcomponents of spending and tax revenues are used, making use of elasticities of disaggregated components of the fiscal variables to output and the price level obtained using the fiscal forecasting model by the Austrian Fiscal Advisory Council (2014).

Our results expose the uncertainty and heterogeneity that is inherent to empirical estimates of fiscal multipliers. In addition to entertaining different SVAR specifications based on Blanchard and Perotti (2002) and Perotti (2004), we also estimate fiscal multipliers from structural Factor Augmented VAR (FAVAR) models. These specifications provide a more adequate framework to account for fiscal foresight and omitted variable biases (Fragetta and Gasteiger, 2014). Furthermore, we also exploit the existing data on government spending and tax composition in Austria in order to obtain additional multiplier estimates. We compare the results for the two most widely used formulations in the literature – the present-value multiplier and the peak multiplier and deliver the first set of credible multiplier estimates for a representative European small open economy after accounting for model uncertainty.

[3] See e.g. Ramey (2016) for a review of the methods used for the identification of exogenous fiscal shocks.

The mean spending multiplier for Austria is estimated at 0.94 for the present-value multiplier and 1.08 for the peak multiplier. The present-value tax multiplier is -0.76 and its peak counterpart is -0.58. Comparing the multipliers to the existing literature, our estimates suggest a stronger reaction of GDP after the increase of government spending as compared to the results for relevant subgroups of countries reported in Ilzetzki et al. (2013). Our estimate of present-value multiplier specification is comparable to that of Denmark (see Ravn and Spange, 2012). As in the case of the study on the Slovenian economy, our results also suggest that peak spending multipliers tend to be higher than their present-value counterparts (see Jemec et al., 2011). The multiplier estimates obtained using the subset of models with relatively superior predictive ability for GDP tend to be smaller in case of present value spending multiplier. Our results also indicate that the models based on subcomponents of government spending and taxes that deliver the best predictive ability for GDP dynamics tend to include compensation of employees, intermediate consumption, gross capital formation, and transfers in kind as part of government expenditures and taxes on production, imports, income, and wealth, and household social contributions. On average, SVAR models of small dimension and using the Cholesky decomposition as an identification device tend to result in relatively lower spending multipliers. On the other hand, using more variables for estimation and employing identification schemes that follow the Blanchard-Perotti or sign restriction approach deliver results with relatively higher values of spending multipliers. For tax multipliers, Blanchard-Perotti identification delivers a lower magnitude of estimates as compared to other specifications. We also find evidence corroborating a conclusion in Ramey (2019) that the specific definition of the multiplier used may lead to significantly different estimates.

The paper is organized as follows. Section 2 briefly presents the methodological setting used to estimate fiscal multipliers, based on SVAR and structural FAVAR models. Section 3 describes the different specification designs assessed for the estimation of fiscal multipliers in Austria. Section 4 presents the results of the analysis in detail and section 5 concludes.

2 Estimating Fiscal Multipliers: SVAR and structural FAVAR models

We can nest the set of models used to estimate fiscal multipliers in the stacked form of a dynamic factor model, following Stock and Watson (2016). A set of q dynamic factors are stacked to yield r static factors in the vector F_t and, abstracting from further deterministic terms (all our models contain a linear time trend), a FAVAR structure is be given by

$$\begin{pmatrix} \underset{n\times 1}{Y_t} \\ \\ \underset{m\times 1}{X_t} \end{pmatrix} = \begin{pmatrix} \underset{n\times n}{\mathbf{I}} & \underset{n\times r}{\mathbf{0}} \\ \\ \underset{m\times n}{\mathbf{\Lambda}^Y} & \underset{m\times r}{\mathbf{\Lambda}^F} \end{pmatrix} \begin{pmatrix} \underset{n\times 1}{\tilde{F}_t} \\ \\ \underset{r\times 1}{F_t} \end{pmatrix} + \begin{pmatrix} \underset{n\times 1}{\mathbf{0}} \\ \\ \underset{m\times 1}{e_t} \end{pmatrix} \tag{1}$$

$$\underset{(n+r)\times(n+r)}{\mathbf{\Phi}(L)} \begin{pmatrix} \underset{n\times 1}{\tilde{F}_t} \\ \\ \underset{r\times 1}{F_t} \end{pmatrix} = \begin{pmatrix} \underset{(n+q)\times(n+q)}{\mathbf{I}} \\ \\ \underset{(r-q)\times(n+q)}{\mathbf{0}} \end{pmatrix} \underset{(n+q)\times 1}{\eta_t} \tag{2}$$

$$\underset{(n+q)\times(n+q)}{\mathbf{A}} \underset{(n+q)\times 1}{\eta_t} = \underset{(n+q)\times(n+q)}{\mathbf{B}} \underset{(n+q)\times 1}{\varepsilon_t} \tag{3}$$

where equation (1) is the measurement equation, equation (2) is the transition equation, and equation (3) is the identification equation, while the (matrix) lag polynomial $\mathbf{\Phi}(L)$ is given by $\mathbf{\Phi}(L) = \mathbf{I} - \mathbf{\Phi}_1 L -$

$\cdots - \mathbf{\Phi}_p L^p$ for matrices $\mathbf{\Phi}_l$, $l = 1, \ldots, p$. The variables in Y_t (output, fiscal variables and other covariates) are assumed to be measured without error by the observed factors $\tilde{F}_t$. X_t contains m observed time series (not contained in Y_t) summarizing information about other macroeconomic and financial phenomena, as well as variables related to labour markets, production and sectoral developments. Variables in X_t are assumed to depend on observed factors $\tilde{F}_t$, unobserved factors F_t and an idiosyncratic component e_t, with matrix $\mathbf{\Lambda}^F$ comprising the corresponding factor loadings. Equation (3) specifies the relationship between reduced-form (η_t) and structural shocks (ε_t). If the number of unobserved factors r is set to zero, the model collapses to a standard SVAR model which can be utilized to implement the methods in Blanchard and Perotti (2002) or Perotti (2004) for structural shock identification. The unobserved factors of the model (F_t) are estimated as principal components and the identification of the model is reached once matrices $\mathbf{A}$ and $\mathbf{B}$ are chosen (see Stock and Watson, 2016).

