IMF Economic Review
Vol. 59, No. 4

Country Heterogeneity and the International Evidence on the Effects of Fiscal Policy

CARLO FAVERO, FRANCESCO GIAVAZZI, and JACOPO PEREGO*

This paper argues that the richer frequency and variety of fiscal policy shocks available in an international sample, which makes the use of this evidence attractive, should be analyzed recognizing the heterogeneity that exists across different countries. The main conclusion of the authors' empirical analysis is that the question "what is the fiscal policy multiplier" is an ill-posed one. There is no unconditional fiscal policy multiplier. The effect of fiscal policy on output is different depending on the different debt dynamics, the different degree of openness, and the different fiscal reaction functions across different countries. Such differences concern not only the size of the multiplier, but sometimes also its sign. There are many fiscal multipliers and an average fiscal multiplier is of very little use to describe the effect of exogenous shifts in fiscal policy on output.
[JEL H60, E62]
IMF Economic Review (2011) **59,** 652–682. doi:10.1057/imfer.2011.25

*Jacopo Perego holds a Master in Economics from Bocconi University and he is currently a PhD student in Economics at NYU. Carlo Favero is president of IGIER and a Fellow of CEPR in the International Macroeconomics Programme. Francesco Giavazzi (Bergamo, 1949) is Professor of Economics at Bocconi University in Milan. He is a research fellow and a trustee of CEPR in London, and a research associate of NBER in Cambridge. A preliminary version of this paper has been presented at the "IMF Fiscal Policy, Stabilization and Sustainability" conference, Florence, June 6–7, 2011. The authors thank Alberto Alesina, Morten Ravn, Luigi Spaventa, participants in the conference, and in the conference in honor of M. Hashem Pesaran, Cambridge, June 2011, for many helpful comments and suggestions. This paper is produced as part of the project Growth and Sustainability Policies for Europe (GRASP), a Collaborative Project funded by the European Commission's Seventh Research Framework Program, contract number 244725.

Measuring the effect of fiscal policy requires collecting a sample of episodes of exogenous shifts in fiscal stance. Such episodes, however, are rather rare at the level of an individual country. This is why, in order to obtain more precise estimates, it is tempting to pool fiscal shocks from different countries and to study their effects in the context of an international panel. Different countries, however, are different: in order to estimate fiscal multipliers using an international panel one must recognize that countries are heterogeneous. This paper considers three sources of heterogeneity: two in the transmission of fiscal shocks and one in how fiscal shocks themselves are generated. The first is specific to the analysis of fiscal policy: countries are heterogeneous in their fiscal reaction functions and therefore in the debt dynamics induced by a shift in fiscal policy. Following a fiscal shock different countries will aim at stabilizing the debt-to-GDP ratio at different levels and over different horizons. The second dimension of heterogeneity comes from different degrees of openness, which affect the way the economy responds to domestic and international shocks. The third is related to heterogeneity in the style of fiscal policy, that is in the contemporaneous correlation of shifts in taxes and spending. The same shift in primary surplus is typically achieved with a different expenditure and taxation mix in different countries. The aim of this paper is to show how the richer frequency and variety of fiscal policy shocks available in an international sample can be analyzed recognizing these sources of heterogeneity.

The (thin) available empirical literature which uses cross-country data to measure the effects of fiscal policy has so far overlooked heterogeneity. In Alesina and Ardagna (2010) and IMF (2010), for instance, fiscal multipliers are estimated by pooling all countries together, leaving the country fixed effect as the unique source of heterogeneity in the panel estimation. Ilzetzki (2011) and Ilzetzki, Mendoza, and Végh (2011) allow for the response to fiscal shocks to be heterogeneous across different groups of countries. However, they do not allow for interdependence, that is for the propagation of fiscal shocks across countries, nor for heterogeneity in debt dynamics.[1]

Recognizing heterogenous debt dynamics is important because fiscal reactions functions might differ depending on the level of debt and on the speed at which fiscal authorities let it accumulate or decline. This point has been made by Favero and Giavazzi (2007), Corsetti, Meier, and Mueller (2009), Leeper (2010) and Zubairy (2011). These papers show that studying the effects of shifts in fiscal policy without tracking the debt dynamics induced by such shifts might lead to fiscal multipliers computed along unsustainable fiscal paths, that is along a path for the debt that is at odds with the beliefs of those who hold government bonds—namely that the government will satisfy its intertemporal budget constraint. In other words, correctly estimated fiscal

[1]Also, in both papers fiscal shocks are identified within a VAR, an identification strategy which runs up against the problem of "non-invertibility" in the presence of fiscal foresight, that is whenever shifts in fiscal policy are anticipated.

