Identification of fiscal SVAR-IVs in small open economies*

Henri Keränen† Sakari Lähdemäki‡

August 24, 2022

-Preliminary and incomplete draft-

Abstract

We propose a novel instrumental variable strategy to identify fiscal shocks in small open economies. Under the assumptions that unexpected changes in trading partners correlate with output of an open economy (relevance) and unexpected fiscal shocks of a small economy are unrelated to its trading partners' forecast errors (exogeneity), we use forecast errors of trading partner economies to proxy unexpected shocks in domestic output. We show that this instrument is relevant and find suggestive evidence that it fulfills the exogeneity assumption better than instruments currently applied in literature. Using this IV strategy, we study the effects of fiscal policy in Canada and euro area small open economies.

Keywords: Fiscal policy, Fiscal multiplier, SVAR-IV, Small open economy

JEL Codes: E62, C26, C32, F41

---

*The authors would like to thank Ilkka Kiema, Samu Kärkkäinen and participants of the IAAE 2021 conference for their useful comments. Roni Mäkelä helped us with data during early parts of this project. Funding provided by the Palkansaajasäätiö is gratefully acknowledged. The authors have no conflicts of interest to declare.

†Economic Policy Council of Finland, VATT Institute for Economic Research; Email: henri.keranen@vatt.fi

‡Labour Institute for Economic Research LABORE; Email: sakari.lahdemaki@labore.fi

1 Introduction

Estimating the structural effects of policy shocks is a central but notoriously hard task in empirical macroeconometrics, given the difficulty of credibly identifying exogenous variation in the policy variable (Nakamura and Steinsson, 2018). At the same time, modern macroeconometrics increasingly uses external instruments to identify dynamic causal effects (Stock and Watson, 2018). Given a valid instrument, researchers can estimate structural relationships in the presence of endogenous variables, which is emblematic of macroeconomics.

In this paper, we identify structural vector autoregressions with external instruments (SVAR-IV) to study the dynamic effects of fiscal policy. Our main contribution is a novel instrument for aggregate output shocks that builds on the small open economy assumption that the policy shocks of a small country do not affect the aggregates of its often larger trading partners. More concretely, we propose to use professional forecast errors of trading partner economies as a proxy for output shocks. These errors are arguably unrelated to the government spending and revenue shocks of a small open economy (exogeneity) while at the same time able to explain output variations in the small domestic economy (relevance). The more open and smaller the economy is, the more confident one can be of these assumptions to hold.

The main concerns related to our instrument are the following. Firstly, there is the concern that the instrument is not relevant. However, the empirical part of this paper shows that our instrument is not weak as trading partner forecast errors are strongly correlated with the unexpected part of domestic output. The other concerns relate to the exogeneity of the instrument.[1] Since the exogeneity of our instrument is crucial for this paper, we address some of the main concerns here.

Firstly, to fix ideas, it is helpful to divide fiscal policy into two separate components: the systematic component and the exogenous policy interventions. The systematic part

[1]In the context of this paper, the exclusion restriction requires that the SVAR structural fiscal shocks are uncorrelated with the instrument (trading partner forecast errors).

reflects the reaction function of fiscal authorities or, in other words, the implicit fiscal rules that govern how fiscal policy endogenously reacts to other (non-policy) shocks in the model. According to Caldara and Kamps (2017), at a quarterly frequency, these reflect, for example, the automatic stabilizers that mainly work through output. The exogenous part, on the other hand, reflects the policy shocks. Following Caldara and Kamps (2017), we use our non-fiscal instrument to estimate the parameters of the tax and spending policy rules that capture the systematic component of fiscal policy.[2] The remaining unexplained part of fiscal policy is taken to be the structural fiscal shocks of the SVAR-model.

Because we consider a small open economy, it is reasonable to think that such an economy's exogenous fiscal policy decisions have no systematic effect on its trading partners. This is similar to the small open economy (SOE) VARs assumption that a small open economy does not affect the world demand (see, for example, Cushman and Zha, 1997).[3] However, it is arguably weaker than the canonical SOE assumption because we only need to assume that the unexpected fiscal policy shocks are uncorrelated with unexpected shocks of the foreign trading partners (foreign block). On the contrary, the SOE VARs often assume that the small open economy does not affect the foreign block at all. Furthermore, since we use (a weighted average of) trading partner forecast errors as a proxy, we need not model the complete VAR dynamics of a set of foreign economies to obtain plausibly exogenous variation that can be used in the structural identification.

However, one concern is that if a small open economy is closely connected to another small open economy, then the exogenous fiscal decisions might cause changes in the output of this trading partner. This could result in a correlation between the true fiscal policy shocks and the unexpected shocks of the trading partner. We argue that this concern should be considered on a case-by-case basis, and no universal truth on this matter should be claimed. In the case of Canada, the SOE assumption is generally adopted in the US-Canada context.

[2] This is in contrast to some of the recent fiscal SVAR literature that utilizes fiscal instruments instead. Fiscal instruments proxy structural policy shocks but must be uncorrelated with non-policy shocks.

[3] To our knowledge, this idea has not been applied in the context of small open economy fiscal policy analysis.

Arguably our instrument is even more likely to fulfill the exogeneity assumption. Considering the euro area small open economies, for example, one could be concerned by the possible spillover effects of Finnish fiscal policy on the Swedish economy. To address this specific concern, we have compared two different versions of our proposed trading partner instrument: a baseline version containing data on all available trading partners (incl. Sweden) and one which is constructed as a weighted average of only G7 country forecast errors. We find that these two versions of the instrument yield very similar results. This fact suggests that one could exclude the most suspicious countries in terms of exogeneity from the sample without a large trade-off in instrument relevance.[4]

It is also worth noting that we need not be concerned whether some of the variation in the instrument is explained by common shocks. Common shocks are acceptable as long as the common variation in the domestic economy and its trading partners is not due to fiscal policy shocks of the small domestic economy. In fact, from an econometricians' point of view, it would be helpful if one could observe two economies subjected to a set of common shocks but with different fiscal policies.

Finally, we also provide statistical evidence that supports the exogeneity assumption related to our instrument. Following Auerbach and Gorodnichenko (2012), we use forecast errors of government spending and investment to proxy the unexpected policy shocks. We then study the relationship between this proxy and our instrument and find suggestive evidence—which is independent of our SVAR application of the instrument—that the exogeneity assumption holds in our small open economy context.

We apply the proposed instrument to identify fiscal SVAR-models that rely on well-established identification schemes. In the classic Blanchard and Perotti (2002) identification, the part of fiscal variables that is not explained by the past lags of modeled variables or the automatic response to contemporaneous variation in output is assumed to be the structural shocks. These shocks are then effectively used as internal instruments in identifying the

[4] In this respect, it is helpful if foreign trade is relatively well diversified among many partner countries, as in the case of Finland.

contemporaneous effects of fiscal shocks on output.[5] Whereas Blanchard and Perotti (2002) rely on external coefficients that capture the systematic contemporaneous response of fiscal variables to output, Caldara and Kamps (2017) propose to estimate these elasticities with external instruments. We follow this latter strategy but utilize our proposed instrument instead.

Preceding SVAR-IV literature has considered the utilization adjusted total factor productivity (TFP) series of Fernald (2014), which for example, Caldara and Kamps (2017) utilize, as the instrument of choice for output shocks. Compared with this prevailing TFP based instrument, the instrument we propose is formed straight from observable data and does not require any strong assumptions on structural forms beyond the small open economy assumption.[6] Furthermore, the analysis in Angelini et al. (2020) raises some concerns that the utilization adjusted TFP series of Fernald (2014), if used as an instrument, does not necessarily fulfill the exogeneity assumption. Based on a proxy for fiscal policy shocks, we provide suggestive evidence that our instrument seems to fulfill the exogeneity assumption more reliably. Moreover, it has been shown by Kurmann and Sims (2021) that revisions made to the utilization adjusted TFP series of Fernald (2014) over time have been substantial.[7] Given these concerns, we see that the instrument we propose is a viable alternative to the TFP instrument in a world of scarce instruments.

We apply the new instrument to two different types of small open economies: Canada and

[5]Following the SVAR model of Blanchard and Perotti (2002), much of the macro-fiscal research has relied on using assumptions about timing restrictions in the real-world setting of fiscal policy to identify fiscal shocks. Another prominent identification strategy has been to use sign restrictions, as in Mountford and Uhlig (2009). Furthermore, another branch of the literature has utilized narratively identified plausibly exogenous policy changes or other shocks in the identification of the effects of fiscal policy. See, for example, Romer and Romer (2010) and Mertens and Ravn (2013) in the context of tax policy changes and Ramey (2011) and Ramey and Zubairy (2018) in which a military-news variable is constructed to identify plausibly exogenous variation in government spending. See Ramey (2019) for a more comprehensive summary of the macro-fiscal literature.

[6]Note also that our instrument is not a narrative instrument as, for example, in Romer and Romer (2010), Mertens and Ravn (2013), or Ramey and Zubairy (2018). That is, we do not choose which shocks are exogenous and which are not. We only assume that the forecasts we use are reasonable and that the resulting forecast errors can proxy unexpected shocks at the quarterly frequency.

[7]Note that the utilization adjusted TFP series is readily available only for the USA. See Comin et al. (2020) for utilization adjusted TFP series for major European countries.

small countries of the euro area. While the intuition behind the instruments for Canada and the small euro area countries is the same, the construction of the instrument for these two differs somewhat in practice. Roughly speaking, more than 70 percent of Canada's exports are destined to the USA. Therefore, the GDP forecast errors for the US economy can be expected to be a good predictor for unexpected movements in Canada's GDP. Contrary to Canada, for example, Finland's exports are more diversified among a number of destination countries. The top 10 export countries form roughly 60 percent of Finland's total exports, whereas at least the top 15 countries are needed to form more than a 70 percent share of total exports.[8] To form an instrument for the small euro area countries that covers roughly a similar share of exports as the instrument for Canada, we gather GDP forecast errors for several countries that are important trading partners of these countries, and weigh these errors by their respective export shares. That is, for the small euro area countries our preferred instrument is a weighted average of the trading partners' forecast errors.

We use our instrument to study fiscal policy in the small open economy context comprehensively. We identify small open economy fiscal SVARs for both Canada and the small euro area countries. Following, for example, Kim and Roubini (2008), Ravn et al. (2012), Forni and Gambetti (2016) and Klein and Linnemann (2019) we add open economy variables such as current account, real effective exchange rate and interest rate to the traditional fiscal 3-variable fiscal SVAR model. This is to better model the open economy dynamics and to control for the monetary policy stance. Canada sets its monetary policy stance independently and has a flexible exchange rate. In contrast, the euro area small open economies operate under fixed exchange rates and a given monetary policy because, as individual countries, they are too small to impact the overall stance of the monetary union. Therefore, these two types of small open economies provide an interesting comparison on the mechanisms underlying fiscal policy in a small open economy.

Results – to be added

[8]These figures are based on averages of export shares over the years 2013-2015.

The rest of the paper is structured as follows. Section 2 discusses the effects of fiscal policy in small open economies. In section 3 we lay out our empirical strategy, discuss the properties of our proposed instrument and review the data we use. Section 4 presents the main results of our analysis and section 5 concludes.

2 Fiscal policy in (small) open economies

A Blanchard and Perotti (2002) style 3-variable model, while perhaps suitable for the USA, which may be seen to resemble a closed economy, does not necessarily adequately capture the dynamic features of an open economy. Moreover, several more recent papers, which study the effects of fiscal policy, study the USA in an open economy setup. These papers include, for example, Kim and Roubini (2008), Ravn et al. (2012), Forni and Gambetti (2016) and Klein and Linnemann (2019). These studies find important evidence over the effects of fiscal policy that the paper by Blanchard and Perotti (2002) cannot reveal, for example, evidence over the twin deficit hypothesis. Considering this paper, it is also crucial to account for open economy factors because all countries studied in this paper can be considered as small open economies. Therefore, the open economy features are most likely even more prominent than in the case of the USA.

In the existing literature, there are rather few studies that empirically examine the effects of fiscal policy in the small open economy context.[9] Additionally, to the studies mentioned above, which focus largely on the USA, several empirical papers have studied fiscal policy with panel data. For example, Beetsma and Giuliodori (2011) study EU countries, Corsetti et al. (2012) concentrate on OECD countries and Ilzetzki et al. (2013) have a panel of 44 countries of different types. All of these papers examine how different background aspects affect the effectiveness of fiscal policy. For example, Ilzetzki et al. (2013) and Beetsma and Giuliodori (2011) study whether the degree of openness of a country affects the effectiveness

[9] There are some studies that concentrate on individual countries. See, for example, Ravn and Spange (2014) who study Denmark's fiscal policy and Čapek et al. (2019) who study Austria.

of fiscal policy. Additionally, these two and Corsetti et al. (2012) all examine whether the exchange rate, being either fixed or flexible, affects the effectiveness of fiscal policy. These papers also study other aspects such as public indebtedness, the health of the financial system, and the development level of a country. While these papers are related to this study, non of them explicitly study fiscal policy's effectiveness in a small open economy.

The empirical literature that studies fiscal policy of open economies refers mainly to two different theories. The older one is the Mundell-Fleming framework, whereas the newer one is the new Keynesian framework. To keep things tractable, we only briefly discuss the predictions of fiscal policy of these two theories.