Various identification methods can be used to retrieve the structural shocks in ε_t. The method pioneered by Blanchard and Perotti (2002) relies on exact restrictions imposed on the error terms of a VAR model which includes GDP, government expenditure and taxes through an identification scheme based on lags in the implementation of fiscal policy. More modern methods (Rubio-Ramírez et al., 2010) use sign restrictions that constrain the direction of the response of variables to particular shocks. Once the structural shocks have been identified, government spending and tax multipliers can be computed. In line with recent literature (e.g. Caggiano et al., 2015; Gechert and Rannenberg, 2014; Ilzetzki et al., 2013; Mountford and Uhlig, 2009), we report present-value (or discounted cumulative) multipliers at lag T,

$$\text{present-value spending multiplier} = \frac{\sum_{t=0}^{T}(1+i)^{-t} y_t}{\sum_{t=0}^{T}(1+i)^{-t} g_t} \frac{1}{g/y}, \tag{4}$$

where y_t is the response of output at time t (in logs), g_t denotes the response of government expenditures at time t (in logs) and g/y is the average share of government expenditures in GDP over the sample. The multiplier is discounted with the interest rate i, which is set to four percent *per annum*.[4] In the context of data at quarterly frequency, we report discounted cumulative multipliers for $T = 4$. The tax multiplier is calculated analogously, after substituting government expenditures in equation (4) with taxes.

If we concentrate on the non-cumulative reaction of GDP, such effects can be summarized using the so-called peak multipliers (see e.g. Blanchard and Perotti, 2002; Caggiano et al., 2015; Fragetta and Gasteiger, 2014; Ramey, 2011b),

$$\text{peak spending multiplier} = \frac{\max_{t=0,\ldots,H}\{y_t\}}{\max_{t=0,\ldots,H}\{g_t\}} \frac{1}{g/y}. \tag{5}$$

In order to account for the business cycle nature of the multipliers (and the known unreliability of results for longer horizons in these specifications), we restrict the horizon to a maximum of two years and set $H = 8$.

3 Model Specifications and Data

Specification choices

As reported in Čapek and Crespo Cuaresma (2020), in the context of estimating multipliers using SVAR specifications, seemingly harmless modelling choices may have a significant effect on the size and precision of fiscal multiplier estimates. In addition to the structural shock identification strategy, these modelling choices include the definition of spending and taxes, the national accounts system employed, the

[4]The discounting does not tend to play a major role for moderate interest rates, while it becomes more important in environments of high interest rates, such as emerging economies. The selection of a four percent interest rate corresponds to a commonly used discount factor of 0.99 per period.

use of particular interest rates or inflation measures in the model, or whether data are smoothed prior to estimation. Using a sample of European countries, Čapek and Crespo Cuaresma (2020) show that the cumulative effects of such apparently innocuous methodological choices can lead to large changes in the estimates of spending and tax multipliers. We explicitly integrate such uncertainty into our estimates for Austria, entertaining the large number of models which can be obtained by combining such possible methodological choices.

Table 1: Modelling choices for the estimation of fiscal multipliers

Dimension	Variants considered
Government data composition	Seven variants, see Table 2; ESA2010 codes and time series in the Appendix A
Deflating index	GDP deflator and HICP (not lagged and lagged by 4 quarters)
Model	VAR and FAVAR models with 3–5 vars. (factors ordered first or last)
Identification strategy	Cholesky ordering (only for spending multipliers), Blanchard-Perotti, sign restrictions
Number of factors	1–2 (FAVARs only)
Lags	1–4 lags

Table 1 lists all the methodological choices considered to construct models aimed at estimating fiscal multipliers for Austria. The set of possible variants is obtained by combining choices relating to (i) the data employed, (ii) the model used, and (iii) the particular specification within the model class. As for the data choices, these mainly concern the composition of government spending and revenues, but can also differ in the choice of the price index used to deflate nominal variables (CPI versus GDP deflator). Since a large part of government spending in Austria is linked to the lagged CPI (e.g. pension payments), we additionally consider lagged CPI (four-quarters lag) as a deflator in our analysis. The basic modelling choices in terms of specification structure are related to (a) the use of a simple VAR model versus employing a specification that incorporates unobserved factors, i.e., a FAVAR model, (b) the selection of variables in the (FA)VAR model, and (c) the choice of the identification strategy. Given a model specification, the technical choice relates to the number of lags in the (FA)VAR equation. For each model specification, we bootstrap 4000 multipliers and use the median as our point estimate.[5] The main analysis includes 1175 different specifications that can be obtained by combining all sensible choices, each yielding a (peak and present-value) spending median multiplier. For the estimation of tax multipliers, Cholesky identification is discarded, since it always results in zero impact multiplier, and thus 587 different specifications are used in the analysis.

Table 2 presents the different compositions of government spending and revenues used to obtain fiscal multipliers. Each choice consists of a specific composition of the government spending and government taxes aggregate. The *Baseline* setting ("Core/Tax Tiny") employs a simple composition which contains just three components of spending (compensation of employees, intermediate consumption and gross capital formation) and two components of revenues (taxes on production, imports, income and wealth).[6] The following two combinations adjust the baseline setting by including also social contributions and subsidies as part of the fiscal aggregate (as in Crespo Cuaresma et al., 2011, for instance). To reflect the

[5] In sign restriction identification schemes, the 4000 solutions are the actual draws. Other identification approaches rely on bootstrapping to compute the 4000 draws. The bootstrap employed builds on resampling raw residuals (with replacement) and subsequent refitting of the model. Portmanteau tests for residual autocorrelation suggest that around two thirds of the estimated models do not exhibit significant residual autocorrelation at any sensible lag.

[6] See Appendix A for the ESA2010 codes corresponding to each component.