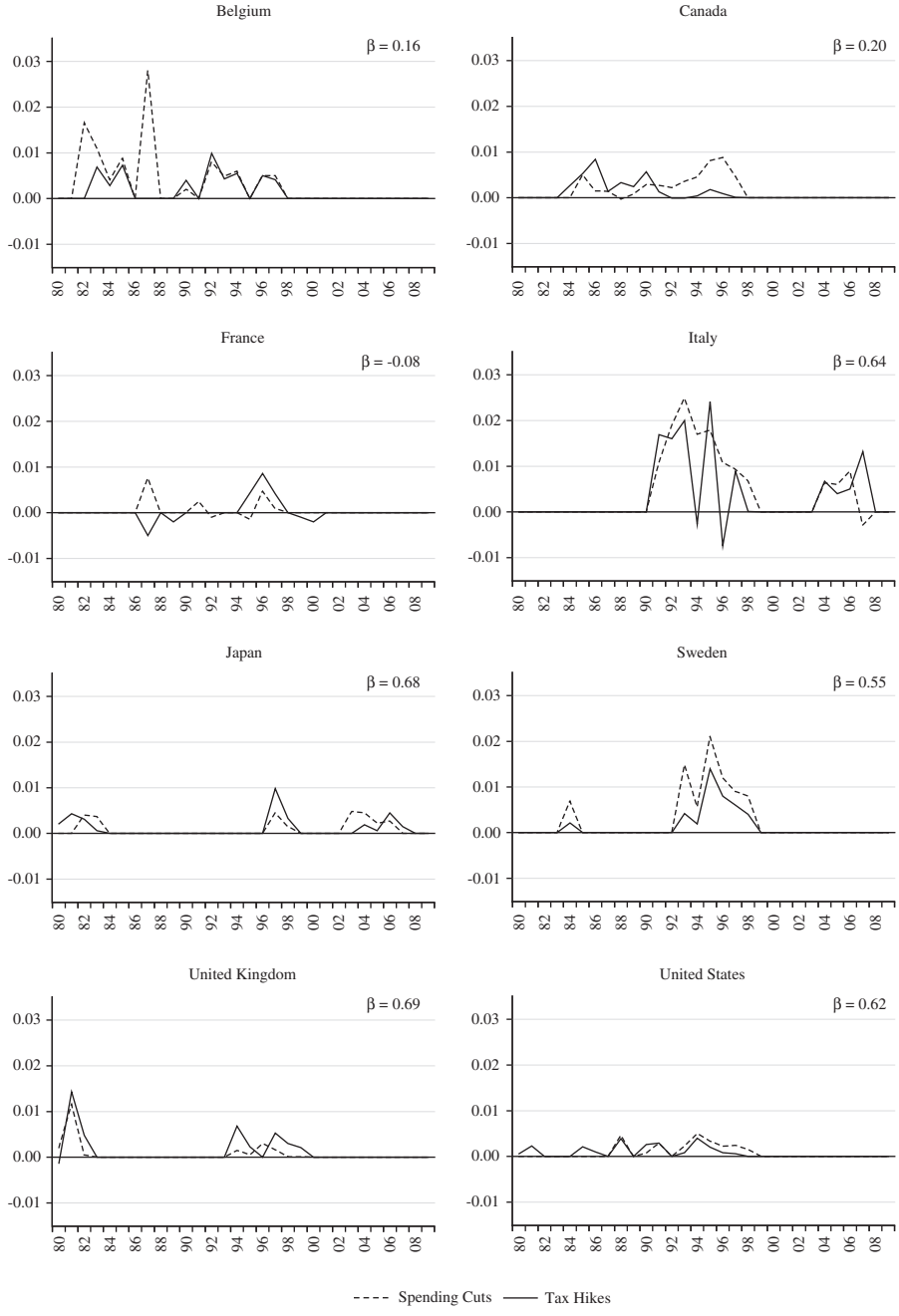
multipliers should not overlook the fact that debt sustainability requires that the government's fiscal actions should satisfy an intertemporal budget constraint. Consider, for example, a positive shift in government spending. Following the shift, the government may respect its budget constraint by adjusting taxes and spending so as to keep the ratio of public-debt-to-GDP stable, or it may delay the adjustment and in the meantime let the debt ratio grow. It may even plan to use the inflation tax. The choice of the policymaker will depend on its preferences, its policy targets, and the initial debt level: different choices will induce different responses of output and other macro variables to the same fiscal shocks. Our results show, for instance, that Canada and the United States typically respond to an increase in debt cutting spending, while Italy and, to a more limited extent, France and Japan, rely on tax increases. Analyses of fiscal policy that do not allow for heterogeneity in debt dynamics across different countries will produce an "aggregate" fiscal multiplier that could be totally irrelevant for the policymakers. As Leeper (2010) correctly argues, "Fiscal policy will shed its alchemy label when the question *What is the fiscal multiplier?* is no longer asked, and analyses of *unsustainable fiscal policies* are no longer conducted."

This paper studies fiscal multipliers estimating a multicountry nonlinear model obtained by augmenting a "global" VAR (GVAR)[2] with each country's (nonlinear) debt-deficit dynamics. The model thus allows for international spillovers and for the possibility that such spillovers, as mentioned above, work differently in different countries. We study the transmission mechanism of a particular set of shifts in fiscal policy, those identified via the "narrative" method in Devries and others (2011). These are, so far, the only available set of narrative multicountry fiscal shocks. As it is well known, the advantage of the narrative identification method is that it avoids the inversion of the MA representation of a VAR, needed to identify structural shocks within a VAR. The narrative identification is therefore robust to the effects of fiscal foresight, that is, to the possibility that shifts in fiscal policy are anticipated (see Hansen and Sargent, 1991; Leeper, Walker, and Yang, 2008, Ramey, 2011). Our main point, however—namely, the importance of allowing for heterogeneity—is independent of the particular identification strategy: it applies identically to the analysis of fiscal shocks identified imposing enough constraints on a structural VAR.

The analysis of narrative fiscal shocks for different countries reveals another source of heterogeneity: tax and spending shocks are typically not independent of one another and the style of fiscal corrections differs across countries. This simple fact is confirmed by the set of fiscal consolidation shocks identified in Devries and others (2011) and reproduced in Figure 1. In this sample, which spans from 1978 to 2009, the contemporaneous correlation of shocks to taxes and government spending is in general different from zero and the relative contribution of revenue increases and

[2]See, for example, Pesaran, Schuermann, and Weiner (2004) and Dees and others (2007).

Figure 1. The Style of Fiscal Consolidations. Fiscal Consolidation (Tax Hikes an Spending Cuts) Shocks Identified in Devries and others (2011). β is the Coefficient of the Regression of Tax Hikes on Spending Cuts



expenditure cuts to an overall shift in fiscal stance differs significantly across countries. In the case of the United States, for example, the historical data tell us that a correction of the primary surplus of 1 percent of GDP is typically achieved with a mix of 60 percent expenditure cuts and 40 percent revenue increases. In the case of Japan, instead, the same adjustment is obtained through a mix of 80 percent in expenditure cuts and 20 percent in revenues increases. Ramey (2011) recognizes this point when she observes that the correlation between revenue and spending shocks may change also within a country. When analyzing the spending shock corresponding to the Korean war she points out that what makes that shock different from the spending shocks that occurred during WWII is that it was accompanied by a contemporaneous increase in taxes, something that did not happen during WWII. This paper explicitly recognizes that shocks to revenues and expenditures are correlated and allows for such correlation to differ across countries. As we shall see, this additional source of heterogeneity has relevant implications for the analysis of the transmission of fiscal policy shocks.

Once we allow for all three sources of heterogeneity—in fiscal reaction functions, in openness, and in the style of fiscal corrections—and we compute multipliers along paths that are not inconsistent with a positive value of government bonds, we find results that sharpen our understanding of the effects of fiscal policy. In particular: (i) international spillovers are important: the effects of global stabilizations (that is, simultaneous fiscal stabilizations in all countries) are different from those of local stabilizations (that is, stabilization performed by each country in isolation). In the case of Canada, for instance, it makes a big difference whether a fiscal consolidation happens contemporaneously also in the United States, or only in Canada; (ii) the initial level of debt and the stability of the debt ratio seem to determine whether a fiscal consolidation is contractionary or expansionary.