According to the Mundell-Fleming framework, fiscal policy in a small open economy with a flexible exchange rate is inefficient because free capital movements prevent interest rates from deviating from the global interest rate. Furthermore, capital inflow increases the demand for domestic currency, raising its valuation that reduces net exports. This offsets the effects of fiscal stimulus. Contrary, if the exchange rate is fixed, according to this theory, fiscal policy is efficient because to maintain the fixed exchange rate, the central bank must increase its money supply which ultimately results in a rise in aggregate income.

According to the new Keynesian framework, fiscal policy in a small open economy with a flexible exchange rate might be as efficient as under a fixed exchange rate depending on the assumptions over monetary policy. This is because, in this theory, the interactions between fiscal and monetary policy are more comprehensively modeled. See, Corsetti et al. (2013). However, adding, for example, financial frictions to the new Keynesian framework (see Corsetti et al. (2013) and Born et al. (2013)), fiscal policy under fixed exchange rates is more effective than under flexible exchange rates.[10]

Within this framework the real exchange rate tends to appreciate after a government spending shock. This is in contradiction against the empirical evidence shown, for example, by Kim and Roubini (2008), Ravn et al. (2012), Bouakez et al. (2014) and Klein and

[10]Both of these studies consider a fraction of households that are excluded from financial markets.

Linnemann (2019). More so, these studies relate to the so called twin deficit hypothesis stating that government deficit and trade balance deficit seem to both increase after fiscal stimulus.[11]

Overall, the new Keynesian framework does not necessarily provide as clear predictions of the effectiveness of fiscal policy as does the traditional Mundell-Fleming framework. This is because different mechanism, assumptions and parametrizations affects how the model react to fiscal stimulus. On the other hand, the traditional Mundell-Fleming framework can be seen to have some key drawback because it lacks, for example, optimizing households and firms. Therefore, it is still largely an empirical question how fiscal stimulus affects small open economies.

Using our instrument, we study the fiscal policy of Canada and small open economies of the euro area. This enables us to compare the effectiveness of fiscal policy in two different types of small open economies. Canada can be seen as a more traditional small open economy with a flexible exchange rate and autonomy over its monetary policy. The euro area's small open economies, on the other hand, are open to trade and capital movements. However, due to their size and membership in a large currency union, they cannot affect the monetary policy stance, and the exchange rate is practically fixed.

According to, for example, Nakamura and Steinsson (2014) its reasonable to assume that these two different types of small open economies react differently to fiscal policy. Therefore, studying these two different types of economies can reveal how different economic institutions affect the dynamics of a small open economy's fiscal policy and its effectiveness. More so, it has been suggested that fiscal policy has a country-specific stabilization role in a currency union, see Gali and Monacelli (2008). The multipliers we estimate provide insight on how effective fiscal stimulus is for a small member of a currency union.

[11] Ravn et al. (2012) show that the inclusion of a deep-habits mechanism in a two-country model can explain these findings. Whereas, Enders et al. (2011) suggest that, considering a standard business cycle model, the exchange rate response depends on the parameterization of trade price elasticity.

3 Empirical strategy

3.1 SVAR specification

Reduced form VAR.— Our empirical starting point is reduced form VAR model for a vector $y' = [g, r, gdp, cab, rer, srate, defl, f_{\Delta g}, f_{\Delta gdp}]$ of nine variables: general government consumption and investment (g), government net revenue (r), GDP (gdp), current account balance as a share of GDP (cab), real exchange rate (rer), (short) nominal interest rate ($srate$), GDP deflator ($defl$) and the one step-ahead forecast of growth in g and gdp ($f_{\Delta g}$ and $f_{\Delta gdp}$). Variables g, r and gdp enter the model in real per capita terms and in natural logarithms. These three are the core three variables familiar from fiscal-SVAR literature, see, for example, Blanchard and Perotti (2002). The inclusion of cab, rer, $srate$ and $defl$ to the model is motivated by their role in the dynamics of open economies as in, for example, Kim and Roubini (2008), Forni and Gambetti (2016) and Klein and Linnemann (2019). Finally, the one-step ahead forecast variables $f_{\Delta g}$ and $f_{\Delta gdp}$ act as controls for possible foresight.[12] In our baseline model we choose not to linearly detrend the main variables of interest.[13]

Accordingly, we specify our reduced form VAR model as follows:

$$y_t = c + \sum_{i=1}^{p} A_i y_{t-i} + u_t, \tag{1}$$

where p is lag length, c is a vector of constants, A_i are the autoregressive coefficient matrices and u_t are reduced form residuals. For Canada, this model is estimated by equation-by-

[12] See, for example, Leeper et al. (2013) and Forni and Gambetti (2014). Considering the impulse responses, we find that inclusion of these variables somewhat alter the responses for Canada whereas they have a smaller effect on the responses for the small euro area countries. It has been also suggested that inclusion of forward-looking variables like interest rates can control for the foresight problem, see, for example Beetsma and Giuliodori (2011). This might partly explain why the results for the euro area countries are similar weather these forecast variables are included or not.

[13] Applied VAR papers vary in how they approach trends in the VAR. In Blanchard and Perotti (2002), a linear and a quadratic trend are included in the main specification. Caldara and Kamps (2017) detrend some of their endogenous variables. Mertens and Ravn (2013) do not add deterministic trends in their main specification, while they test their results' robustness with such a specification. Gertler and Karadi (2015) do not add deterministic trends in their main specification. According to Kilian and Lütkepohl (2017), while a VAR in levels is asymptotically valid even under true cointegration relations, its finite sample bias can be considerable. This bias is even more severe when a deterministic trend is included in the model.

equation OLS. For the euro area small open economies (Finland, Austria, the Netherlands, Belgium and Portugal) we estimate the VAR model in (1) equation-by-equation in panel form similarly to Ilzetzki et al. (2013) with country fixed effects but with common VAR coefficients on the lag terms.[14] Importantly, we consider *rer* and *srate* as exogenous variables in the case of euro area economies whereas for Canada these variables are treated as endogenous.

In lag length selection, we rely on both the standard information criteria and the partial autocorrelation function of the residuals to specify a model that both contains enough information and has no autocorrelation in the residuals. For both the euro area small open economies and Canada, our baseline specification has a lag length of 5.[15] For constructing confidence intervals we utilize the residual-based moving block bootstrap proposed by Brüggemann et al. (2016) that is shown to be applicable for Proxy-SVARs / SVAR-IVs. See Online Appendix for more details on the construction of confidence intervals.

Structural form.— Caldara and Kamps (2017) develop an analytical framework under which different identification schemes to estimating fiscal multipliers can be considered. Their SVAR framework builds on the idea of characterizing the systematic component of fiscal policy and then retrieving shocks to fiscal policy as the unexplained part in the VAR residual of the policy variable. To put their framework in more concrete terms, consider that in the data there is a positive relationship between *policy* and *output* but that *policy* may systematically react to changes in *output*. Any SVAR identification scheme must then decompose this positive co-movement between shocks to *policy* and other shocks that move *output* (Caldara and Kamps, 2017). This structural decomposition then in turn determines the estimated effect of *policy* on *output*.

In the popular Blanchard and Perotti (2002) identification scheme government spending shock is identified under the assumption that the systematic component is zero at the

[14]Euro area small open economies are chosen firstly because they are economically small compared to large euro area countries and secondly due to the timing of switching to the euro regime and related data availability.

[15]In the Online Appendix, we also examine robustness to different choices relating to the specification of the reduced form VAR.

quarterly frequency. It is thought that government spending does not react to other shocks contemporaneously due to implementation lags and thus government spending shock can be recovered as simply the reduced form residual from the VAR. Similarly, Blanchard and Perotti (2002) identify shocks to government revenues based on institutional knowledge of tax and transfers systems in order to construct the systematic component as the automatic response of revenues with respect to output.

We build our analysis on the contribution of Caldara and Kamps (2017) that non-fiscal proxies can be used to identify the systematic component of fiscal policy. Suppose we have an instrumental variable m_t that satisfies the following conditions:

$$E[m_t e_t^{\textit{non-policy}}] = \Gamma \neq 0 \tag{2}$$

$$E[m_t e_t^{\textit{policy}}] = 0, \tag{3}$$

where $e_t^{\textit{non-policy}}$ are the unexpected non-policy shocks and $e_t^{\textit{policy}}$ are the unexpected policy shocks. Given that these conditions of relevance and exogeneity with respect to non-policy and policy hold, we can use m_t as an instrument and estimate the contemporaneous elasticities of the policy variables with respect to, for example, output via 2SLS instead of relying on external estimates of the elasticities. This strategy is in contrast to some of the other studies in recent fiscal-SVAR literature that directly instrument for policy shocks. With non-fiscal proxies the identification strategy explicitly hinges on capturing the systematic component of fiscal policy.

Across the paper, we consider two different structural specifications which we label as BP and CK, for Blanchard and Perotti (2002) and Caldara and Kamps (2017) respectively. For both identification schemes we focus on a simple form of fiscal rule in that non-policy shocks can affect fiscal policy variables contemporaneously only through their effect on output. Studying US data, Caldara and Kamps (2017) argue that this simple form captures well the systematic component of fiscal policy. The only difference between BP and CK identifications in our paper is that in the former we impose a zero restriction on the contemporaneous output elasticity of government spending and investment while in the latter we do not. As discussed for example in Caldara and Kamps (2017) and Ravn and Spange (2014), even small deviations from zero in this parameter can yield considerably different results in estimated effects of government spending shocks. Here we are able to study the potential differences. When it comes to the revenue side of fiscal policy we estimate the output elasticity of net revenues using the external IV strategy and our proposed instrument also in BP identification. Originally in BP identification this parameter is assigned outside the model. We also allow for a direct effect from government spending shocks to revenues.

3.2 Non-fiscal instrument for output

Proposed non-fiscal instrument— We propose to use trading partner forecast errors of output as an instrumental variable for domestic output in small open economies. The essential assumptions needed for our instrument are the following. Firstly, the professional forecasts are sensible in the sense that forecast errors capture meaningful and unexpected variation in trading partner economies. Secondly, these forecast errors have explanatory power on the unexpected changes of aggregate output of the country of interest. Thirdly, the fiscal policy shocks of this domestic country do not explain the unexpected variation in its trading partners output. Formally, the latter two assumptions were laid out in equations (2) (relevance) and (3) (exogeneity) respectively.

In the Online Appendix we have a more detailed discussion on the validity of our instrument. Furthermore, we describe how we test instrument validity and show adequate test results. Overall our instrument fulfills the underlying assumptions well. We test relevancy with a robust F-test related to the first stage of 2SLS estimation of the structural coefficients. We also study relevancy by conducting a test where domestic forecast errors of output are explained with our instrument. This test that is not related with the SVAR-model also suggests that our instrument is related with unexpected variation in output.

The exogeneity assumption is harder to test statistically but we conduct tests motivated by Auerbach and Gorodnichenko (2012) who use OECD forecast errors of government spending and investment as proxies for fiscal shocks. Using this type of a proxy we test if our instrument's variation can be explained with fiscal shocks. These tests suggest that our instrument seems fulfill also the exogeneity assumption.

Alternative non-fiscal instrument in the literature— Earlier literature has utilized the quarterly utilization adjusted TFP series of Fernald (2014) which is available for the United States. This series is used as an instrument, for example, in Caldara and Kamps (2017) and in Angelini et al. (2020) when estimating the output elasticity of fiscal variables in a similar setting to ours. In the Online Appendix we discuss in detail the possible caveats related to this instrument.

In contrast to the TFP instrument, a major advantage of our instrument is that it is derived rather straightforwardly from observable data. This means that the decisions made by the researcher have presumably a lot smaller impact on the final instrument series and thereby also the elasticity estimate. Furthermore, the tests we utilize suggest that the instrument formed from the forecast observations fulfills the necessary assumptions more convincingly than the TFP series based instrument does.

3.3 Data

VAR variables.— We collect data from Statistics Canada, Eurostat and OECD. For g, r, gdp we rely on national statistical agencies in Statistics Canada and Eurostat while rest of the variables are collected from OECD. We limit ourselves to studying periods with relatively stable macro-institutions, that is, we wish to avoid possible structural breaks. Thus, for Canada, our sample period is 1986Q1-2019Q4 which is motivated by the start of the Great Moderation in the mid 1980s. The sample period for Euro area economies is 1999Q1-2019Q4, i.e. the EMU period. However, for Austria data are available only for 2001Q1-2019Q4. All data are seasonally adjusted when applicable. For Canada we use real government spending but for European economies we deflate government spending and investment by the GDP deflator since deflators for the individual series are not available for all countries of interest.

Instrument.— For our instrument we collect data on past macroeconomic forecasts of professional forecasters. The forecasts we consider are often reported in levels. Since the level forecasts of real variables rely on base years that change over forecast vintages and are also conditional on the then available and later revised information on past levels of the variables, we transform all level forecasts to log-difference forecasts as follows. Let $F_t[.]$ denote a forecast operator and v_{t+1} is the value of variable v in period $t+1$. Then $F_t[v_{t+1}]$ is the forecast of the value of v in period $t+1$ made in period t. Using this notation we can write the forecasted log-difference made at time t as

$$F_t[ln(v_{t+1}) - ln(v_t)]. \tag{4}$$

Forecast errors of variable v at time $t+1$ are then obtained as the difference between the realized log-difference and the forecasted one, i.e.

$$\{ln(v_{t+1}) - ln(v_t)\} - \{F_t[ln(v_{t+1}) - ln(v_t)]\}. \tag{5}$$

For Canada, we use quarterly forecasts of the US economy from the Survey of Professional forecasters (SPF). In SPF, each quarter after the first release of the previous quarter's GDP figure, a panel of professional forecasters is asked to provide forecasts of several macro-variables of the US economy. The survey is released in the middle of each quarter. In constructing the instrument, we use the mean forecast of one quarter ahead of real GDP. The real GDP forecast is transformed to a log-difference forecast as outlined above.