Table 2: Government spending and revenues composition

Tag	Gov't spending composition	Gov't revenues composition
core/tax tiny (*Baseline*)	Compensation of employees, intermediate consumption, and gross capital formation	Taxes on production, imports, income, and wealth
core/tax small net soc.t.		*Baseline* adjusted for actual social contributions
core/net tax small		*Baseline* adjusted for social contributions and subsidies
corefix+soc.t.kind/tax mid	*Baseline* (gross fixed capital) + transfers in kind	*Baseline* + household social contributions
corefix+soc.t.kind/net tax mid		*Baseline* + household social contributions adjusted for subsidies
corefix+soc.t.kind/net tax large		*Baseline* + household social contributions adjusted for subsidies and transfers
core/net tax all	*Baseline* + acquisitions of assets	*Baseline* + household social contributions adjusted for subsidies and transfers (incl. capital transfers)

Note: We use seven sets of compositions of government spending and revenues. Starting from "core/tax tiny", which is the *Baseline* composition (shaded in grey), the other composition sets add extra spending and/or revenue items. These are ordered from narrower to broader sets, comprising different spending and/or revenue items. The corresponding tag is constructed with abbreviations of spending composition separated from abbreviations of revenue composition using a slash "/" . The term "core" refers to the *Baseline* spending composition, "corefix" highlights the use of fixed capital formation. The abbreviations for taxes range from "tiny", with only several items, to "all", with a broad selection of revenue items. For specific ESA codes for each composition set, see Appendix A.

particularities of the Austrian economy, we also use other composition choices reflecting the importance of transfers in kind, household social contributions, subsidies, and transfers for the country. Deviating from the existing literature, so as to cover the specific case of Austria, we introduce three new data compositions, whose tag starts with "corefix" in Table 2. The inclusion of social transfers in kind in this government spending aggregate accounts for the fact that social transfers in kind amount to more than 8% of overall government spending in the country. Due to their use to finance large parts of the healthcare and social protection system, changes in the provision of social transfers in kind create important economic spillovers (for example by substituting private expenditure for old-age and long-term care) that should be considered in the analysis. The particular revenue compositions used reflect the importance of household social contributions, subsidies and transfers for overall disposable household income in Austria. Following Muir and Weber (2013), we also entertain models based on government spending aggregates that contain acquisitions of assets and a battery of adjustments regarding social contributions, subsidies, and transfers (including capital transfers).

The Cholesky identification strategy identifies a fiscal shock using a particular ordering based on the contemporaneous responses across shocks. The first and most exogenous variable is assumed to be government spending, followed by GDP, inflation (in VAR models with four and five variables), taxes, and the interest rate (in VAR models with five variables only). Since GDP is ordered before taxes, the impact tax multiplier is zero by construction, so we use the identification strategy based on the Cholesky decomposition exclusively for spending multipliers. The Blanchard-Perotti identification scheme follows Blanchard and Perotti (2002) for VAR models with three variables and Perotti (2004) for specifications with more variables. The output and price elasticities required to carry out the identification procedure

Table 3: Output and price elasticities of spending and tax composition

	Output elasticity	**Price elasticity**
Spending compositions		
core	0	-0.542
corefix+soc.t.kind	0	-0.542
Revenue compositions		
tax tiny	0.832	-0.005
tax small net soc.t.	2.375	1.923
net tax small	2.725	2.355
tax mid	0.721	0.064
net tax mid	1.579	1.127
net tax large	1.750	1.344
net tax all	2.205	1.856

Note: Elasticities are calculated using the fiscal forecasting model by Austrian Fiscal Advisory Council (2014). Compositions in Table 2. For detailed ESA codes for each composition, see Appendix A.

in Blanchard and Perotti (2002) are computed for every net tax and spending composition specification using the fiscal forecasting model of the Austrian Fiscal Advisory Council (see Table 3). The model partitions government revenue and expenditure into around 120 budget items that are corrected for structural breaks and then projected individually (see Austrian Fiscal Advisory Council, 2014). We shock the model in the year 2019 using a 1 % increase in real GDP to obtain estimates of output elasticities and a 1% increase in the price level for price elasticities. The real GDP shock is decomposed into its subcomponents (tax bases) so as to represent an average historical shock in the country. The reaction of the individual budget items is then aggregated to the corresponding compositions (see Table 2 and Appendix A) using the average weights of these items during the period 2000–2019. As a last step, for the case of the output elasticity, we deflate the nominal budget reactions using the rise in inflation induced by the GDP shock. For the price elasticity estimates, we substract one (the size of the original shock) to the percentage reaction in the price level.

Our implementation of sign restrictions identifies three shocks: the business cycle shock is identified by requiring the impulse responses of output and taxes to be positive for at least the four quarters following the shock. The tax shock is identified by a positive response of taxes for at least the four quarters following the shock (and the shock is required not to meet the identifying restrictions for the business cycle shock). For the identification of a government spending shock, the responses of government spending need to be positive for at least the four quarters following the shock (and the shock is required not to meet the identifying restrictions for the business cycle shock).

The identification strategies mentioned above are unable to explicitly address the issue of fiscal foresight. If a fiscal policy change is known before its (official) implementation and economic agents react accordingly, the reaction in the real economy may be apparent earlier. This timing mismatch is known as fiscal foresight and essentially amounts to a limited information problem (Fragetta and Gasteiger, 2014). Forni and Gambetti (2014) suggest to remedy the problem by extending the VAR model with principal components (as estimates of unobservable factors), which are calculated from a broad range of additional time series containing relevant information. We add one or two principal components to the VAR specification with three variables, making the model a proper FAVAR specification. We estimate the principal components with the aid of 26 additional time series that relate to macroeconomic dynamics, financial markets, and the labour market.[7]

Additionally, we add dummy variables to the baseline specification so as to reflect the impact and consequences of the Great Recession on the economic variables used in the models. We add a dummy taking value one for the period 2008Q4–2009Q2 and a step dummy starting from 2009Q1 until the end of the sample.[8]

Data

The main source of data is Eurostat, while some financial variables used for the estimation of the unobserved factors are sourced from the European Central Bank. We use time series of the corresponding disaggregated components of government spending and tax revenues to construct the various fiscal variables required to estimate our models (see Appendix). For extended versions of the VAR model with four and five variables, we also use inflation and the interest rate. The data cover the period span from the first quarter of 2001 to the fourth quarter of 2018, yielding 72 quarterly observations. If available, seasonally adjusted variables are employed. If seasonally adjusted data are unavailable, we use the X-13 toolbox to remove seasonal patters from those variables that contain a seasonal component.[9] All the time series for spending and tax categories, as well as GDP, are obtained from the source in nominal terms and subsequently deflated using the corresponding deflator (see Table 1).[10] The corresponding fiscal variables and GDP enter the (FA)VAR models in logs, while inflation and the interest rate are added to the VAR without further transformation (i.e., in percentage points). The methodological framework employed for the identification of fiscal shocks, which corresponds to the standard specifications used in the modern literature on fiscal multipliers, implies that the variables in the VAR model are assumed to be stationary or trend-stationary (i.e., stationary around deterministic linear trend). All time series used to estimate the factors are transformed to reach stationarity prior to obtaining estimates of the factors.[11]