Beyond contributing to the empirical literature on the macroeconomic effects of fiscal policy our results could also be used to discriminate between alternative theoretical models. For instance, as suggested by Perotti (2011), the finding of a fiscal multiplier smaller or larger than one can discriminate between a neoclassical and a new-Keynesian model. In neoclassical models with lump-sum taxation where government spending is pure waste and produces no externality, a shift in expenditures affects the economy only via a wealth effect. As spending rises, the need to satisfy the government intertemporal budget constraint makes the present value of taxes rise correspondingly. Note that this channel is overlooked in models that estimate fiscal multipliers omitting the government's intertemporal constraint. Forward-looking agents see their after-tax labour income reduced and will therefore cut down their consumption of both goods and leisure. Consumption falls and GDP increases (depending on the elasticity of labor supply) less than the increase in government spending: the output multiplier is smaller than one. In contrast, in a new-Keynesian model, in response to a rise in government spending consumption increases and the output multiplier is typically larger than one, provided that monetary policy does not put too

much weight on output, so that the expansion in output and labor demand are sufficient to generate an increase in the real wage.

The rest of the paper is organized as follows. In Section I we describe a model that allows for heterogeneity in fiscal multipliers. In Section II we provide some evidence on three sources of heterogeneity: in the style of fiscal corrections, in fiscal reaction functions, and in openness. Section III presents our empirical results and discusses what difference all of this makes. Section IV examines the robustness of our results to the potential endogeneity of narrative shocks. Section V concludes.

I. Measuring International Fiscal Multipliers

As mentioned in the introduction, this paper does not address the issue of the identification of fiscal policy shocks. We instead focus our attention on the transmission mechanism of fiscal shocks using a particular set of narrative shocks identified in Devries and others (2011), who apply the narrative method originally proposed by Romer and Romer (2010, hereafter R&R) to identify shifts in fiscal policy in a group of 15 OECD countries.

The literature that pools fiscal shocks from different countries to study their effect on output (for example, Alesina and Ardagna, 2010; IMF, 2010) typically estimates a panel of cross-country pooled output equations. The growth rate of real GDP is regressed on a set of current and lagged values of the fiscal shocks and on lagged GDP growth. In particular, IMF (2010) estimates, using a first version of the Devries and others (2011) shocks, on the sample of 15 OECD countries[3] the following specification:

$$\Delta y_{i,t} = \alpha + \beta \Delta y_{i,t-1} + B(L)\varepsilon^{g}_{i,t} + C(L)\varepsilon^{\tau}_{i,t} + \lambda_i + \nu_t + \mu_{i,t}. \tag{1}$$

The equation includes a full set of country dummies, λ_i, to account for differences in trend growth rates across countries and time dummies, ν_t, to account for global shocks, such as shifts in oil prices or the global business cycle. The simulation of the dynamic effects of the structural shocks (shifts in taxes, $\varepsilon^{\tau}_{i,t}$ and in expenditure, $\varepsilon^{g}_{i,t}$) generates a single fiscal multiplier restricted to be the same for all countries.

Differently from this approach, we study the effects of fiscal shocks in our panel of countries allowing for heterogeneity in the style of fiscal corrections, in the countries' degree of openness and in their debt-deficit dynamics. The model is specified to contain the minimal set of macroeconomic variables necessary to pin down the debt-deficit dynamics endogenously:

$$\begin{aligned} X'_{i,t} &= C_{i,1} + C_2 X'_{i,t-1} + \varphi_{i,1} B_{i,t-1} + \varphi_{i,2} Z_{i,t} + \varphi_{i,3} Z_{i,t-1} + \gamma^{g}_{i} \varepsilon^{g}_{i,t} + \gamma^{\tau}_{i} \varepsilon^{\tau}_{i,t} + \eta_{i,t} \\ &\quad \text{if } i \neq us, \\ X_{i,t} &= C_{i,1} + C_{i,2} X_{i,t-1} + \varphi_{i,1} B_{i,t-1} + \gamma^{g}_{i} \varepsilon^{g}_{i,t} + \gamma^{\tau}_{i} \varepsilon^{\tau}_{i,t} + \eta_{i,t} \quad \text{if } i = us. \end{aligned} \tag{2}$$

[3]Australia, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Portugal, Spain, Sweden, the United Kingdom, and the United States.

with

$$\begin{aligned} X'_{i,t} &\equiv [y_{i,t}, g_{i,t}, \tau_{i,t}, i_{i,t}, p_{i,t}, s_{i,t}], \\ X_{i,t} &\equiv [y_{i,t}, g_{i,t}, \tau_{i,t}, i_{i,t}, p_{i,t}], \\ Z_{i,t} &\equiv [y^*_{i,t}, s^*_{i,t}], \\ \varphi_i &\equiv [\varphi_{i,1}, \varphi_{i,2}, \varphi_{i,3}] \end{aligned} \quad ,$$

where $y_{i,t}$, $g_{i,t}$, $\tau_{i,t}$, $p_{i,t}$, $s_{i,t}$ are (respectively) the logs of real output, real government expenditures, real government revenues, the GDP deflator, and the real exchange rate relative to the U.S. dollar, $i_{i,t}$ is the nominal average cost of financing the gross government debt. We append to the model the following set of identities that determine the dynamics of the gross debt-GDP ratio, B

$$\begin{aligned} \tilde{Y}_{i,t} &= e^{y_{i,t}+p_{i,t}} \\ \widetilde{gr}_{i,t}, &= (\tilde{Y}_{i,ti,t} - \tilde{Y}_{i,ti,t-1})/\tilde{Y}_{i,ti,t-1}, \\ B_{i,t} &= B_{i,t-1}\left(\frac{1+i_{i,t}}{1+\tilde{gr}_{i,t}}\right) + \frac{e^{g_{i,t}} - e^{\tau_{i,t}}}{e^{y_{i,t}}} + \zeta_{i,t} \end{aligned} \quad ,$$

where $\tilde{Y}_{i,t}$ is the level of nominal output (where throughout a $\sim$ denotes nominal variables). We use the vector of global variables $Z_{i,t}$ to parsimoniously model comovements in open economies by adopting the GVAR approach proposed by Pesaran, Schuermann and Weiner (2004): a country-specific exogenous international variable, $y^*_{i,t}$, is constructed for each country through a weighted average of the (log of) foreign output

$$y^*_{i,t} = \sum_{j=1}^{n-1} w_t^{ij} y_{j,t} \quad i \neq j,$$

where the weights w_t^{ij} are based on trade shares—the share of country j in the total trade of country i measured in U.S. dollars with $w_t^{ii} = 0$. The current value and the first lag of $y^*_{i,t}$ are included in the specification of each country's VAR to capture international comovements in the cycle. We adopt the same procedure to model relative prices, using each country's real exchange relative to the U.S. dollar to construct the following global variable