For Euro area economies, our data for the instrument is from the OECD Economic Outlooks (EOs), which contain forecasts of several macroeconomic variables for a number of countries. OECD EOs are published twice a year in June and December. For EOs published before 2003S2, only annual and semi-annual forecasts are available, whereas, from

2003S2 onwards, the EOs contain both annual and quarterly forecasts for a large subset of variables. When quarterly forecasts are not available, we interpolate the semi-annual growth rate (log-difference) forecast over the two quarters. In other words, we assume that the semi-annual growth rate forecast is constant across the period so that we can simply divide the semi-annual log-difference by two into two quarterly log-differences.

Since, unlike in Canada's case, no single country has an overwhelming share in the exports of a typical euro area economy, we combine forecast errors from several countries into a single instrument. In doing so, we weight trading partner forecast errors by their share in domestic exports. The quarterly data on export weights are from OECD International trade statistics.

4 Results

In this section we show results from the SVAR models described in section 3 for Canada and the small euro area countries. Considering Canada, we follow Corsetti et al. (2012) and Klein and Linnemann (2019) and study a sample starting from mid-1980s. This is to exclude the turbulent period before the Great Moderation and to ensure that the policy environment stays stable within our sample. In the panel of the small euro area countries the sample for each individual country starts from 1999. We choose this period because all countries in the panel enter the EMU period at this time and stay in the Euro regime the whole sample period.

In Table 1 we report the estimated structural parameters from CK identification (baseline) and BP style identification. The estimated size of the output elasticity of net revenue for Canada is 3.81 (CK, column 4) or 3.58 (BP, column 2) and for the euro area small open economies 1.43 (CK, column 8) or 1.42 (BP, column 6). Canada's estimate of output elasticity of net revenue is notable larger than the estimate for small euro area countries within both identification schemes. Also, the CK style identification produces a notably larger estimate for Canada than BP identification whereas for the small euro area countries the estimates are practically the same.

The parameter that represents the estimate of output elasticity of government spending is somewhat smaller for Canada than the small euro area countries with CK identification. In both cases the estimate is negative (columns 3 and 7). This parameter is restricted to zero in the BP identification (columns 1 and 5).

For Canada, the estimates of the output elasticity of net revenue is notably larger than those shown in Perotti (2005), which are derived as in Blanchard and Perotti (2002). This estimate, 1.86, is the only comparable estimate we find for Canada. According to the results in Caldara and Kamps (2017), estimating this parameter with a non-fiscal proxy seems to produce notably larger values for this elasticity also for the USA. Also Angelini et al. (2020) estimate large elasticities in a similar framework for the US. For the small euro area countries, the estimates are closer to earlier estimates. For example, Burriel et al. (2010) use an elasticity of 1.54 for the euro area.

Considering both, Canada and the small euro area countries, the first stage robust F-statistic of the small open economy fiscal model is clearly over the rule of thumb value, 10, which indicates that the instrument we use is not weak (see Online Appendix B for more).

[Table 1 here.]

4.1 Impulse responses

In Figure 1 and Figure 2 we depict the impulse responses for Canada and the euro area small open economies using both BP and CK style identifications of the structural form. For both cases we depict impulse responses and related 0.68 intervals. The black lines and shaded areas represent impulse responses and confidence intervals for CK identification and red dashed lines and dot-dashed lines represent BP identification. All impulse responses are depicted 20 quarters ahead. The impulse responses represent development of the endogenous variables to a 1 percent of GDP fiscal shocks (g and r). We first discuss the impulse responses from CK identification that are our baseline results. We then discuss the main differences in the impulse responses between these two identification schemes.

Canada - spending shock.— A shock to government spending first raises GDP but after 15 quarters the impulse response turns negative. A government spending shock has first a deflationary effect but after 5 quarters the effect on inflation turns positive. Also the real exchange rate follows a similar pattern. Current account starts to decrease straight from the first quarter after the shock and while varying over time the impulse response stays overall negative. The nominal short-term interest rate decreases on impact but roughly after two years it starts to increase. However, after 15 quarters it turns negative again. Government revenues decrease after fiscal expansion. This in turn results in a budget deficit $(r - g)$ after a government spending shock. More so, we find that a government spending shock results in both a budget and a current account deficit. This is in line with the twin-deficits hypothesis.

These results are somewhat in line with the Mundell-Fleming framework. The current account decreases which means that capital flows in. At the same time the real exchange rate appreciates. Also the short-term interest rate, while first drops, raises after 5 quarters and then again starts to decrease. There are some distinct features also. In the standard Mundell-Fleming framework inflation is fixed. However, here it seems that a government spending shock leads also to inflationary pressure. Moreover, GDP raises on impact and the effect stays positive for several years. However, in the long run the impulse response turn negative. Overall, it seems that the decrease of the current account is not enough to fully offset the effects of expansionary fiscal policy as the standard Mundell-Fleming framework would suggest.

Canada - revenue shock.— A negative shock to net revenue raises GDP and the effect seems to last almost 5 years. A negative net revenue shock starts to slightly raise inflation roughly after two years from the shock. The real exchange rate first decreases sharply but roughly after two years from the shock the response turns positive. Current account increases on impact but then quickly turns negative. The impulse response of nominal short-term interest rate first drops but roughly after two years it rises slightly. The effects of this expansionary revenue shock on $defl$, rer, cab and $srate$ are qualitatively quite similar to the spending shock but more muted.

Compared to a government spending shock a net revenues shock seems to also stimulate the economy with a smaller force. However, the cost of this type of stimulus seems to be much smaller. The revenue shock, while negative on impact, has rather little effect on the budget deficit $(r - g)$ in the medium-term.

[Figure 1 here.]

Euro area small open economies - spending shock.— A shock to government spending first clearly raises GDP (CK identification). Quite quickly the impulse response starts to decrease slowly towards zero. The overall effect is positive and lasts for several years. A government spending shock seems to have a small negative effect on domestic inflation. The impulse response for current account is clearly negative the whole period. A spending shock has a clearly negative effect on the budget balance. Therefore, these results are in line with the twin-deficits hypothesis.

Here we assume that the responses of euro area inflation, the exchange rate and the short-term nominal interest rate are exogenous, because all countries belong to a large monetary union. Presumably the actions of one small member does not affect the decisions of the central bank. This indeed seems to be largely the case if these variables are treated as endogenous in the model.

Overall, it seems that fiscal stimulus raises GDP in the short run. However, the decrease in current account indicates that at least part of the stimulus spills to other countries. That is, imports rise which dilutes the effect of fiscal stimulus.

Euro area small open economies - revenue shock.— Expansionary shock to net revenue has first a very small positive effect on GDP but afterwards the effect turns negative. It seems that the shock has a persistent effect on government revenue and in-turn after a few quarters the impulse response of government spending starts to decrease as well. In addition, a shock that lowers net revenues seems to add deflationary pressures. For the current account the impulse response decreases at first but in the long run it turns somewhat positive. The effect on budget deficit is negative. Therefore, roughly the 10 first quarters these results are in line with the twin-deficits hypothesis.

[Figure 2 here.]

Difference in impulse responses between CK and BP identification —We also find that the choice between BP or CK style identification might alter the estimated impulse responses remarkably and thereby also the estimates of fiscal multipliers (see section 4.2).[16] This difference is due to the additional zero restriction in BP identification rather than due to our instrument. In CK style identification we have an estimated value for the output elasticity of government spending which in BP is set to zero. To examine the differences between these two identification schemes we depict in Figure 3 the impact multiplier of government spending as a function of the elasticity of government spending (g) with respect to output (gdp) for both, Canada and small euro area economies. Clearly, the size of this elasticity has a remarkable effect on the impact multiplier of government spending. Moreover, it is largely this parameter that explains the differences in the impulse responses between these two identification schemes.

Considering Canada, for example, there is a clear difference in the impulse response of GDP to a spending shock between these identification schemes. CK identification produces a larger initial impact on GDP than BP identification while after 8 quarters these impulse responses are quite similar. Considering the euro area small open economies, the difference in impulse response of GDP to a spending shock stays evident the whole period studied. Also other impulse responses differ more clearly than in the case of Canada. In both cases CK identification produces a larger impact than BP identification. Considering a revenue

[16] This tendency, among these two identification schemes, to produce different impulse responses has been pointed out also by others, see, for example, Ravn and Spange (2014).

shock, all the impulse responses are seemingly similar for Canada and the euro area small open economies.

The proposed instrument produces similar sized estimates of the output elasticity of net revenues in both identification schemes. This finding causes a predicament. The BP style zero restriction of the government spending response to output within a quarter is well established in the literature. Moreover, the CK style identification does not produce statistically significant estimates of this coefficient. It seems reasonable that one could just set this parameter to zero. However, whether this parameter is zero or not has an impact on the impulse responses of main interest. Therefore, while the parameter might be statistically insignificant, the effect on the impulse responses and thereby the fiscal multipliers is remarkable.

4.2 Fiscal multipliers

After the financial crisis it has become customary to report multipliers in the fashion of Mountford and Uhlig (2009) who calculate present value cumulative fiscal multipliers. This type of a multiplier represents the dynamic multiplier as it takes into account both the dynamic response of output and the dynamic response of the fiscal variable after an shock to the latter. Accordingly, we derive the fiscal multipliers reported in this paper following Mountford and Uhlig (2009). For example, the government spending multiplier at horizon h is calculated as

$$Multiplier_h = \frac{\sum_{j=0}^{h} \frac{x_j}{(1+i)^j}}{\sum_{j=0}^{h} \frac{g_j}{(1+i)^j}} \frac{1}{\overline{g/x}}, \tag{6}$$

where h is the time horizon, i is the risk-free rate, x_j is the response of output at period j and g_j is the response of government spending at period j. The scaling factor $\overline{g/x}$ is the sample average of government spending divided by GDP. For simplicity, we use a zero discount rate. As Ramey (2019) mentions, the interest rates typically used produce very similar multipliers.

Fiscal stimulus - government spending— In Figure 4 we depict the cumulative multipliers.

The x-axis represents the horizon h the multiplier is calculate up to. The black line and shaded area represent point estimates and confidence intervals for CK identification and the red dashed line and dot-dashed lines represent BP identification. In panel A we show the cumulative multiplier for Canada and in panel B the cumulative multiplier for euro area small open economies.

The government spending multiplier for Canada (CK identification) is clearly positive and near 1 almost the whole period. Considering BP identification, the cumulative multiplier is clearly smaller, near 0.5. In both cases after 10 quarters the cumulative multiplier start to decrease slowly. Overall, government spending based fiscal stimulus seems to be quite effective in Canada.

The euro area small open economies government spending multiplier (CK identification) is small in the first quarter but the one year cumulative multiplier is already above 0.5. Afterwards the multiplier stay near 0.5 while it starts to drop after 5 quarters. Accordingly, government spending based fiscal stimulus can boost the economy but it is rather expensive or in other word inefficient. Considering BP identification, the cumulative multiplier is clearly smaller. While positive on impact after 12 quarters the multiplier turn slightly negative.

The cumulative multipliers for Canada and for the euro area small open economies differ from each other. While both multipliers are larger than 0 the multiplier for Canada is larger. This is in contradiction with the traditional Mundell-Fleming framework which suggests that fiscal policy is inefficient in a small open economy with a flexible exchange rate (Canada) and that fiscal stimulus can be effective in a small open economy with a fixed exchange rate (small euro area economies).

The smaller multiplier in the euro area small open economies might be partly explained because the effectiveness of stimulus seems to be diluted because a larger part of the stimulus leaks to other countries via imports. However, it is likely that this channel dilutes also fiscal stimulus in Canada. Another explanation might be that in Canada monetary policy can react in a way that amplifies fiscal policy whereas for a small monetary union member monetary policy is exogenous and might therefore be, for example, too tight.

[Figure 4 here.]

Fiscal stimulus - tax cuts— TBA

4.3 Robustness

Reduced form VAR specification— Evidently also the reduced form VAR specification has an effect on the impulse responses and thereby estimates of the fiscal multiplier. However, in an Online Appendix where we repeat the estimation using different choices of the lag length we find quite similar impulse responses. Overall as the estimated impulse responses using different lag lengths mostly fall on the 0.68 confidence interval of the baseline specification.[17] Differences in impulses responses seem to be driven more by the choice of identification scheme than by changes to the reduced form VAR specification.

TBC

5 Conclusion

In this paper we have proposed a novel instrument for aggregate output that is based on professional forecast errors in trading partner economies. The instrument we propose is suitable for a small open economy setting where the domestic economy has a negligible effect on the economies of its trading partners while at the same time the developments in these trading partners are highly relevant for the domestic economy.