4 Fiscal Multipliers in Austria: The Role of Forecasting Performance and Specification Choices

The estimated fiscal multipliers for Austria are summarized in Table 4. We make use of out-of-sample predictive accuracy as a validation device of the models used in our exercise. We utilize the last four observations of our GDP series as an out-of-sample period and compute the mean absolute error (MAE) of one-step-ahead GDP predictions for all specifications used to obtain multiplier estimates, after estimating the models using a sample that excludes the out-of-sample observations. The results of this forecasting exercise allow us to refine the inference on Austrian expenditure and tax multipliers by concentrating on the estimates corresponding to the set of models with best predictive ability.

The mean present-value spending multiplier over all models is 0.94 and reduces to 0.87 if we focus on the group of best models according to predictive ability (specifications corresponding to the 40% best models in terms of MAE). Generally, peak spending multipliers are larger than present-value spending

[7] See the Appendix A for the list of the time series used to estimate the factors.

[8] See the Appendix for the results without crisis dummies and with different dummification strategies for the Great Recession period.

[9] We employ the X-13 Toolbox for Seasonal Filtering by Yvan Lengwiler in *Matlab File Exchange*. The default setting lets TRAMO select additive or multiplicative filtering and then decomposes the series into a trend, cycle and seasonal component using X-11, with additive outliers allowed, as well as trading day dummies.

[10] Revenue categories are not available in real terms. In order to investigate the effects of deflating with different price indices while keeping consistency, we choose to source all time series in nominal terms and deflate them with the same deflator.

[11] See the Appendix A for the transformations carried out in each of the time series used to estimate the factors.

Table 4: Fiscal multiplier estimates

Multiplier type	min	16-th p.	mean	median	84-th. p	max
Spending multiplier (present value)	-1.81	0.63	0.94	0.99	1.22	2.43
— best 40%	-1.38	0.52	0.87	0.89	1.21	2.15
Tax multiplier (present value)	-2.30	-1.28	-0.76	-0.82	-0.23	1.92
— best 40%	-2.30	-1.23	-0.76	-0.84	-0.24	1.11
Spending multiplier (peak)	0.25	0.87	1.08	1.06	1.30	2.22
— best 40%	0.25	0.83	1.07	1.03	1.34	1.99
Tax multiplier (peak)	-2.17	-0.90	-0.58	-0.58	-0.19	-0.02
— best 40%	-2.17	-0.90	-0.59	-0.57	-0.22	-0.05

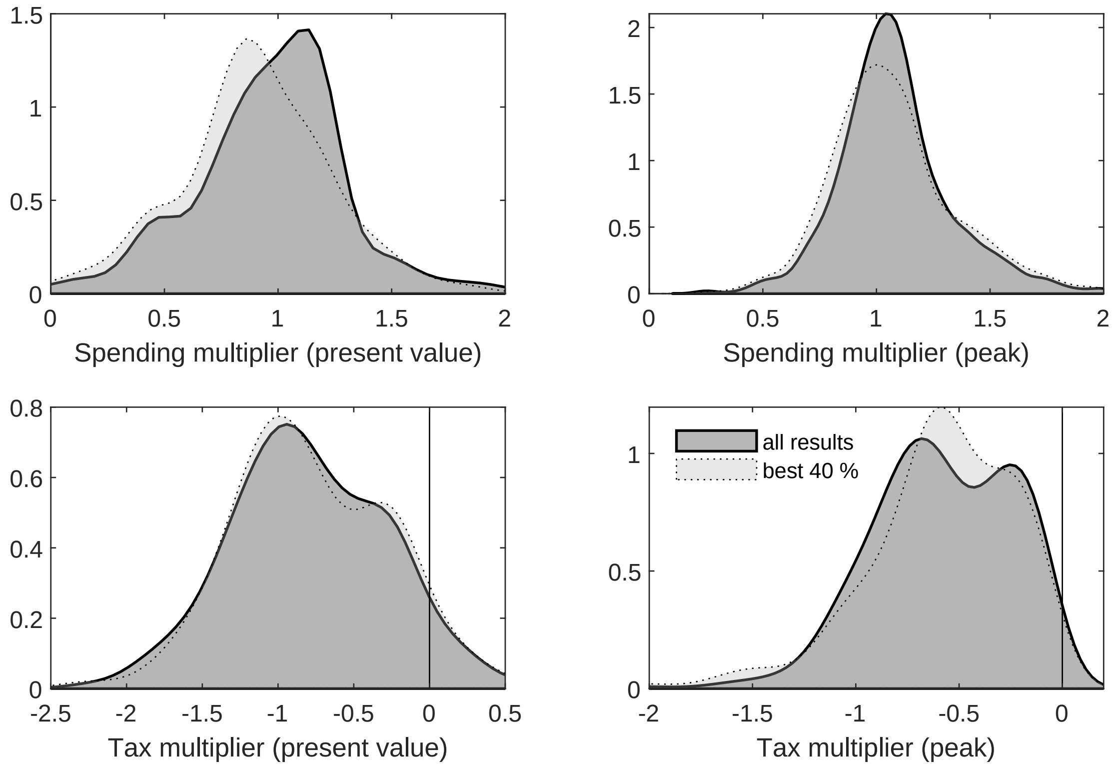
Note: Descriptive statistics of the full set of results based on 1175 spending and 587 tax median multipliers estimates. The group based on the 40% best-forecasting models consists of 465 spending and 236 tax multipliers. See Figure 1 for kernel densities.

multipliers. The mean peak spending multiplier is 1.08 over all models and 1.07 in the group of models with best predictive power. As for the tax multipliers, the value of present-value tax multiplier is -0.76 across all models and also concentrating on the models with particularly good forecasting ability. The mean peak tax multiplier is -0.58 for the whole set of specifications entertained and -0.59 once we concentrate on the models with best forecasting performance. Our findings support the hypothesis that spending multipliers are larger (in absolute value) than tax multipliers. The smoothed densities of the estimated multipliers are presented in Figure 1 for the full sample of fiscal multiplier estimates, as well as for the top 40% models in terms of out-of-sample predictive ability.