$$s^*_{i,t} = \sum_{j=1}^{n-1} w_{ij} s_{j,t} \quad i \neq j.$$

To track the country-specific debt dynamics we must first recognize that the equation which determines the evolution over time of the debt-income ratio is highly nonlinear. The fact that this equation is nonlinear is the reason why it must be tracked by means of endogenous variables rather than simply augmenting the VAR with the government debt series. These endogenous

variables are those determining the path of government debt: the cost of debt service, the growth rate, and the primary deficit. In what follows, we derive the debt dynamics in terms of *gross* debt and, by doing that, we slightly depart from previous work such as Bohn (1998), who uses *net* government liabilities as his definition of public debt. We use gross debt for several reasons. First, statutory debt limits, when they exist, are usually imposed on gross debt. Second, gross debt is the measure which is more largely available to the public and, for this reason, it is more likely to be the one entering the information set of economic agents, hence influencing their decisions when responding to fiscal shocks. Third, there is an inherent difficulty in evaluating government assets, most of which do not have a market price to be used as a reference. The last reason is technical: in two of the countries in our sample, Sweden and the United Kingdom, the net debt series turns negative for some years. This is a problem because whenever the net debt comes close to zero in our simulation it induces an exploding path for the cost of debt service, hence making the system unstable and the simulation unfeasible.

In order to track the debt dynamics, we start from the two following identities:

$$\begin{aligned} \tilde{B}_{i,t} &\equiv \tilde{B}^{n}_{i,t} + \tilde{A}_{i,t}, \\ \tilde{B}^{n}_{i,t} &\equiv \tilde{B}^{n}_{i,t-1} + \tilde{D}_{i,t} + \tilde{I}_{i,t} + \mu_{i,t}, \end{aligned} \tag{3}$$

where $\tilde{B}_{i,t}$, $\tilde{B}^{n}_{i,t}$, $\tilde{A}_{i,t}$, $\tilde{D}_{i,t}$ and $\tilde{I}_{i,t}$ denote, respectively, the nominal levels of gross debt, net debt, government assets, the primary deficit, and net interest payments. The error term, $\mu_{i,t}$, is to be interpreted as a zero-mean vector of statistical discrepancies. From Equation (3), by adding and subtracting $\tilde{A}_{i,t-1}$, we obtain

$$\tilde{B}_{t} \equiv \tilde{B}_{i,t-1} + \tilde{D}_{i,t} + \tilde{I}_{i,t} + \Delta\tilde{A}_{i,t} + \mu_{i,t}. \tag{4}$$

Dividing both sides of Equation (4) by nominal GDP, $\tilde{Y}_{i,t}$ (and dropping the tilde to denote ratios to GDP) we have

$$B_{i,t} \equiv \frac{\tilde{B}_{i,t-1} + \tilde{I}_{i,t}}{\tilde{Y}_{i,t}} + D_{i,t} + \nu_{i,t} + \mu_{i,t}. \tag{5}$$

$\nu_{i,t} = \Delta\tilde{A}_{t}/\tilde{Y}_{t}$ denotes the component in the change of gross public debt which reflects the change in the value of government assets. Since we have no economic model to determine the evolution of government assets, we shall assume that $\nu_{i,t}$ is an exogenous random variable. For notational convenience we define $\zeta_{i,t} \equiv \nu_{i,t} + \mu_{i,t}$. From Equation (5) we get

$$B_{it} = B_{it-1}\left(\frac{1 + i_{it}}{1 + \tilde{g}r_{i,t}}\right) + \frac{e^{g_{i,t}} - e^{\tau_{i,t}}}{e^{y_{i,t}}} + \zeta_{i,t}. \tag{6}$$

Note that the specification of the set of identities that determine the debt dynamics imposes a nonnegativity constraint on the cost of financing the debt, a feature that will turn out to be very useful when simulating the model over periods of very low interest rates. Note also that, as mentioned above, conditional on $X_{i,t} \equiv [y_{i,t}, g_{i,t}, \tau_{i,t}, i_{i,t}, p_{i,t}]$ and $\zeta_{i,t}$, the system is closed, which means that we have expressed the dynamics of gross debt/GDP ratio, $B_{i,t}$ in terms of endogenous variables only.[4]

Model (2) has a few features that are worth discussing:

- it includes a nonlinear debt feedback.[5] Therefore, the impact of fiscal shocks on output depends on the initial conditions at which such shocks occur, as represented by the preshock debt level B_{it-1}.[6] Following a fiscal shock, however, debt stabilization is not imposed: the coefficients on the debt feedback are freely estimated. Note that these coefficients are allowed to be different across countries, so that our specification can accommodate heterogeneous debt-deficit dynamics. One restriction we impose on the $\varphi_{i,1}$ coefficients is that, for every country, debt only appears in the equations for $g_{i,t}$, $\tau_{i,t}$, $i_{i,t}$ and $p_{i,t}$. Debt is dropped from the equation for $y_{i,t}$ because in that equation—and in that equation only—it turns out not to be statistically significant;
- it allows to compute impulse responses to fiscal shocks keeping track of the debt dynamics. If $\varepsilon^{g}_{i,t}$ and $\varepsilon^{\tau}_{i,t}$ are validly identified shocks, the only additional assumption required to track the debt dynamics by appending Equation (6) to the VAR is that $\zeta_{i,t}$ is strongly exogenous. $\zeta_{i,t}$, captures the statistical discrepancies in the debt-deficit dynamics reconstructed using the OECD variables, and is the only additional shock that needs to be added to the VAR in order to compute the debt dynamics;
- $\varepsilon^{g}_{i,t}$ and $\varepsilon^{\tau}_{i,t}$—the shocks identified in Devries and others (2011) with the narrative method, thus not requiring the inversion of the Moving Average representation of a VAR—are directly included in the VAR and impulse responses with respect to these shocks can be directly derived from the joint simulation of Equation (2) and the associated identities;

[4]In order to check how closely our debt-dynamics equation tracks the actual path of debt-GDP ratios for the eight countries in our sample, we have brought the system (6) to the data and simulated it forward starting in 1980, feeding it with the actual values of $X_{i,t}$ and $\zeta_{i,t}$. Figure 3 shows that the debt dynamics produced by this simulation are virtually not distinguishable from the actual ones.