This instrument, we argue, has a number of desirable properties when compared to the prevailing instrument for output used in the related literature which is the utilization-adjusted TFP series of Fernald (2014). Our instrument is quite simple to construct from observable data and there is arguably a much smaller role for the researcher in making

[17] When constructing confidence intervals we utilize the residual-based moving block bootstrap proposed by Brüggemann et al. (2016). See Online Appendix for details.

choices that affect the resulting series. We provide evidence for the validity of our proposed instrument. Test results for euro area small open economies and Canada show that the instrument is relevant and not weak. More so, we also find suggestive evidence that the instrument seems to fulfill the necessary exogeneity assumption. In comparison we raise some questions about the exogeneity of the utilization-adjusted TFP series with respect to government spending shocks. Furthermore, we show that revisions made to the TFP series have had a potentially large effect on its performance as an instrument. A disadvantage of our proposed instrument is that it is valid only for small open economies.

We apply the proposed instrument in estimating small open economy fiscal SVAR models using the SVAR-IV methodology for Canada and a panel of euro area small open economies. As a result of the fiscal SVAR analysis we get estimates for impulse responses and corresponding fiscal multipliers. We find that some of the impulse responses as well as government spending multipliers are sensitive to differences in SVAR identification schemes we consider (Blanchard and Perotti, 2002; Caldara and Kamps, 2017). However, broadly taken, the fiscal multipliers we calculate are in line with what one would expect in a small open economy context.

Results – TBA

Finally, we want to emphasize that the instrument we propose can potentially be used also in other contexts. There is no reason why one would be limited to utilize our instrument in estimating fiscal multipliers in a SVAR model. As an example, our instrument and the structural identifications we utilize can also be applied in local projections as well (Plagborg-Møller and Wolf, 2021). Presumably there are many other potential applications where this instrument might be useful.

References

Andrews, Isaiah, James H Stock, and Liyang Sun. 2019. "Weak Instruments in Instrumental Variables Regression: Theory and Practice." *Annual Review of Economics* 11 727–753.

Angelini, Giovanni, Giovanni Caggiano, Efrem Castelnuovo, and Luca Fanelli. 2020. "Are Fiscal Multipliers Estimated with Proxy-SVARs Robust?." Bank of Finland Research Discussion Papers 13/2020.

Auerbach, Alan J, and Yuriy Gorodnichenko. 2012. "Fiscal Multipliers in Recession and Expansion." In *Fiscal Policy after the Financial Crisis*, 63–98, University of Chicago Press.

Basu, Susanto, John G. Fernald, and Miles S. Kimball. 2006. "Are Technology Improvements Contractionary?" *American Economic Review* 96 (5): 1418–1448.

Beetsma, Roel, and Massimo Giuliodori. 2011. "The effects of government purchases shocks: review and estimates for the EU." *The Economic Journal* 121 (550): F4–F32.

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

Born, Benjamin, Falko Juessen, and Gernot J Müller. 2013. "Exchange Rate Regimes and Fiscal Multipliers." *Journal of Economic Dynamics and Control* 37 (2): 446–465.

Bouakez, Hafedh, Foued Chihi, and Michel Normandin. 2014. "Fiscal policy and external adjustment: New evidence." *Journal of International Money and Finance* 40 1–20.

Brüggemann, Ralf, Carsten Jentsch, and Carsten Trenkler. 2016. "Inference in VARs with Conditional Heteroskedasticity of Unknown Form." *Journal of Econometrics* 191 (1): 69–85.

Burriel, Pablo, Francisco De Castro, Daniel Garrote, Esther Gordo, Joan Paredes, and Javier J Pérez. 2010. "Fiscal Policy Shocks in the Euro Area and the US: an Empirical Assessment." *Fiscal Studies* 31 (2): 251–285.

Caldara, Dario, and Christophe Kamps. 2017. "The Analytics of SVARs: a Unified Framework to Measure Fiscal Multipliers." *The Review of Economic Studies* 84 (3): 1015–1040.

Čapek, Jan, Jesús Crespo Cuaresma, Johannes Holler, and Philip Schuster. 2019. "Fiscal multipliers in a small open economy: the case of Austria." *Oxford Economic Papers*.

Comin, Diego A, Javier Quintana Gonzalez, Tom G Schmitz, and Antonella Trigari. 2020. "Measuring TFP: The Role of Profits, Adjustment Costs, and Capacity Utilization." Working Paper 28008, National Bureau of Economic Research.

Corsetti, Giancarlo, Keith Kuester, and Gernot J Müller. 2013. "Floats, Pegs and the Transmission of Fiscal Policy." *Series on Central Banking Analysis and Economic Policies no. 17.*

Corsetti, Giancarlo, Andre Meier, and Gernot J Müller. 2012. "What determines government spending multipliers?" *Economic Policy* 27 (72): 521–565.

Cushman, David O, and Tao Zha. 1997. "Identifying Monetary Policy in a Small Open Economy Under Flexible Exchange Rates." *Journal of Monetary Economics* 39 (3): 433–448.

Enders, Zeno, Gernot J Müller, and Almuth Scholl. 2011. "How Do Fiscal and Technology Shocks Affect Real Exchange Rates?: New Evidence for the United States." *Journal of International Economics* 83 (1): 53–69.

Fernald, John G. 2014. "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity." *Federal Reserve Bank of San Francisco Working paper Series* 2012-19.

Forni, Mario, and Luca Gambetti. 2014. "Sufficient Information in Structural VARs." *Journal of Monetary Economics* 66 124–136.

Forni, Mario, and Luca Gambetti. 2016. "Government Spending Shocks in Open Economy VARs." *Journal of International Economics* 99 68–84.

Gali, Jordi, and Tommaso Monacelli. 2008. "Optimal Monetary and Fiscal Policy in a Currency Union." *Journal of International Economics* 76 (1): 116–132.

Gertler, Mark, and Peter Karadi. 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity." *American Economic Journal: Macroeconomics* 7 (1): 44–76.

Ilzetzki, Ethan, Enrique G Mendoza, and Carlos A Végh. 2013. "How Big (Small?) Are Fiscal Multipliers?" *Journal of Monetary Economics* 60 (2): 239–254.

Kilian, Lutz, and Helmut Lütkepohl. 2017. *Structural Vector Autoregressive Analysis.* Cambridge University Press.

Kim, Soyoung, and Nouriel Roubini. 2008. "Twin deficit or twin divergence? Fiscal policy, current account, and real exchange rate in the US." *Journal of international Economics* 74 (2): 362–383.

Klein, Mathias, and Ludger Linnemann. 2019. "Tax and spending shocks in the open economy: Are the deficits twins?" *European Economic Review* 120 103300.

Kurmann, Andre, and Eric Sims. 2021. "Revisions in Utilization-Adjusted TFP and Robust Identification of News Shocks." *Review of Economics and Statistics* 103 (2): 216–235.

Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack R Porter. 2021. "Valid t-ratio Inference for IV." Working Paper 29124, National Bureau of Economic Research.

Leeper, Eric M, Todd B Walker, and Shu-Chun Susan Yang. 2013. "Fiscal Foresight and Information Flows." *Econometrica* 81 (3): 1115–1145.

Mertens, Karel, and Morten O Ravn. 2013. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States." *American Economic Review* 103 (4): 1212–47.

Montiel Olea, José Luis, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics* 31 (3): 358–369.

Mountford, Andrew, and Harald Uhlig. 2009. "What are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics* 24 (6): 960–992.

Nakamura, Emi, and Jon Steinsson. 2014. "Fiscal Stimulus in a Monetary Union: Evidence from US Regions." *American Economic Review* 104 (3): 753–92.

Nakamura, Emi, and Jón Steinsson. 2018. "Identification in Macroeconomics." *Journal of Economic Perspectives* 32 (3): 59–86.

Perotti, Roberto. 2005. "Estimating the Effects of Fiscal Policy in OECD Countries." Discussion Paper No. 4842, Centre for Economic Policy Research.

Plagborg-Møller, Mikkel, and Christian K. Wolf. 2021. "Local Projections and VARs Estimate the Same Impulse Responses." *Econometrica* 89 (2): 955–980.

Ramey, Valerie A. 2011. "Identifying Government Spending Shocks: It's all in the Timing." *The Quarterly Journal of Economics* 126 (1): 1–50.

Ramey, Valerie 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.

Ramey, Valerie A, and Sarah Zubairy. 2018. "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data." *Journal of Political Economy* 126 (2): 850–901.

Ravn, Morten O, Stephanie Schmitt-Grohé, and Martín Uribe. 2012. "Consumption, government spending, and the real exchange rate." *Journal of Monetary Economics* 59 (3): 215–234.

Ravn, Søren Hove, and Morten Spange. 2014. "The effects of fiscal policy in a small open economy with a fixed exchange rate." *Open Economies Review* 25 (3): 451–476.

Romer, Christina D, and David H Romer. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review* 100 (3): 763–801.

Solow, Robert M. 1957. "Technical Change and the Aggregate Production Function." *The Review of Economics and Statistics* 39 (3): 312–320.

Stock, James H, and Mark W Watson. 2018. "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments." *The Economic Journal* 128 (610): 917–948.

Stock, James, and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models*, edited by Andrews, Donald W.K. 80–108, New York: Cambridge University Press.

Tables & Figures

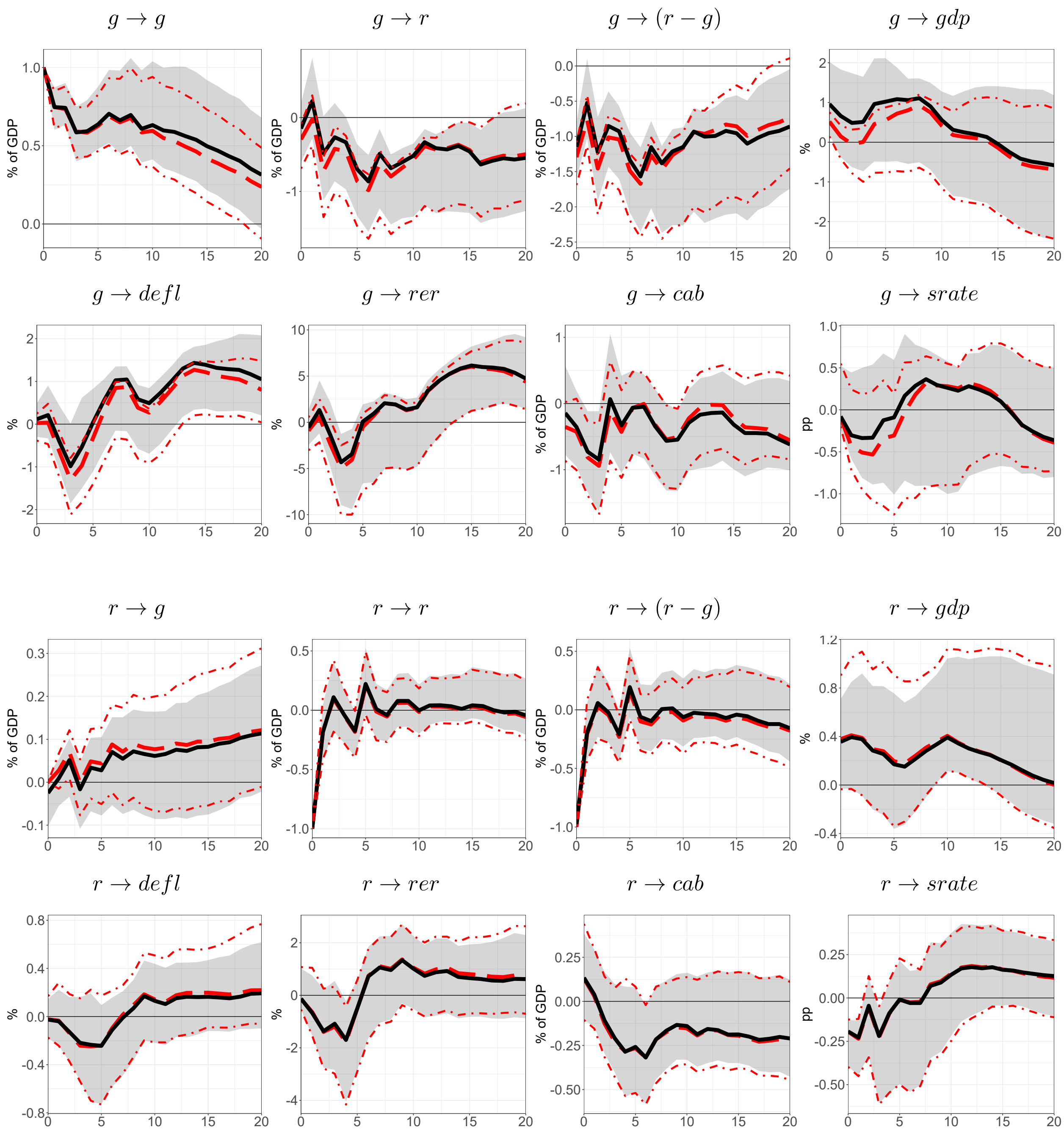
Table 1: 2SLS estimates on the output elasticities of fiscal variables.