With the exception of present-value spending multiplier, comparing the means of the multiplier distributions across all models and focusing on the models with best predictive ability delivers very similar results. However, within certain types of specifications, sizeable differences can be found when zooming into the group of models which have a higher predictive power. The most pronounced differences between variants of the same type of specification are depicted in Figure 2, which shows the empirical densities of present value spending multiplier for the full sample and for subsets based on predictive ability (best 20%, 40%, 60%, and 80% models), split in four panels depending on the number of lags of the (FA)VAR. The (FA)VAR models with one or two lags tend to higher values of the spending multiplier. The first two panels of Figure 2 demonstrate that the modes of the distributions are almost 1.2. In contrast, models with three or four lags results in a distribution of spending multipliers with a mode around 1. However, concentrating on the best specifications according to predictive ability, the distribution of multipliers in the models with one or two lags is concentrated around significantly lower values. The mode of the distribution for models with one lag (first panel) is around 0.9, whereas the mode of the distribution for models with two lags is below 0.8. These findings suggest that although some specifications tend to deliver values of spending multipliers larger than 1, many of these disappear once we focus on models which predict well. The patterns observed in first two panels of Figure 2 help explain the differences between distributions in the first panel of Figure 1.

Table 5 summarizes the share of models with best forecasting performance in the full set of specifications by variable definition. The data composition which tends to improve forecasting performance for GDP data is the composition tagged ”corefix+soc.t.kind/tax mid”, which covers 16.8% of the models in the top 40% specifications by predictive ability. Adjusting the revenue part of this composition by subsidies, social benefits other than social transfers in kind, and other current transfers, is the composition (tagged ”corefix+soc.t.kind/net tax large”) that leads to the relatively worst predictive ability, covering only 9.7% of the models among the top 40%. However, as the results for the last composition in the table (”core/net tax all”) show, broader compositions do not necessarily lead to worse predictive ability.

Figure 1: Fiscal multiplier estimates: kernel densities



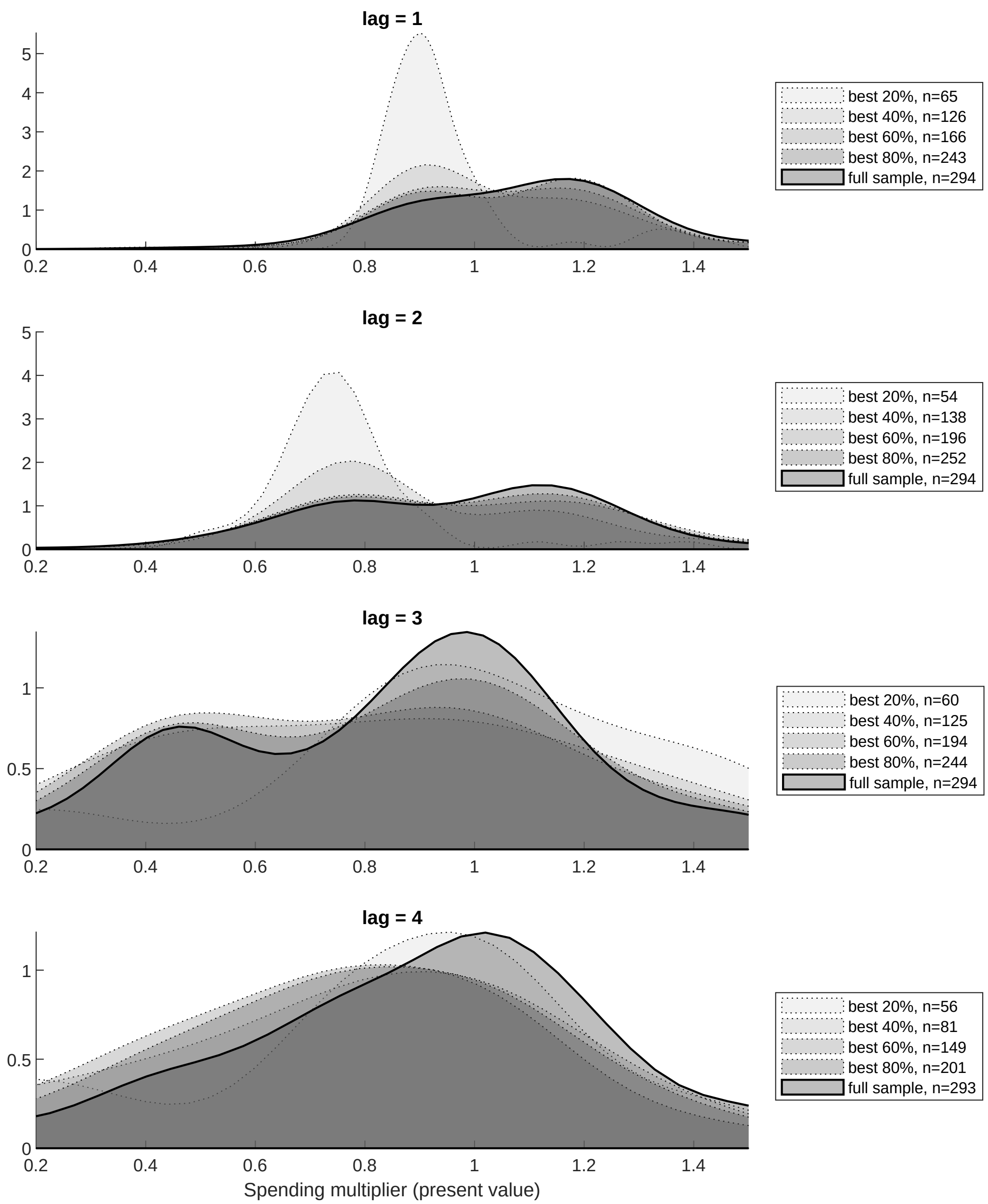
Note: The dark density corresponds to the full set of results, the light density refers to the top 40% best models in terms of predictive ability. See also notes to Table 4.