[5]Ghosh and others (2011) also find evidence of the importance of a debt feedback and of nonlinearities in an international panel of advanced economies. Zubairy (2011) allows for a debt feedback in a DSGE model, while Killian and Vigfusson (2010), in the case of oil shocks, and Auerbach and Gorodnichenko (2011) also allow for the presence of asymmetries and nonlinearities in a VAR model.

[6]Impulse responses and their associated confidence intervals are computed using the simulation technique described in Favero and Giavazzi (2007).

- the degree of openness is allowed to differ across countries by letting the coefficients $\varphi_{i,2}$ and $\varphi_{i,3}$ be country-specific;
- the United States is treated as a closed economy. This is not an identifying restriction. We have imposed that foreign variables have no impact on the U.S. economy to be able to compare our results with the existing empirical evidence that typically analyzes the effects on fiscal policy in the United States using a closed economy specification. When the validity of the closed economy restrictions for the United States is tested statistically, the hypothesis that all the relevant coefficients are zero could not be rejected;
- finally the model allows for the correlation between revenue and spending shocks and for heterogeneity across countries in the "style" of fiscal policy. When a fiscal adjustment of 1 percent of the GDP is simulated in country i, we pair a shock of size $1/1+\beta$ to $\varepsilon^{g}_{i,t}$ with a shock of size $\beta/1+\beta$ to $\varepsilon^{\tau}_{i,t}$, where β is computed using the fact that $E_t(\varepsilon^{\tau}_{i,t}|\varepsilon^{g}_{i,t}) = \beta\varepsilon^{g}_{i,t}$.

II. Three Sources of Heterogeneity in Fiscal Multipliers

Our specification allows for all three sources of heterogeneity discussed in the introduction. We consider each of them in detail, before illustrating our empirical results.

Heterogeneity in the Style of Fiscal Corrections

The international exogenous fiscal consolidation identified in Devries and others (2011) are tax increases and spending cuts implemented to reduce the budget deficit and to put the public debt on a sustainable path.[7] Such shocks are identified for a group of OECD countries using the records available in official documents to identify the size, timing, and principal motivation for the fiscal actions taken by each country. This identification strategy applies to a panel of countries; the idea originally proposed in R&R who used Presidential speeches, Congressional reports, and other public records to identify all major U.S. postwar tax policy actions. However, the Devries and others (2011) shocks differ from R&R's in two important dimensions. R&R focus *only* on revenue shocks and identify *two* main types of legislated exogenous tax changes: those driven by long-run motives, such as to foster long-run growth, and those aiming to deal with an inherited budget deficit. Devries and others (2011) instead consider *both* expenditure and revenue shocks and focus *only* on fiscal actions motivated by the objective of reducing a budget deficit.

This observation raises the possibility of a truncation problem in the Devries and others (2011) shocks series. A truncation would arise if deficit-driven fiscal expansion episodes were dropped. Although we cannot

[7]The data set is available on the IMF website (www.imf.org/external/pubs/cat/longres.aspx?sk=24892.0).

check for truncation in general, there are two countries in our sample for which this is possible, the United States and the United Kingdom. Consider the case of the United States, for which the Devries and others (2011) shocks can be compared with the R&R narrative shocks. The R&R narrative shocks show both positive and negative observations; these shocks, however, are constructed aggregating tax shocks that are deficit-driven and tax shocks driven by a long-run growth motive. Deficit-driven fiscal expansions never occur in the R&R sample, because all tax shocks driven by the *long-run* motive are expansionary (that is, negative tax shocks) and all the *deficit-driven* tax shocks are contractionary (that is, positive tax shocks). Therefore the R&R deficit-driven shocks, which are directly comparable to those identified by Devries and others (2011), show no evidence of truncation. Moreover in the R&R identification, *deficit-driven* tax shocks and *long-run* tax shocks are virtually orthogonal (their correlation is −0.08). The same observation—namely the fact that the series of deficit-driven tax shocks is almost exclusively composed of tax increases—applies to the narrative series of deficit-driven shocks identified by Cloyne (2011) for the United Kingdom. Note, however, that orthogonality of *deficit-driven* tax shocks and *long-run* tax shocks makes the multiplier computed using only deficit-driven fiscal shocks unbiased but does not make it directly comparable with the one computed using the R&R's series. The former is a multiplier with respect to deficit-driven fiscal shock *only*. The latter, instead, is relative to a generic fiscal shock, either long-run or deficit driven, obtained by imposing the restriction that the output responses to long-run motivated tax changes and to deficit-driven tax changes are identical.

The original Devries and others (2011) sample includes 15 OECD countries. The data are annual and extend from 1978 to 2009.[8] In this sample 173 episodes of fiscal consolidation are identified. In what follows, however, we restrict our attention to a subsample of eight countries: Belgium, Canada, France, Italy, Japan, the United Kingdom, Sweden and the United States. This choice is constrained by the availability of the data needed to track the debt dynamics—such as general government gross debt and interest payments—which for some of the countries in the original Devries and others (2011) sample are available only for a too short time span.

As it is apparent from Figure 1, revenue shocks and expenditure shocks are correlated, and the fiscal mix historically used to achieve a correction in the budget is heterogeneous across countries. In the case of the United States, for example, the historical data tell us that a correction of the primary surplus of 1 percent of GDP is typically achieved with a mix of 60 percent expenditure cuts and 40 percent revenue increases. In the case of Japan, instead, the same adjustment is obtained through a mix of 80 percent in expenditure cuts and 20 percent in revenues increases.

[8]The source for all variables, other than the narrative fiscal shocks, is the OECD database.

The evidence in Figure 1 has two important implications. First, it tells us that, for basically all the countries considered, the simulation of the effects of a shock to government spending, assuming no contemporaneous shift in taxes, would violate the historical pattern. Such an experiment would describe a situation that does not exist in the data—because $\varepsilon^{\tau}_{i,t}$ shocks have never occurred independently of $\varepsilon^{g}_{i,t}$ shocks, at least in this sample. This observation casts strong doubts on the usefulness of using the narrative shocks identified in Devries and others (2011) to study the effects of tax-based adjustments *separately* from those of expenditure-based adjustments. If the identified spending and revenue shocks have a specific pattern of correlation, that specific pattern should be preserved when simulating the effect, for instance, of a tax shock. In other words, it would be difficult to interpret the effect of a tax shock which is assumed to take place independently of an expenditure shock since such an occurrence has never been observed in the sample from which the data are drawn. Second, the evidence described in Figure 1 implies that, when studying the international evidence on the effects of a fiscal correction, one should allow for this source of heterogeneity in policy, that is for the different styles of such corrections across countries. A shift in the primary surplus equivalent to 1 percent of GDP is not achieved with the same mix in all countries. This restriction, which is implicitly imposed both in Devries and others (2011) and in R&R, violates the heterogeneity present in the data.