	Canada				Pooled EUR			
	BP		CK		BP		CK	
	g	r	g	r	g	r	g	r
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
gdp	**0**	3.58**	-0.280	3.81**	**0**	1.42	-0.490	1.43
		(1.67)	(0.365)	(1.74)		(0.948)	(0.785)	(0.962)
g		-0.810				-0.030		
		(0.498)				(0.118)		
$g - a_g gdp$				-1.03*				-0.052
				(0.590)				(0.127)
Standard-Errors	Newey-West ($L = 3$)				2-way clustered (Country & Time)			
Observations	131	131	131	131	367	357	357	357
Adjusted R^2	0.994	0.961	0.993	0.961	0.997	0.992	0.997	0.992
F-stat. (1st stage), gdp		16.6	15.7	16.7		16.7	17.2	16.8
Constant	✓	✓	✓	✓				
Country FE					✓	✓	✓	✓

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

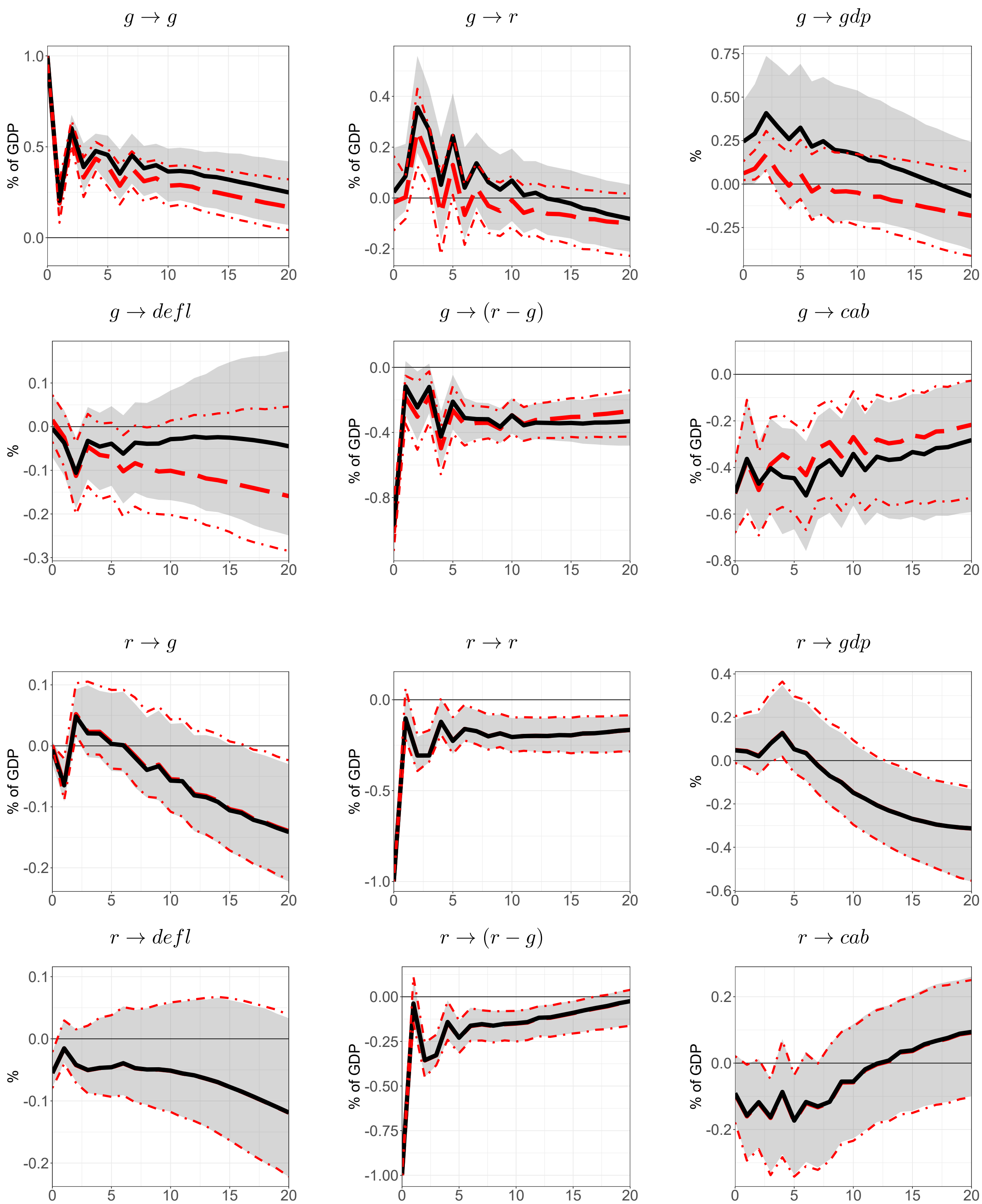
Notes: This table presents 2SLS estimates on the elasticities of fiscal variables g and r with respect to gdp which is instrumented by trading partner forecast errors of output. All models include 5 lags of VAR variables as controls.

Figure 1: SVAR impulse responses to 1% of GDP fiscal shocks, Canada.



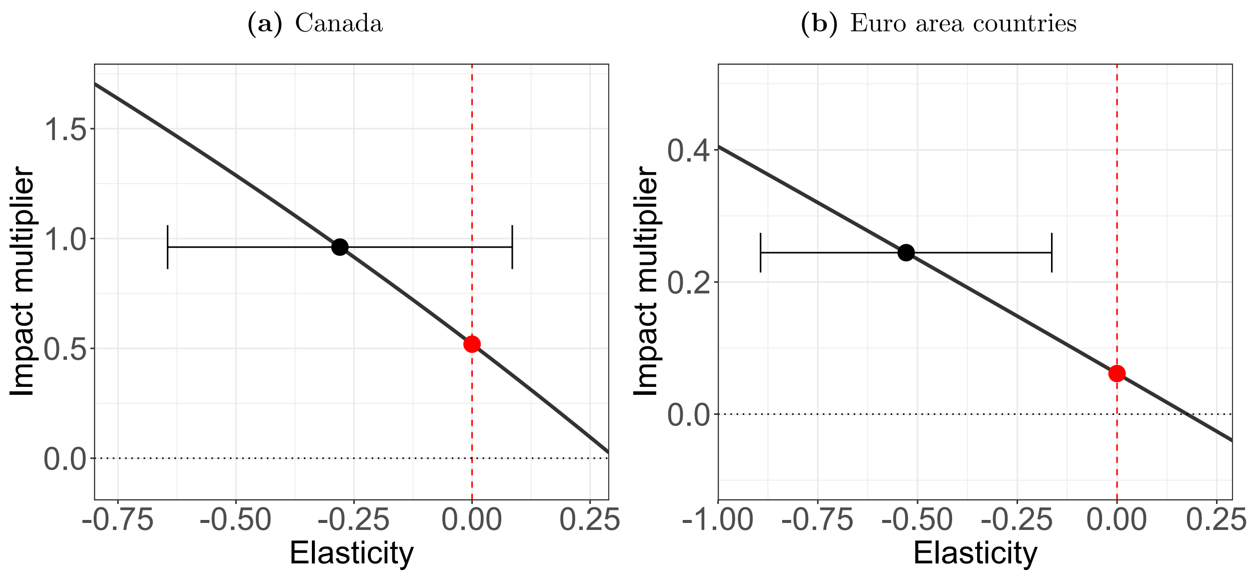
Notes: This figure plots SVAR impulse responses to fiscal shocks that either increase g by 1% of GDP or decrease r by 1% of GDP. VAR specification has 5 lags of g, r, gdp, $defl$, rer, cab and $srate$ (and forecast variables). Sample is 1986Q1-2019Q4. Black line and shaded area represent impulse responses and confidence intervals for CK identification while red dashed line and dot-dashed lines represent BP identification. Residual-based moving block bootstrap 0.68 confidence intervals with 1000 draws. Horizontal axis has quarters from 0 to 20.

Figure 2: Pooled SVAR impulse responses to 1% of GDP fiscal shocks, Euro area countries.



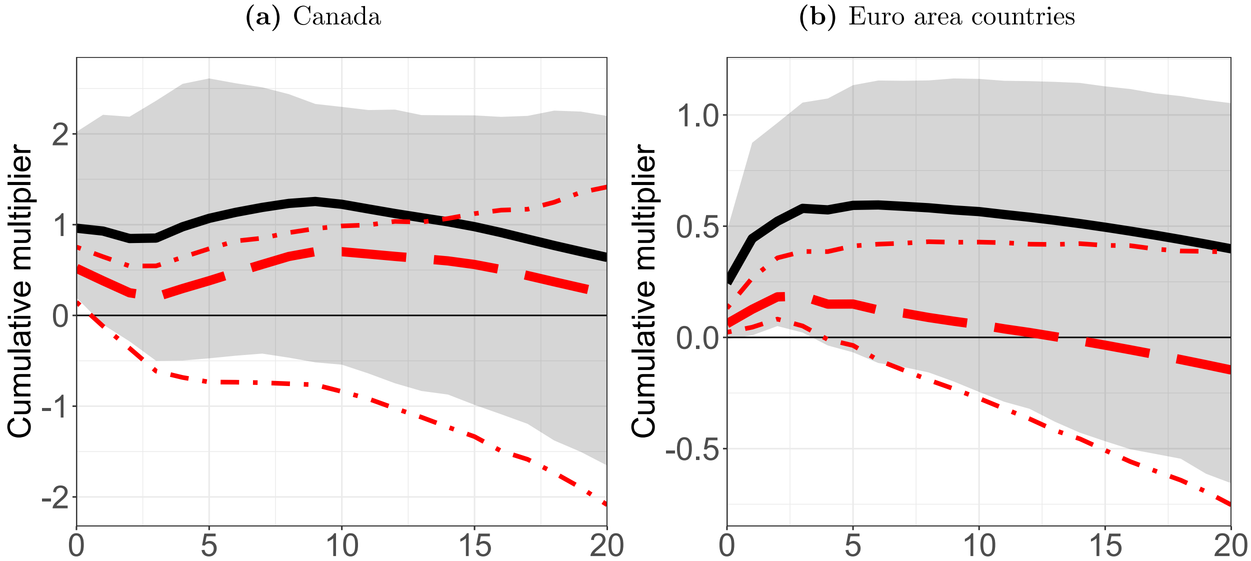
Notes: This figure plots pooled SVAR impulse responses to fiscal shocks that either increase g by 1% of GDP or decrease r by 1% of GDP. VAR specification has 5 lags of g, r, gdp, $defl$, rer, cab and $srate$ (and forecast variables). Sample is 2001Q1-2019Q4 for Austria and 1999Q1-2019Q4 for Belgium, Finland, Portugal and the Netherlands. Black line and shaded area represent impulse responses and confidence intervals for CK identification while red dashed line and dot-dashed lines represent BP identification. Residual-based moving block bootstrap 0.68 confidence intervals with 1000 draws. Horizontal axis has quarters from 0 to 20.

Figure 3: Elasticity of g w.r.t. gdp and government spending multiplier.



Notes: This figure plots the impact multiplier of government spending as a function of the elasticity of government spending g with respect to output gdp (solid black line) that is consistent with the reduced form VAR model. Red-dot corresponds to the zero restriction imposed on this elasticity by the BP identification while black-dot is the estimated elasticity along with its one standard error confidence interval from Table 1.

Figure 4: Cumulative government spending multipliers.



Notes: This figure plots cumulative government spending multipliers (Equation (6)) calculated from SVAR impulse responses to government spending shocks for up to 20 quarters from the initial shock. Black line and shaded area represent point estimates and confidence intervals for CK identification while red dashed line and dot-dashed lines represent BP identification. Residual-based moving block bootstrap 0.68 confidence intervals with 1000 draws. Horizontal axis has quarters from 0 to 20.

Online Appendix

Appendix A - Bootstrap confidence intervals

We form confidence intervals by residual-based recursive-design bootstrap (for more see, for example, Kilian and Lütkepohl, 2017). The idea is to draw bootstrap samples from the recursive VAR residuals and then recursively generate bootstrap realizations of the variables in the VAR. To generate these realizations recursively from the estimated VAR one also needs to set some initial values for the VAR variables.

Recently much attention is shown towards the inference of impulse response functions. Jentsch and Lunsford (2019) claim that the confidence intervals shown in Mertens and Ravn (2013) are too small because the wild bootstrap they utilize is asymptotically invalid. Jentsch and Lunsford (2019) propose that a valid residual-based bootstrap method should be preferred against wild bootstrap. Accordingly, such a method is the residual-based moving block bootstrap proposed by Brüggemann et al. (2016). Jentsch and Lunsford (2019) show that the impulse responses in Mertens and Ravn (2013) are not statistically significant under asymptotically valid confidence intervals even within the 68% significance level. However, they only consider standard Efron (1979) percentile intervals. That is, the intervals are calculated directly from the sample of bootstrapped impulse responses. Mertens and Ravn (2019) argue in a reply to this critic that the simulation evidence that supports the use of Brüggemann et al. (2016) bootstrap method in finite samples bases on Hall (1992) percentile intervals which allows for possible finite sample bias. Using Hall (1992) percentile intervals and residual-based moving block bootstrap Mertens and Ravn (2019) show that their impulse responses are again significantly different from zero at the 68% level. However, the 95% significance level that Mertens and Ravn (2013) originally find can not be recovered.

Based on the above discussion we utilize residual-based moving block bootstrap proposed by Brüggemann et al. (2016) and use percentile intervals that are not sensitive for possible finite sample bias. We also amend the block-design to include the instrument series as Jentsch and Lunsford (2019) propose. Contrary to the standard recursive-design bootstrap the block-design allows for heteroskedasticity in the error terms. However, in a SVAR it does not allow for serial correlation.