Data compositions which lead to models featuring particularly good predictive ability are the *Baseline* (”core/tax tiny”), the “corefix+soc.t.kind/tax mid”, and the “core/tax small net soc.t.” variants (see Table 2 for a description of data composition and the Appendix A for ESA codes).

Figure 3 shows multiplier estimates across different sets of government spending and revenue compositions. While most of the empirical densities for spending multipliers are relatively similar, tax multipliers seem to be more sensitive to varying composition of government spending and taxes. For the case of the spending multiplier (see top panels of Figure 3), models using the composition that includes acquisition of assets (“core/net tax all”, inspired by Muir and Weber, 2013) lead to a distribution of multiplier estimates that has a similar mean as that of other data composition choices, but is more spread around the mode. This indicates that adding acquisition of assets as part of spending composition leads to a less precise point estimate of the spending multiplier across models.

Our results further highlight that for tax multipliers, the choice of a particular group of fiscal variables in the model may have a larger effect on multiplier estimates than in the case of spending multipliers. The empirical distributions of some multiplier estimates tend to be rather flat for certain cases, while a composition set including capital transfers “core/net tax all”, delivers more precise peak tax multiplier estimates (albeit relatively low in magnitude). The lower magnitude of tax multiplier is also due to a potentially misleading identification of exogenous shocks, especially for a revenue variable (net taxes) that

Figure 2: Spending multiplier densities based on forecasting performance, split over lags of the (FA)VAR



Note: Kernel densities estimated on subsets of multipliers according to the number of lags in the (FA)VAR equation. The darkest density corresponds to the full set of results, the lighter ones correspond to subsets of models by predictive ability (best 20%, best 40%, best 60%, and best 80%).

Table 5: Data composition and forecasting performance

	Count		Percentage	
	total	**best 40%**	**total**	**best 40%**
core/tax tiny	168	77	14.3	16.6
core/tax small net soc.t.	168	76	14.3	16.3
core/net tax small	168	57	14.3	12.3
corefix+soc.t.kind/tax mid	168	78	14.3	16.8
corefix+soc.t.kind/net tax mid	168	62	14.3	13.2
corefix+soc.t.kind/net tax large	168	45	14.3	9.7
core/net tax all	167	70	14.2	15.1
total	1175	465	100%	100%

Note: **Count** contains numbers of existing specifications across different spending/tax compositions. **Percentage/best 40%** illustrates the relative representation of various spending/tax composition sets among the best 40% specifications. For a graphical representation of all results based on selected compositions, see Figure 3.

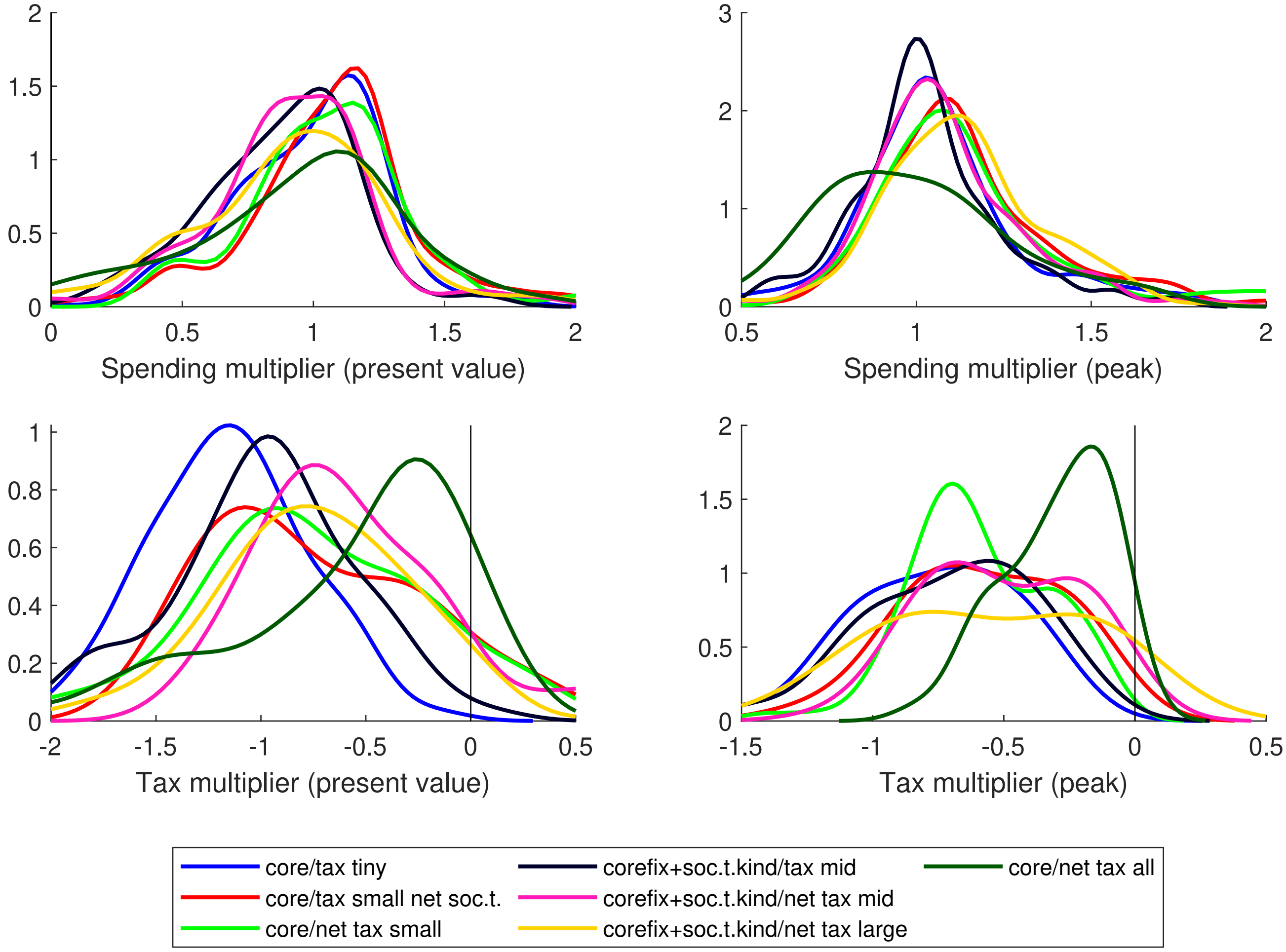
includes capital transfers. In recent years, virtually all of the variation in capital transfers in Austria has been due to sizable banking support programs, which arguably had only mild effects on GDP. This leads to more precise but lower magnitudes of (net) tax multipliers once capital transfers are included, however providing little information on how more common types of taxes affect output. While the "core/net tax all" composition delivers the lowest average magnitude of the estimate of the present value tax multiplier, the Baseline composition "core/tax tiny" delivers the highest one. More inclusive specifications ("tax small net soc.t." and "net tax small") tend to deliver estimates closer to zero, which are estimated with less precision.