To illustrate the importance of this point we have run an experiment focusing on the United States only. Consider a regression of output growth on a distributed lag of fiscal shocks estimated to evaluate the impact on output of (i) a tax shock of 1 percent of GDP simulated setting expenditure shocks to zero (the experiment run by R&R), and (ii) an adjustment of the primary surplus of 1 percent of the GDP obtained using the historical mix of shifts in taxes and in expenditure. In practice, we have estimated the following two models, where $i = US$ and $A(L, q)$ is a lag polynomial of degree q[9]:

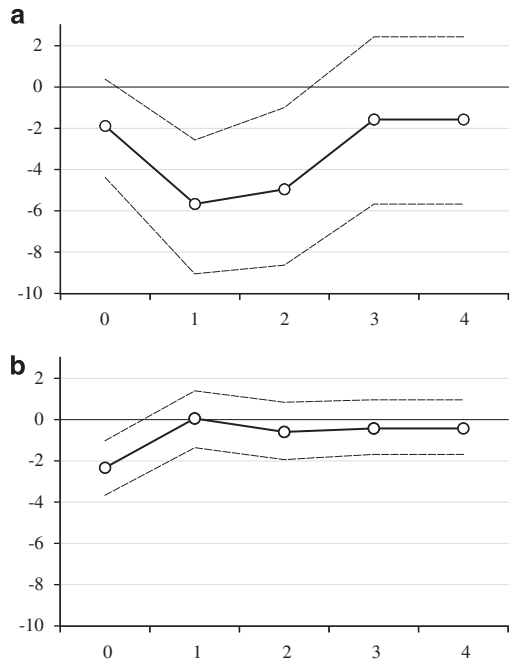
$$\Delta y_{i,t} = \alpha + \beta \Delta y_{i,t-1} + B(L,2)\varepsilon^{\tau}_{i,t} + \mu_{i,t}, \tag{7}$$

$$\Delta y_{i,t} = \alpha + \beta \Delta y_{i,t-1} + B(L,2)\varepsilon^{g}_{i,t} + C(L,2)\varepsilon^{\tau}_{i,t} + \mu_{i,t}. \tag{8}$$

The results are reported in Figure 2. The multiplier obtained from Equation (7), reported in the left-hand panel, is estimated by simulating a shock to $\varepsilon^{\tau}_{i,t}$ equivalent to 1 percent of GDP. The multiplier obtained from Equation (8), reported in the right-hand panel, is estimated by simulating a shock of $1/1+\hat{\beta}$ to $\varepsilon^{g}_{i,t}$ and a shock of $\hat{\beta}/1+\hat{\beta}$ to $\varepsilon^{\tau}_{i,t}$. The coefficient $\hat{\beta}$ comes from the estimation of $\varepsilon^{\tau}_{i,t} = \alpha + \beta\varepsilon^{g}_{i,t} + \nu_{i,t}$ in the sample. In this second experiment the overall simulated shift in fiscal policy still amounts to 1 percent of GDP, but it now reflects the fiscal policy style observed in the

[9]Where the lag-polynomial is defined as $M(L,q) = \sum_{i=0}^{q} \beta_q L^q$.

Figure 2. The United States: Impulse Responses of Output to a Shift in Fiscal Policy Corresponding to 1 Percent of GDP: (a) Tax Hikes only; (b) Tax Hikes and Spending Cuts



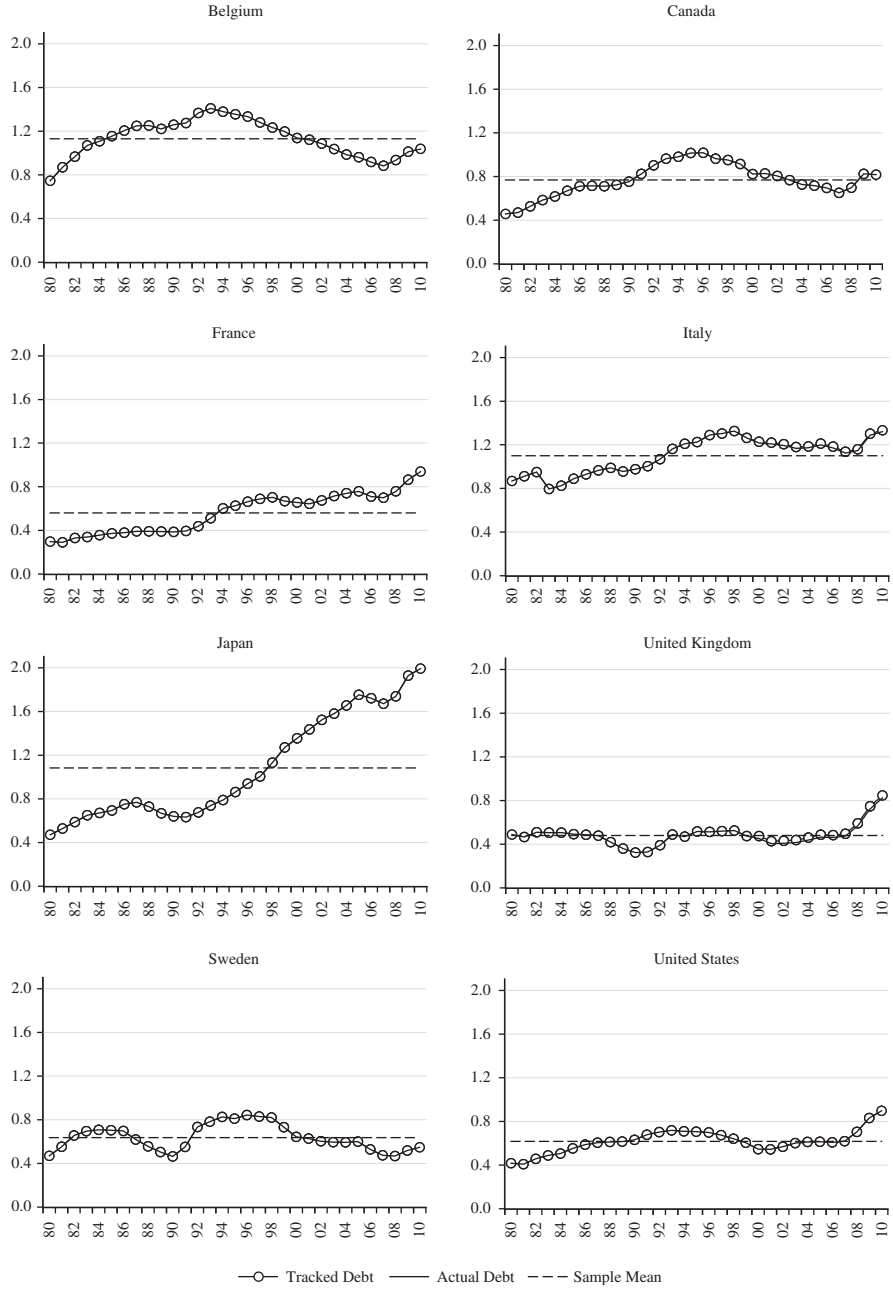
data. As Figure 2 shows, the two multipliers are quite different. The multiplier obtained using the actual style of fiscal corrections differs from zero only in the first year and is much smaller than that obtained by simulating an isolated tax shock, which is negative and significant for three years following the shock.