Kilian and Lütkepohl (2017) state that, while Efron (1979) percentile intervals are the most commonly used bootstrap confidence intervals, when modeling structural impulse responses, these should be used with caution. This is because the finite sample distribution of structural impulse response estimators is not necessarily normal. Especially, the possible finite sample bias reflects into the accuracy of Efron (1979) percentile intervals. Accordingly, this leads to low coverage of the bootstrap confidence intervals and furthermore it is not uncommon that the impulse response estimates set outside of the bootstrapped 95% confidence interval. Kilian and Lütkepohl (2017) explain that in many applications the assumption, that parameter estimates of a VAR are not systematically different from their true values, is violated because of the finite sample bias. Moreover, the bootstrap estimates tend to amplify the overall bias. This leads to a bootstrap distribution that is centered in a 'wrong place'.

Kilian and Lütkepohl (2017) review percentile intervals that allow for bias and asymmetry. They state, that in finite samples Hall (1992) percentile intervals are often less accurate than the percentile t-intervals. Furthermore, Hall (1992) percentile intervals are not systematically more accurate than Efron (1979) standard percentile intervals. Therefore, we infer that percentile t-intervals are most accurate confidence intervals in finite samples. The percentile t-intervals Kilian and Lütkepohl (2017) review, are equal-tailed percentile-t intervals (Efron, 1982) and symmetric percentile-t intervals (Hall, 1992). These are computationally heavy since a nested bootstrap round is needed in order to compute a bootstrap standard error estimate for each individual bootstrap impulse response estimates. According to Kilian and Lütkepohl (2017), the symmetric percentile-t interval is often but not always more accurate than the equal-tailed percentile-t interval.[18]

Another way to improve confidence interval accuracy which Kilian (1998) proposes is to first correct for finite sample bias in the least squares VAR parameter estimates by bootstrap and then form bootstrap intervals using Efron (1979) percentiles (for more see Kilian and Lütkepohl, 2017). Kilian and Lütkepohl (2017) note that in typical sized samples bias adjustment is not necessarily enough to ensure accurate inference when the VAR model includes a deterministic trend. The simulation results in Kilian (1999) recommend the use of Kilian (1998) method in finite samples.

After all, it is not completely clear which method to use over another in a finite sample while the (stylized) simulation based evidence supports the use of Kilian (1998). Given this and the above discussion we utilize the beforehand bias correction proposed by Kilian (1998) with Efron (1979) percentiles.

[18]The review concerning the different percentile intervals given in Kilian and Lütkepohl (2017) raises an anxiety that Jentsch and Lunsford (2019) and Mertens and Ravn (2019) both still possibly draw their confidence intervals inaccurately. Moreover, Brüggemann et al. (2016) test their bootstrap method only using Hall (1992) percentile intervals. It would be very useful to see how this method behaves in finite samples, for example, when percentile-t intervals are used.

Appendix B - Detailed discussion about the validity of the proposed instrument

In this appendix we discuss the possible caveats related to the prevailing TFP series based instrument. Discuss the validity of our instrument. Give a description over the tests we utilize to validity our instrument and discuss the results considering these tests. We also show how the different vintages of the Fernald (2014) TFP series seem to fulfill the assumptions underlying the validity of a non-fiscal instrument.

TFP instrument.—Before we proceed to a more detailed discussion of our proposed instrument it is useful to briefly highlight key differences between our proposed instrument and the currently leading instrument for output in the related literature: the quarterly utilization adjusted TFP series of Fernald (2014) which is available for the United States. This series is used as an instrument, for example, in Caldara and Kamps (2017) and in Angelini et al. (2020) when estimating the output elasticity of fiscal variables in a similar setting to ours. In contrast to the TFP instrument, a major advantage of our instrument is that it is derived rather straightforwardly from observable data. This means that the decisions made by the researcher have presumably a lot smaller impact on the final instrument series and thereby also the elasticity estimate. Furthermore, the tests we utilize suggest that the instrument formed from the forecast observations fulfills the necessary assumptions more convincingly than the TFP series based instrument does.

Construction of the utilization adjusted TFP series of Fernald (2014) involves a number of assumptions. Since TFP growth is that part of aggregate output growth that cannot be explained by changes in inputs, one first needs to at least implicitly decide over a specification of the aggregate production function. Ever since Solow (1957) it has been standard practice to calculate TFP growth as the difference between the growth in aggregate output and a weighted average of input growth rates where each input is weighted by its corresponding factor share. The Solow residual can be calculated this way under the assumption that firms minimize costs and, crucially, make no profits in which case factor shares are the correct weights. In a recent paper, Comin et al. (2020) illustrate how the zero-profit assumption as well as ignoring adjustment costs can bias the calculation of TFP. By relaxing the zero-profit assumption they show that one underestimates long-run TFP growth and overestimates its volatility and cyclicality if one uses the conventional method of calculating TFP growth.

Furthermore, to form the utilization-adjusted TFP series one needs to adjust the TFP series for capacity utilization. This involves additional decisions over how to account for or how to model/proxy the utilization rate. For example, to form their utilization adjusted TFP series Basu et al. (2006) as well as Fernald (2014) use changes in hours per worker as their utilization proxy under the assumption that it is one-to-one related to changes in capacity utilization. More recently, Comin et al. (2020) use firm surveys on capacity utilization as a more direct proxy.

Clearly any of the choices made in the process of modeling the utilization adjusted TFP series have an effect on the final TFP series. In fact, as documented in Kurmann and Sims (2021) the revisions made to the utilization adjusted TFP series as a result of new releases of data as well as methodological tweaks made by the author seem to have a remarkable effect on the resulting series of Fernald (2014).

It is also worth noting that when using a TFP series of any kind as an instrument for output one is effectively using a filtered series of output as an instrument for output. This perhaps helps to ensure the relevance of the instrument but also raises concerns about the series' exogeneity. In a fiscal SVAR for instance, for any TFP series to be a valid instrument, the series should be strictly orthogonal to fiscal shocks. Given how the TFP series is constructed, this relies critically on the filtering process to correctly filter out all demand-driven variation in the original output series (under the assumption that fiscal shocks affect demand). Yet, we show in the next subsection that the TFP series of Fernald (2014) can be predicted by government spending forecast errors of professional forecasters. This, we argue, might indicate that the exogeneity assumption does not hold for the TFP instrument. Angelini et al. (2020) allude to a similar issue.

Relevancy and exogeneity of the proposed instruments

Relevance.— The argument over why the instrument we propose should satisfy the relevance assumption is simply the following: for a country that is open to trade and whose exports form a large share of its GDP, an unexpectedly good (bad) development in its trading partners has an effect on its output through, for example, growing (declining) the demand for its exports.[19] The intuition of the relevance assumption is quite clear. The remaining question is whether this relation is strong enough to ensure that the problems arising from weak instruments are diluted.

The concept of weak instruments, which has been highlighted lately, is directly related to the relevance assumption. The weak instruments problem springs from the low correlation between the instrument and the corresponding endogenous regressor. The reason for why the weak instrument problem is highly prominent is that, if the instrument is indeed weak, it hinders inference due to the low efficiency of the estimator. Importantly, it also magnifies the finite sample bias which could mean that the estimates are far off from the true values.

[19]Note also that any common/global exogenous output shocks strengthen this relation and does not dilute the exogeneity assumption as long as it is unrelated with the fiscal policy shocks.

Considering the fact that most time series samples are quite finite, extra care should be taken to confirm that the instruments in use are not weak.

To test for the relevance and weakness of an instrument in 2SLS estimation, one can examine the F-statistics of the first stage regression. The test statistic we use is the efficient F-test proposed by Montiel Olea and Pflueger (2013).[20] As argued in Andrews et al. (2019), this test should be used instead of the standard F-statistic which assumes homoskedasticity when detecting weak instruments. Accordingly, in the case with only one instrument and one endogenous variable it is sufficient to use the following rule of thumb; efficient F statistic > 10; or to rely on the critical values provided in Stock and Yogo (2005).

The above test and the related rule of thumb can be used to detect weak instruments. This has, as we understand, more to do with the possible bias of the point estimator (location) than inference of this estimator. As Lee et al. (2021) point out this rule of thumb, while largely used, does not ensure that the standard t-statistics based inference is reliable. We think that the location of the structural parameters of the SVARs is the most important concern while naturally the related distributions are also important. Therefore, we utilize the more conventional procedures and base our inference on standard t-statistics as the efficient F-statistics is larger than 10.

We study the relevance assumption in more detail in section 4 where we actually utilize our instrument in estimating SVAR models for both Canada and euro area small open economies. Here we show evidence related to the relevance assumption of our instrument that can be in a sense considered more general compared to the F-statistics from the first stage regressions. Table A1 (panel A) reports results from regressions of our proposed instrument on GDP forecast errors and for comparisons sake the Fernald (2014) TFP series for both the sample used in Caldara and Kamps (2017) and a more recent sample for USA. The forecast errors of output can be seen as proxying unexpected GDP movements. The results show that in all cases for Canada and euro area small open economies, our instrument is significantly related with unexpected GDP movements. This is also true for both TFP instruments and unexpected GDP movements in USA, which is the expected result, for example, based on the first stage statistics shown in Caldara and Kamps (2017).

Exogeneity.— . The reasoning why our instrument should satisfy the exogeneity assumption is the following. Since the countries we consider are argued to be relatively small, unexpected shocks to their fiscal variables should not have a notable effect on their trading partners economic development contemporaneously. This arguably holds especially for a country with multiple trading partners. This is because, even if the unexpected fiscal shock

[20]Note that in the case where we have only one endogenous regressor and one instrument, this test reduces to the heteroskedasticity robust F-test.

is a direct government purchase (import) from a certain trading partner, the fact that the instrument is formed from unexpected changes in several trading partners, should dilute this relation.[21]

While the intuition behind the exogeneity assumption is credible given that we consider only relatively small economies, it would be convenient if this assumption could also be tested for statistically. Unfortunately, the exogeneity assumption cannot be straightforwardly tested. This is because we do not directly observe the exogenous shocks we are interested in. However, to test exogeneity, we gather one-period ahead forecast errors of government spending and investment from professional forecasts in the same manner as we do for the trading partner instrument. We then regress this proxy for the unexpected fiscal shock on the instrument.[22]

In Panel B of Table A1 we show results from these regressions for euro area small open economies and Canada (columns 1-2). In the same table we also show results for the utilization-adjusted TFP series of Fernald (2014) in the case of the USA. We show the results using the series which is collected from Caldara and Kamps (2017) (column 3) and a more recent 2022 revision of the same utilization-adjusted TFP series (column 4) regressed against SPF forecast errors of government spending growth for the USA.[23] As these regressions show, the coefficients for euro area small open economies and Canada are all insignificant. This, we argue, supports our claim that the instrument we propose is indeed exogenous.

For the USA there is a statistically significant relation between Fernald (2014) TFP series and the errors but not between the instrument Caldara and Kamps (2017) utilize and the errors. The utilization adjusted TFP series has seen many revisions over the years and this might according to Kurmann and Sims (2021) have a considerable effect on estimation results. To ensure that the result in Table A1 Panel B column 4 considering Fernald (2014) series is not just a one time wonder we show in the next section coefficient estimates and confidence intervals from the same regression using different vintages of the Fernald (2014) utilization adjusted TFP series. We find that overall the coefficient estimates are significant and positive across vintages.

[21]The exogeneity assumption also requires that an unexpected change in the trading partners economy does not result in an unexpected shock in government spending or revenue. Since the government is not likely to see in real time whether a trading partner is doing better than expected it is unlikely that it would make such unexpected spending or revenue decision, which would be systematically related to the unexpected changes in trading partners economies.

[22]For example, Auerbach and Gorodnichenko (2012) use the OECD forecast errors of government spending and investment as proxies for fiscal shocks.

[23]Utilization-adjusted TFP growth of Fernald (2014) are from the March 3rd 2022 vintage and downloaded from https://sites.google.com/site/fernaldjg/TFP. All other vintages used in this paper are from this source.

Table A1: Relevance and exogeneity of the instrument.

Panel A: Relevance

	Dependent variable: Forecast error of Δgdp			
	CAN (1)	EUR (2)	US(CK) (3)	US (4)
Trading partner forecast error instrument	0.463***	0.956***		
	(0.103)	(0.126)		
Δ Utilization adjusted TFP			0.085***	0.045***
			(0.027)	(0.016)
Observations	92	390	152	204
Adjusted R^2	0.188	0.366	0.133	0.038
Country FE		✓		

Panel B: Exogeneity

	Dependent variable: Instrument			
	CAN (1)	EUR (2)	US(CK) (3)	US (4)
Forecast error of Δg (OECD)	-0.048	-0.029		
	(0.107)	(0.029)		
Forecast error of Δg (SPF)			-0.414	0.829***
			(0.338)	(0.312)
Observations	92	390	101	153
Adjusted R^2	-0.008	-0.010	0.007	0.037
Country FE		✓		

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Notes: In Panel A each column reports OLS estimates from a regression of professional forecast errors of quarterly growth of GDP on different instruments. For Canada and European economies these forecast errors of Δgdp are from OECD Economic Outlooks and for the US they are from the Survey of Professional forecasters (SPF). Trading partner forecast error instrument is constructed from US GDP forecast errors (SPF) for Canada and from export-share weighted mean of OECD trading partner forecast errors (OECD) for European economies. Utilization-adjusted TFP series of Fernald (2014) is either from Caldara and Kamps (2017) (CK in column 3) or from the March 3rd 2022 vintage (column 4). In Panel B, each column reports OLS estimates from a regression of different instrument (now the dependent variable) on the professional forecast errors in the growth of the sum of general government consumption and investment. For Canada and European economies, these forecast errors of Δg are from the OECD Economic Outlooks and for the US they are from the SPF. Heteroskedasticity robust standard errors are used in columns (1), (3) and (4). Column (2) has two-way (Country $\times$ Time) clustered standard errors.