Turning to the effects of using different econometric specifications, identification strategies, and number of variables (see Figure 4), on average, models with three variables and a shock identification design based on the Cholesky decomposition tend to result in lower spending multiplier estimates compared to models which employ more variables and different identification schemes. Whereas VAR models with 3 variables or models estimated with Cholesky ordering lead to present value median spending multipliers centered around 0.8, following more modern approaches yield spending multiplier estimates with a median above unity. However, sign restriction and Blanchard-Perotti identification strategies tend to have higher variance around the mean and deliver therefore less precise estimates. For tax multipliers, which do not include estimates based on Cholesky identification, the patterns indicate that those based on the Blanchard-Perotti identification scheme tend to be smaller in magnitude. In the case of present value tax multipliers, the estimates calculated using Blanchard-Perotti identification are less precise and have a higher frequency of outlying values.

5 Conclusions

This paper estimates fiscal multipliers for Austria, a stereotypical advanced small open economy, with a focus on the dimension of model uncertainty that emanates from the choice of a particular econometric model to obtain point estimates of the reaction of GDP to shocks in fiscal variables. We present a comprehensive framework which allows to assess the effects of different multiplier definitions and choices related to the data, the model employed, and further technical choices associated with the specification of the model exert on fiscal multiplier estimates.

Figure 3: Multiplier densities and data composition, based on all results

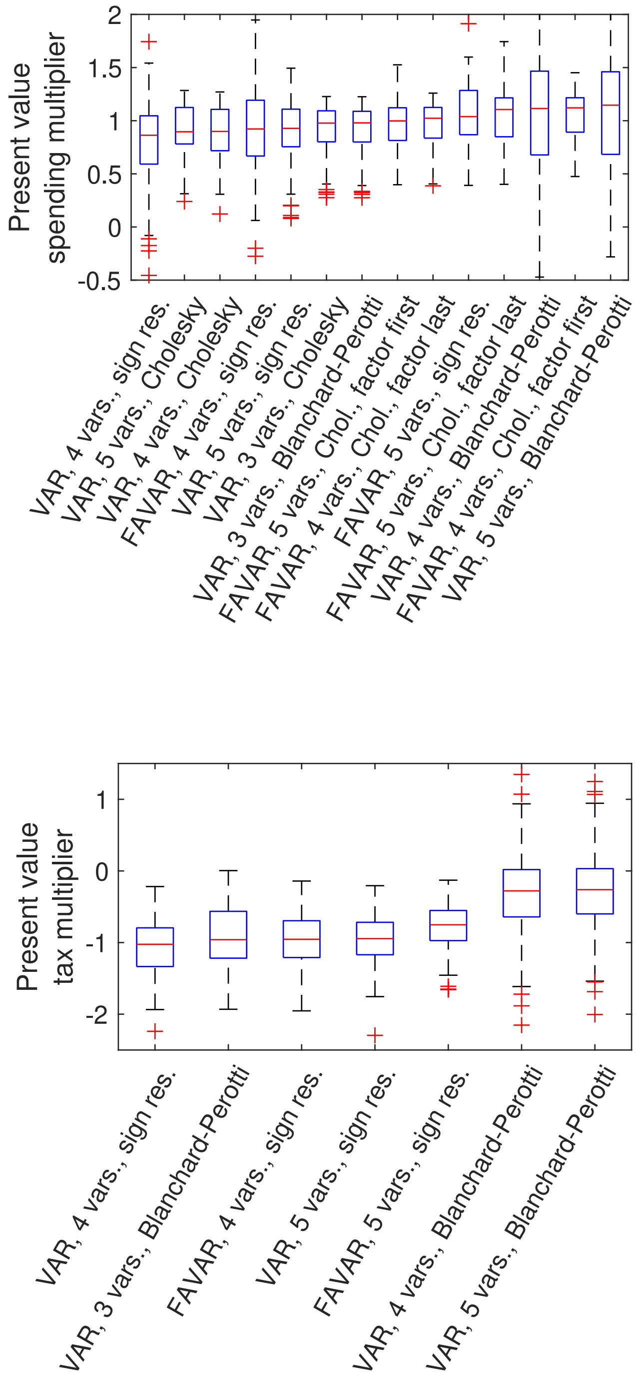


Note: For details on data compositions, see Table 2.

The mean present-value spending multiplier over all models entertained is 0.94 and reduces to 0.87 once we focus on the best models according to out-of-sample predictive ability. Generally, estimates of the peak spending multiplier for Austria tend to be larger than present-value spending multipliers. The mean peak spending multiplier is 1.08 and 1.07 if calculated on the basis of the group of models with best predictive performance. As for the tax multipliers, the mean of the present-value tax multiplier is -0.76, with no effect of selecting models with best predictive ability. The mean peak tax multiplier is -0.58 for all specifications used and -0.59 once we concentrate on the models with the best forecast performance.

Splitting our results based on the number of lags in the (FA)VAR model, our findings suggest that even though some specifications tend to lead to values of spending multiplier larger than unity, many of these are discarded once we focus on models which predict well. Comparable results are found when we focus on forecasting performance and split models over different compositional definitions of government expenditures and taxes. The particular composition that delivers the highest percentage of models that predict well uses compensation of employees, intermediate consumption, gross capital formation, and transfers in kind as part of government expenditures and taxes on production, imports, income, and wealth, and household social contributions.