In the light of this difference, we conclude that multipliers should be estimated based on the historical correlation between shifts in taxes and in spending, rather than artificially setting to this correlation to zero. This is nothing new: the simulation of reduced form models such as a VAR, not respecting the historical pattern of correlations present in the data, would be subject to the Lucas (1976) critique.

Heterogeneity in Fiscal Reaction Functions

Countries are also heterogeneous in their fiscal reaction functions: following a shift in fiscal policy, different countries will aim at stabilizing the debt-to-GDP ratio at different levels and over different horizons. In other words, the effects of a shift in fiscal policy will depend on the country-specific debt-deficit dynamics: Figure 3 illustrates that these dynamics is clearly heterogeneous across the eight countries in our sample.

Figure 3. Tracking Debt Dynamics. Actual and Simulated Debt Dynamics. The Simulated Debt Dynamics is Obtained by Dynamic Simulation of the Government Budget Constraint with Observed Data



Heterogeneity in Openness

A third dimension of heterogeneity is related to the different degrees of openness, because openness determines the size of the multiplier and the extent to which an economy is affected by international fluctuations. Openness varies a lot across the eight countries in our sample. The United States is the closest of all. In most empirical investigations on the effect of fiscal policy it is treated as a closed economy: we shall not depart from this hypothesis, assuming that the U.S. economy is unaffected by international fluctuations (as we mentioned above, when this assumption is tested it is not rejceted on our sample). This, however, is not true for smaller economies where the effect of a shift in fiscal policy, at home or abroad, will depend on the international economic environment in which such a shift takes place. It has been argued (see for example Perotti, 2011) that the sharply different response of the Irish economy to the two consolidations carried out during the 1980s—which resulted in a deep recession in 1981–82 and in an economic boom five years later—were associated with the very different economic conditions prevailing at the time in Ireland's main trading partner, the United Kingdom.

III. Results

The presentation of our results is organized in four subsections. We start by discussing the robustness of fiscal multipliers estimated on panels of countries. We then explain why it is important to keep track of debt dynamics and we show this with a case study of the United States. Finally, in Section "Computing the Effects of Fiscal Policy Allowing for Heterogeneity," we report our results.

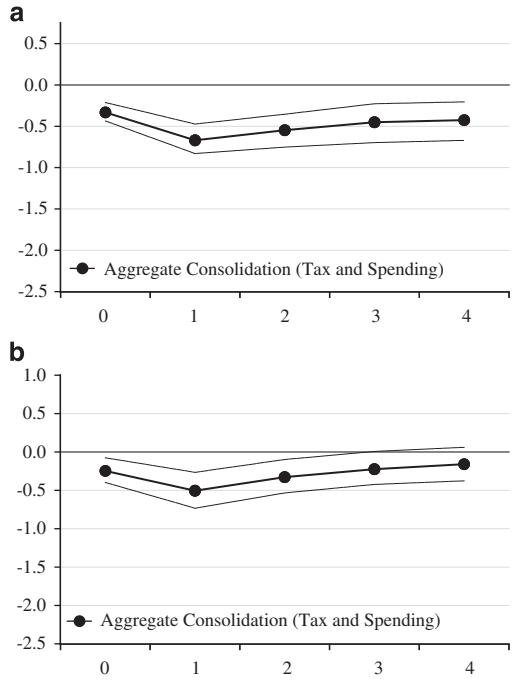
On the Robustness of International Fiscal Multipliers

We start our empirical analysis by replicating the available international evidence on the fiscal transmission mechanism (for example, Alesina and Ardagna, 2010; IMF, 2010), which, as we said, is typically based on the panel estimation of a cross-country output equation. The specification is a regression of the growth rate of real GDP on a set of current and lagged values of fiscal shocks and lagged GDP growth. In particular, IMF (2010) estimates, on their sample of 15 OECD countries, the following equation:

$$\Delta y_{i,t} = \alpha + \beta \Delta y_{i,t-1} + B(L,2)\varepsilon^{g}_{i,t} + C(L,2)\varepsilon^{\tau}_{i,t} + \lambda_i + \nu_t + \mu_{i,t}. \quad (9)$$

The equation includes a full set of country dummies, λ_i, to account for differences in trend growth rates across countries and time dummies, ν_t, to account for global shocks, such as shifts in oil prices or the global business cycle.

Figure 4. (a) Devries and Others Replication: Single Equation, 15 OECD Countries (b) Devries and Others Estimated on a Subsample of Eight Countries



We replicate the results of the IMF study by reporting in Figure 4a the multiplier with respect to an aggregate fiscal shock, $\varepsilon^{g}_{i,t} + \varepsilon^{\tau}_{i,t}$, obtained by imposing $B(L,2) = C(L,2)$. When aggregate shocks are considered, the estimated multiplier is statistically significant but smaller than 1.

The simple empirical model described by Equation (9) imposes very strong restrictions. The effects of fiscal consolidations are assumed to be identical across countries: the only heterogeneity allowed for is that captured by the fixed effects in the panel estimation. We doubt that this global fiscal multiplier is a useful concept for the selection of the structural model to be used for policy advice. The following assumptions, in particular, appear to be very restrictive:

- fiscal shocks are assumed to be identical across all countries. No heterogeneity in the fiscal policy mix is allowed for;
- the responses of output to fiscal shocks are computed overlooking their effects on the dynamics of debt. The specification thus rules out the possibility that fiscal dynamics differ across countries characterized by different debt levels. It also shuts down another possibly important effect, namely the effect that fiscal shocks can exert on interest rates;

- fiscal multipliers are assumed to be the same in small and open, and large and less open economies. Moreover, the model leads to the evaluation of the effect of a *global* fiscal shock, and it is silent on the effect of *local* fiscal shocks, that is shocks that occur in one country only.