Additional results on the use of Fernald (2014) utilization-adjusted TFP series as an instrument

Next we study the different vintages of the Fernald (2014) TFP series. We try to test the relevancy and exogeneity of this series if it is considered as an TFP instrument in a similar fashion as in the previous section of this appendix. Note that the newest vintages of this series should be the most accurate version of it.

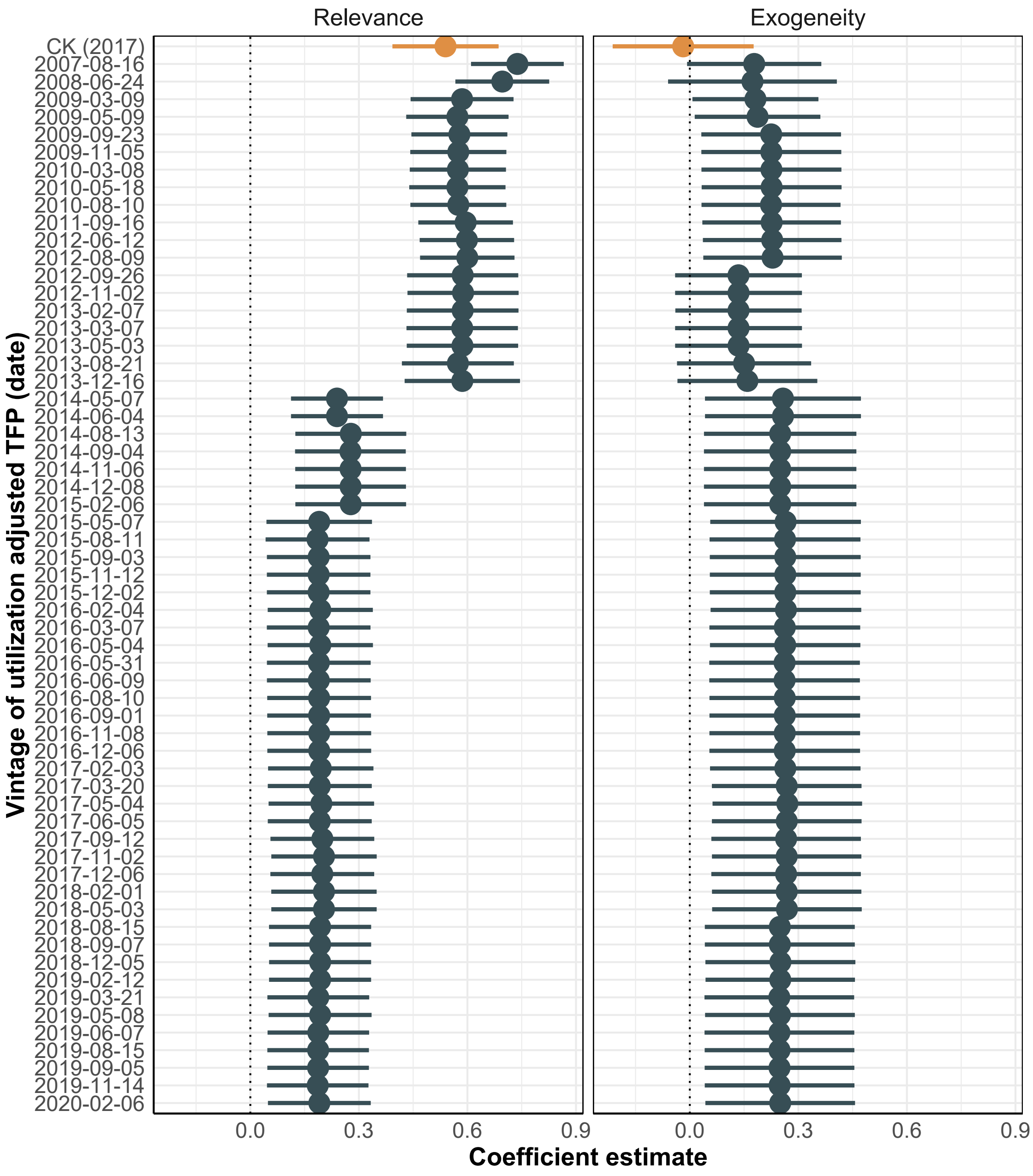
In Figure A1 we plot coefficient estimates from similar type of regressions as in Table A1 for different vintage of the Fernald (2014) TFP series. In panel A we report the coefficients and their standard error bands for relevance and in panel B the coefficients and their standard error bands for exogeneity. In panel A the coefficients before 2014 vintages are all clearly significantly different from zero. After that while significant there is a clear decrease in the size of the coefficient towards zero. Similarly in panel B before 2014 the coefficients are sometimes insignificant. But after 2014 the coefficients are significant indicating possible violation of the exogeneity assumption. Note that we mark the Caldara and Kamps (2017) instrument with yellow in Figure A1.

Furthermore, in Figure A3 we show the weak instrument test results for the TFP instrument. For these calculations we use the Caldara and Kamps (2017) data. The earlier vintages of the instrument are not weak whereas sometime during 2015 the revisions have weakened the instrument. Yet, in Figure A2 we depict estimated values of certain structural parameter of the identification schemes used in this study. Again depending on the vintage one gets very different sized coefficient estimates for these parameters. Furthermore, none of them are quite similar with the ones we get with the Caldara and Kamps (2017) instrument.

One more interesting and puzzling thing is that Caldara and Kamps (2017) refer to Fernald (2014) when describing the forming of their TFP instrument. However, we are not able to find practically any correlation between any revision of Fernald (2014) TFP series and the instrument Caldara and Kamps (2017) use. Caldara and Kamps (2017) do not provide a detailed description how they obtain their instrument. Anyway, it is not directly one of the series Fernald (2014) provides.[24] Assuming that Caldara and Kamps (2017) only used the same method as Fernald (2014) and not directly one of the ready-made series one would still expect strong correlation between these series. Furthermore, those series that Fernald (2014) provides are related with the forecast errors which suggest that these series as instruments might not be exogenous. Contrary, the instrument Caldara and Kamps (2017) use anyway seems to be exogenous according to this test, despite that it should be a Fernald (2014) style TFP series.

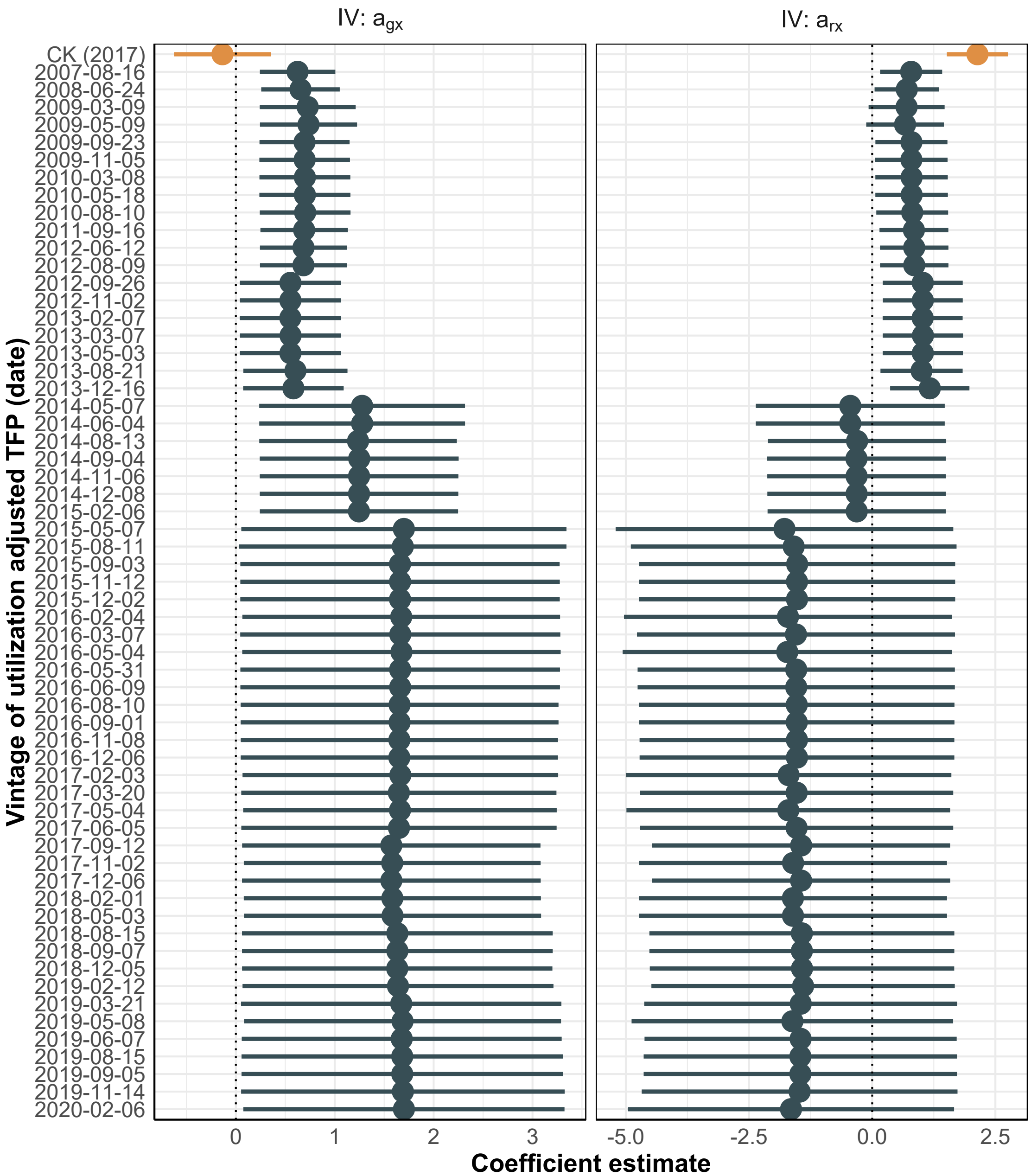
[24] One possibility is that Caldara and Kamps (2017) use some old revision or earlier version of the Fernald (2014) TFP series which is not available anymore.

Figure A1: Coefficient estimates from repeated regressions of Table A1 on the relevance and exogeneity of the utilization-adjusted TFP instrument for output using US data.



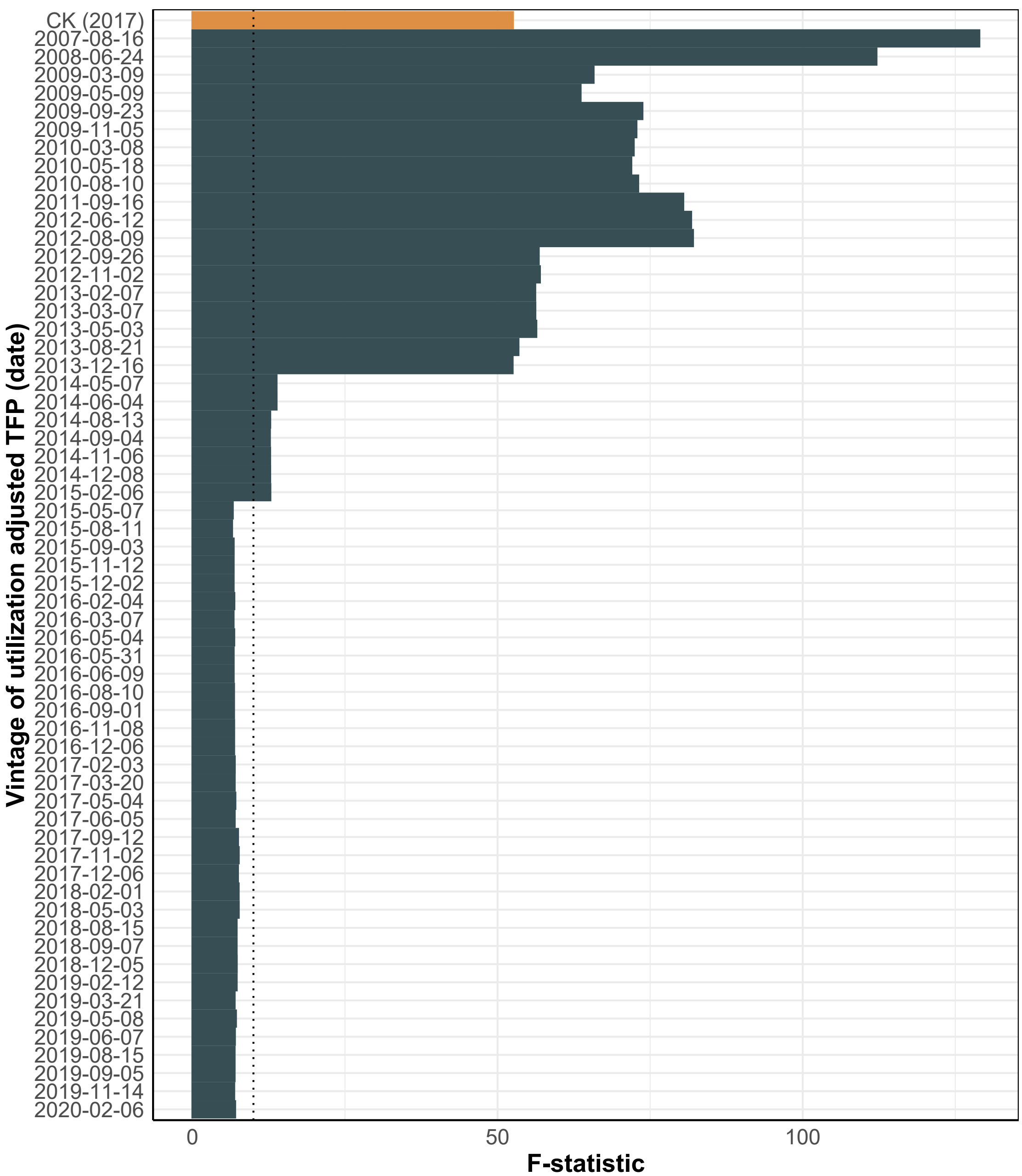
Notes: The first column plots the point estimates from repeated regressions of real GDP on the utilization adjusted TFP using data from different vintages of the series of Fernald (2014) except in the first row where data from Caldara and Kamps (2017) is used. The second column plots estimates from a model where the utilization adjusted TFP series is regressed on forecast errors of government spending. Models in both columns include four lags of real GDP, government spending and revenues as controls. Errorbars plot the 95% confidence intervals of the estimated coefficients. Forecast errors of government spending are constructed from the Survey of Professional Forecasters data. Robust Eicker-Huber-White standard errors are used throughout.