Figure 4: Fiscal multipliers by model and identification strategy types



Note: Boxplots are sorted by the median multiplier, the central (red) mark of the boxplot. The bottom and top edges of the box indicate the 25th and 75th percentiles.

On average, multipliers obtained from models that require few variables and use Cholesky identification for the structural shocks tend to result in lower estimates of the spending multiplier. On the other hand, using more variables for estimation and employing identification schemes that follow the Blanchard-Perotti approach or sign restrictions deliver higher estimates of spending multipliers. For tax multipliers, Blanchard-Perotti identification delivers estimates of lower magnitude as compared to other specifications.

In line with conclusions in Ramey (2019), we find that the specific method used to obtain multipliers can make a big difference in terms of inference. Given the scarce evidence on multipliers in developed small open economies, the results we present for Austria have a value of their own for policymakers and fiscal authorities.

References

Afonso A and Sousa RM (2011) The macroeconomic effects of fiscal policy in Portugal: a Bayesian SVAR analysis. *Portuguese Economic Journal* 10(1), 61–82

Auerbach AJ and Gorodnichenko Y (2012) Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy* 4(2), 1–27

Austrian Fiscal Advisory Council (2014) Fiscal Forecasts of the Austrian Fiscal Advisory Council: Forecasting Principles and Methods

Barrell R, Holland D and Hurst I (2012) *Fiscal Consolidation: Part 2. Fiscal Multipliers and Fiscal Consolidations*. Organisation for Economic Co-Operation and Development (OECD), OECD working papers edition

Beetsma R and Giuliodori M (2011) The Effects of Government Purchases Shocks: Review and Estimates for the EU. *The Economic Journal* 121(550), F4–F32

Benetrix AS and Lane PR (2009) The impact of fiscal shocks on the Irish economy. *The Economic and Social Review* 40(4), 407

Blanchard O and Perotti R (2002) An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *The Quarterly Journal of Economics* 117(4), 1329–1368

Blanchard OJ and Leigh D (2013) Growth forecast errors and fiscal multipliers. *The American Economic Review* 103(3), 117–120

Breuss F, Kaniovski S and Schratzenstaller M (2009) Macro-economic effects of the fiscal stimulus measures in Austria. *Austrian Economic Quarterly* 14(4), 205–216

Caggiano G, Castelnuovo E, Colombo V and Nodari G (2015) Estimating Fiscal Multipliers: News From A Non-linear World. *The Economic Journal* 125(584), 746–776

Caldara D and Kamps C (2017) The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers. *The Review of Economic Studies* 84(3), 1015–1040

Čapek J and Crespo Cuaresma J (2020) We just estimated twenty million fiscal multipliers. *Oxford Bulletin of Economics and Statistics* 82(3), 483–502

Crespo Cuaresma J, Eller M and Mehrotra A (2011) The Economic transmission of fiscal policy shocks from Western to Eastern Europe. *BOFIT Discussion Papers* 2011(12), 3

Forni M and Gambetti L (2014) Sufficient information in structural VARs. *Journal of Monetary Economics* 66, 124–136

Fragetta M and Gasteiger E (2014) Fiscal Foresight, Limited Information and the Effects of Government Spending Shocks. *Oxford Bulletin of Economics and Statistics* 76(5), 667–692

Gechert S (2015) What fiscal policy is most effective? A meta-regression analysis. *Oxford Economic Papers* 67(3), 553–580

Gechert S and Rannenberg A (2014) Are Fiscal Multipliers Regime-Dependent? A Meta Regression Analysis. *IMK Working Paper* (139)

Hebous S (2011) The effects of discretionary fiscal policy on macroeconomic aggregates: a reappraisal. *Journal of Economic Surveys* 25(4), 674–707

Ilzetzki E, Mendoza EG and Végh CA (2013) How big (small?) are fiscal multipliers? *Journal of Monetary Economics* 60(2), 239–254

Jemec N, Kastelec A and Delakorda A (2011) *How Do Fiscal Shocks Affect the Macroeconomic Dynamics in the Slovenian Economy*. Prikazi in analize. Banka Slovenije

Karras G (2014) Fiscal policy spillovers through trade openness. *Journal of Economic Integration* 29(3), 563–581

Kilponen J, Pisani M, Schmidt S, Corbo V, Hledik T, Hollmayr J, Hurtado S, Júlio P, Kulikov D, Lemoine M et al. (2015) Comparing fiscal multipliers across models and countries in Europe. Technical report

Koch S, Monárová Z and Reiter M (2019) Macro-economic effects of the fiscal stimulus measures in

Austria. *Mimeo, Institute for Advanced Studies, Vienna*

Mineshima A, Poplawski-Ribeiro M and Weber A (2014) Size of fiscal multipliers. In *Post-Crisis Fiscal Policy*. MIT Press, 315–372

Mountford A and Uhlig H (2009) What are the effects of fiscal policy shocks? *Journal of Applied Econometrics* 24(6), 960–992

Muir D and Weber A (2013) Fiscal Multipliers in Bulgaria: Low But Still Relevant. Technical report, IMF Working Paper No. 13/49

Perotti R (2004) Estimating the effects of fiscal policy in OECD countries. Technical Report December, IGIER Working Paper No. 276

Ramey VA (2011a) Can Government Purchases Stimulate the Economy? *Journal of Economic Literature* 49(3), 673–685

Ramey VA (2011b) Identifying government spending shocks: It's all in the timing. *The Quarterly Journal of Economics* 126(1), 1–50

Ramey VA (2016) Macroeconomic shocks and their propagation. In *Handbook of macroeconomics*, volume 2. Elsevier, 71–162

Ramey VA (2019) Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? *Journal of Economic Perspectives* 33(2), 89–114

Ravn SH and Spange M (2012) The effects of fiscal policy in a small open economy with a fixed exchange rate: The case of Denmark. Technical report, Danmarks Nationalbank Working Papers

Rubio-Ramírez JF, Waggoner DF and Zha T (2010) Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *The Review of Economic Studies* 77(2), 665–696

Schuster P (2019) On fiscal multipliers in New Keynesian small open economy models. *Mimeo*

Stock JH and Watson MW (2016) Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In *Handbook of macroeconomics*, volume 2. Elsevier, 415–525