Figure A2: Estimates of certain structural parameters from SVAR-IV using US data and different vintages of Fernald (2014) TFP series as an instrument.



Notes: The columns plot the point estimates of certain structural parameters from a SVAR-IV where the adjusted TFP instrument is utilized. Estimates for the output elasticity of net revenue are shown in the right column and estimates for the output elasticity of government spending are shown in the left column. The different estimates represent different vintages of the TFP series of Fernald (2014) except in the first row where data from Caldara and Kamps (2017) is used. Models in both columns include four lags of real GDP, government spending and revenues as controls. Errorbars plot the 95% confidence intervals of the estimated coefficients. Forecast errors of government spending are constructed from the Survey of Professional Forecasters data. Robust Eicker-Huber-White standard errors are used throughout.

Figure A3: The first stage F-statistics for the output elasticity of net revenue estimates of the SVAR-IV using US data and different vintages of Fernald (2014) TFP series as an instrument.



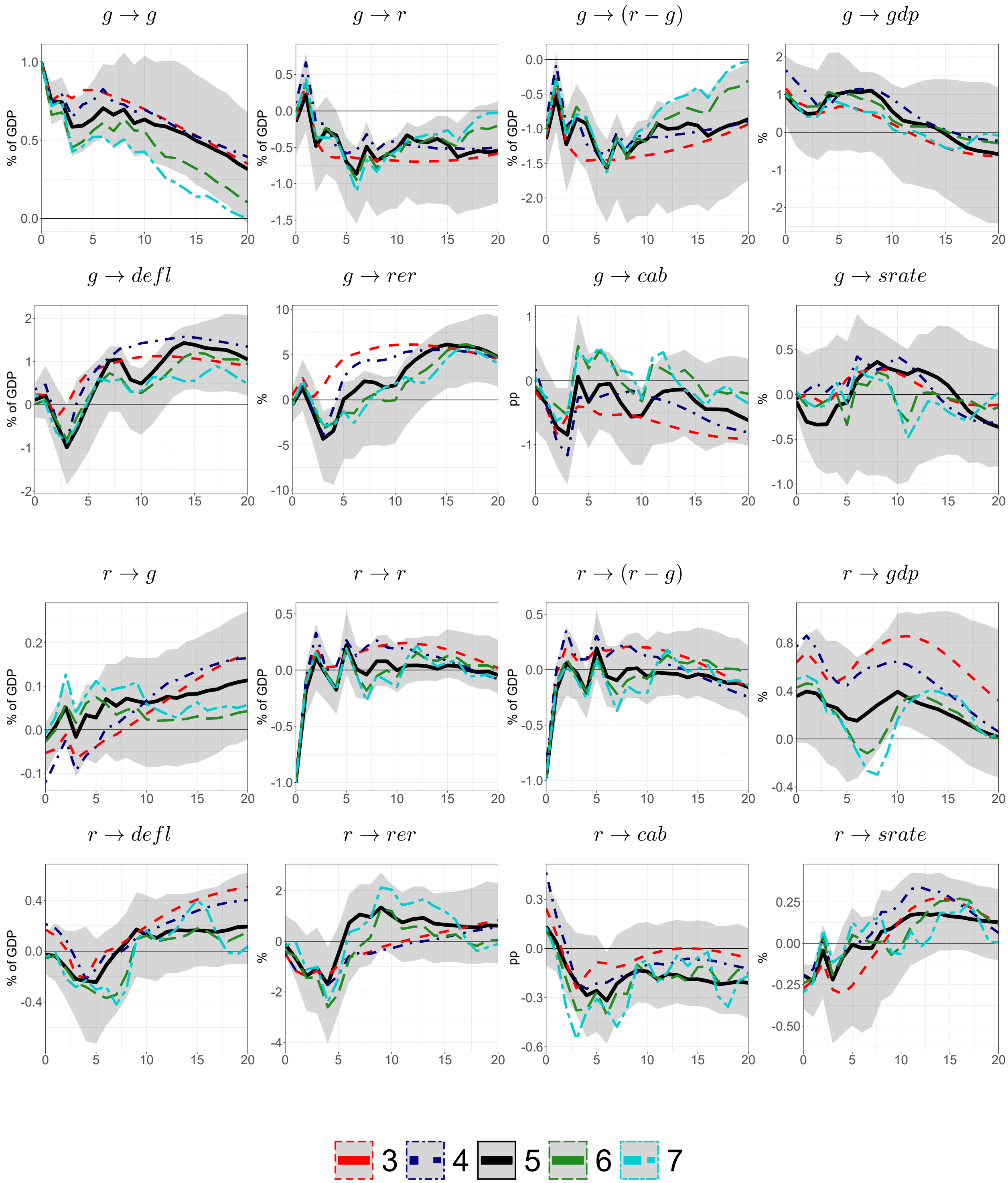
Notes: Each bar represents the first stage efficient F-statistics related with the output elasticity of net revenue estimates from SVAR-IV using US data and different vintages of Fernald (2014) TFP series as an instrument. The first bar (orange) represents a model where data from Caldara and Kamps (2017) is used.

Appendix C - Robustness to reduced form VAR specification

Figures A4 and A5 depict the impulse responses from SVARs with different lag lengths and different identification schemes for Canada and the euro area small open economies.

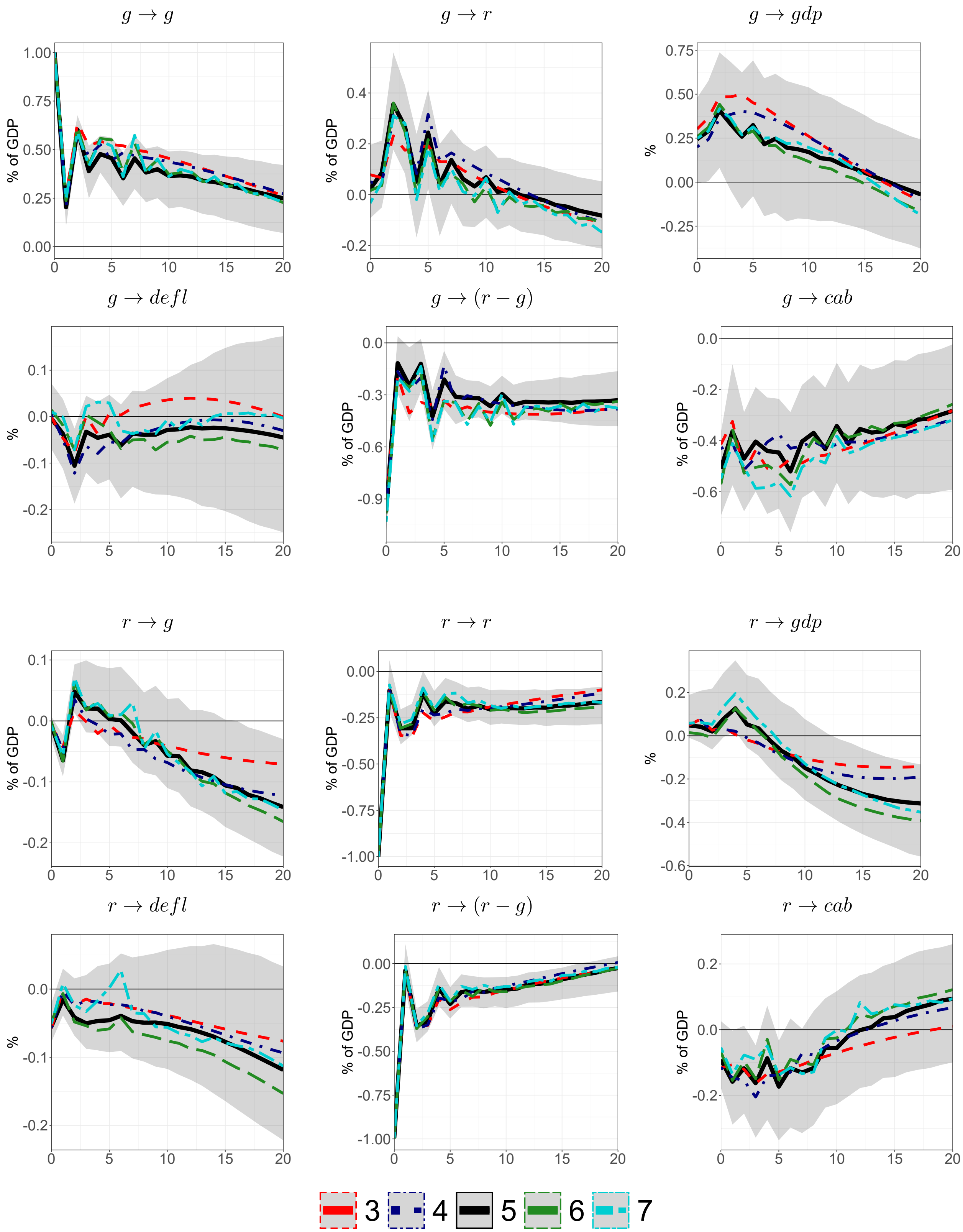
Overall it seems that the instrument is robust to choices related with VAR specification, such as whether the endogenous variables are in log levels or detrended and also what is the lag order of the reduced form VAR.

Figure A4: SVAR impulse responses to 1% of GDP fiscal shocks, Canada.



Notes: This figure plots SVAR impulse responses to fiscal shocks that either increase g by 1% of GDP or decrease r by 1% of GDP. Baseline VAR specification has 4 lags of g, r, gdp, $defl$, rer, cab and $srate$ (plus foresight). Sample is 1985Q1-2019Q4. Black line and shaded area represent impulse responses and confidence intervals for baseline CK identification while colored lines are IRFs using different lag lengths in the reduced form VAR. Moving block bootstrap 0.68 confidence intervals are for baseline specification (4 lags). Horizontal axis has quarters from 0 to 20.

Figure A5: SVAR impulse responses to 1% of GDP fiscal shocks, Euro area countries.



Notes: This figure plots SVAR impulse responses to fiscal shocks that either increase g by 1% of GDP or decrease r by 1% of GDP. Baseline VAR specification has 4 lags of g, r, gdp, $defl$, rer, cab and $srate$ (plus foresight, some variables exog.). Sample is 1999Q1-2019Q4. Black line and shaded area represent impulse responses and confidence intervals for baseline CK identification while colored lines are IRFs using different lag lengths in the reduced form VAR. Moving block bootstrap 0.68 confidence intervals are for baseline specification (4 lags). Horizontal axis has quarters from 0 to 20.

Supplementary References

Efron, Bradley. 1979. "Bootstrap Methods: Another Look at the Jackknife." *Annals of Statistics* (7): 1–26.

Efron, Bradley. 1982. *The Jackknife, the Bootstrap, and other Resampling Plans.* Volume 38. SIAM.

Hall, Peter. 1992. *The Bootstrap and Edgeworth Expansion.* Springer-Verlag, New York.

Jentsch, Carsten, and Kurt G Lunsford. 2019. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States: Comment." *American Economic Review* 109 (7): 2655–78.

Kilian, Lutz. 1998. "Small-Sample Confidence Intervals for Impulse Response Functions." *Review of Economics and Statistics* 80 (2): 218–230.

Kilian, Lutz. 1999. "Finite-Sample Properties of Percentile and Percentile-t Bootstrap Confidence Intervals for Impulse Responses." *Review of Economics and Statistics* 81 (4): 652–660.

Mertens, Karel, and Morten O Ravn. 2019. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States: Reply." *American Economic Review* 109 (7): 2679–91.