On the Importance of Tracking Debt Dynamics

To illustrate the importance of keeping track of the debt dynamics induced by a shift in fiscal policy, we start by considering a restricted version of our general empirical model obtained setting $\varphi_{i,1}$, $\varphi_{i,2}$ and $\varphi_{i,3}$ to zero. Equation (10) encompasses the single equation specification used in the Devries and others (2011) study, but it also allows to keep track of the debt dynamics when computing impulse responses, thus checking whether multipliers are computed along divergent fiscal paths. Otherwise it replicates the Devries and others study in that no debt feedback is imposed

$$X_{i,t} = C_{i,1} + C_{i,2}X_{i,t-1} + \gamma_i^g \varepsilon_{i,t}^g + \gamma_i^\tau \varepsilon_{i,t}^\tau + \mu_{i,t} \tag{10}$$

with $X_{i,t} = [y_{i,t}, g_{i,t}, \tau_{i,t}, \tau_{i,t}, p_{i,t}]$. The usual set of identities in Equation (6) is appended to Equation (10) in order to track debt dynamics endogenously. (Note that because we now keep track of debt dynamics the sample is restricted to only eight countries). Equation (10) can be interpreted as a set of stacked closed economy VARs: no exchange rate is included and no common fluctuations among different components of $X_{i,t}$ across countries is allowed for ($\varphi_{i,2} = \varphi_{i,3} = 0$). Moreover, if panel restrictions are imposed, such that, for every country i, $C_{i,1} = C_1$, $C_{i,2} = C_2$, $\gamma_i^g = \gamma^g$ and $\gamma_i^\tau = \gamma^\tau$, Equation (10) can be reinterpreted as an approximation of the truncated MA representation of Equation (9).

We have estimated the system (10) on data from our sample of eight countries. Figure 4b shows that the estimated multipliers replicate very closely those obtained with the IMF specification, Equation (9), and reported in Figure 4a.

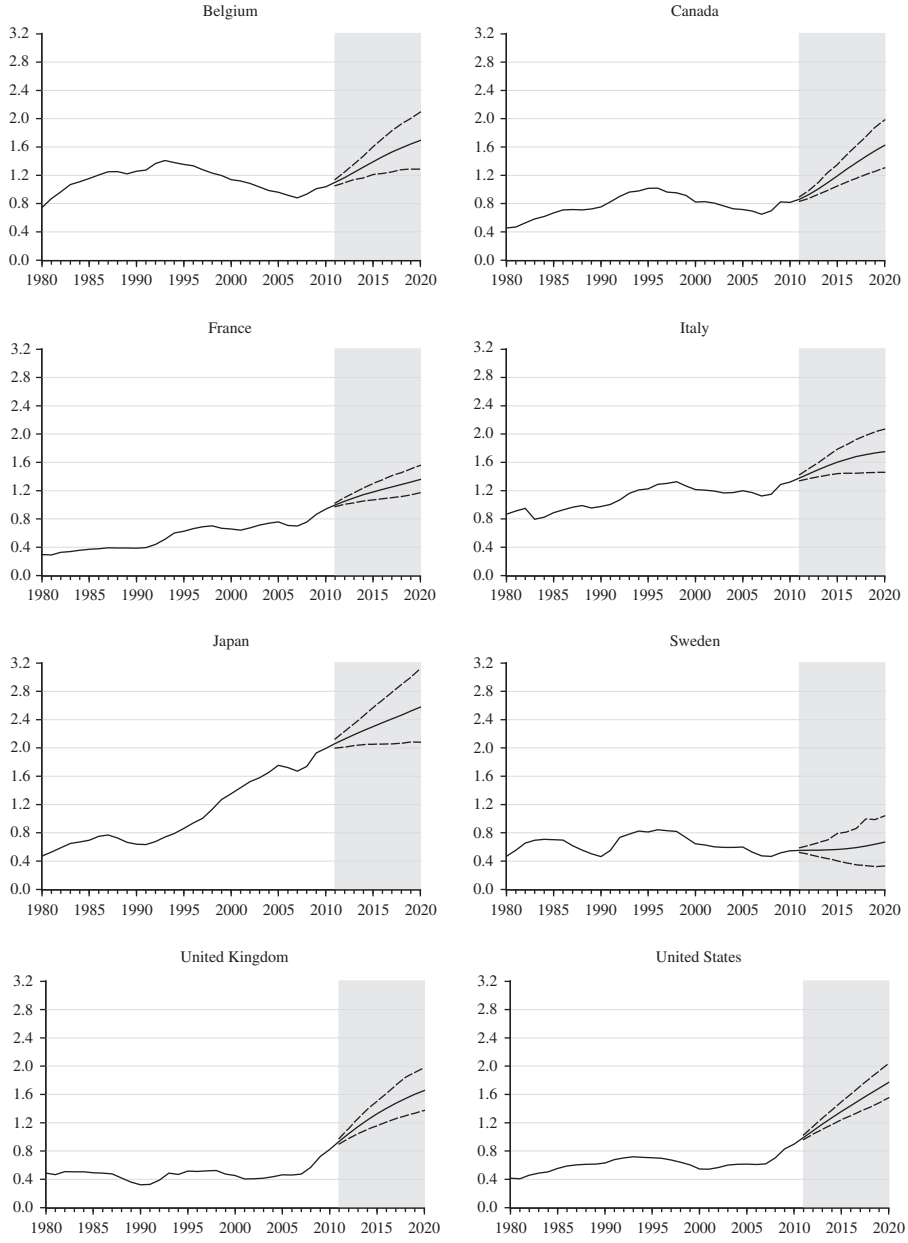
Figure 5 reports the simulated debt dynamics for each of the countries in the sample: for most of the countries—with the only clear exception of Sweden—the common multiplier is computed along an unstable debt path.

We now come to the core of the paper. We shall estimate fiscal multipliers in a model that allows for a debt feedback, international comovements, and cross-country heterogeneity. Before attacking this problem, however, we show a case study of the United States to document the error one can make by omitting the debt feedback.

The Effects of Overlooking the Debt Feedback: A Case Study of the United States

This section illustrates the importance of keeping track of the effects of fiscal policy on debt when estimating fiscal multipliers. We study the United States as a closed economy for the reasons explained above. We start by estimating

Figure 5. Debt Dynamics Out-of-Sample Simulations (Shaded Area)



two models for the U.S. economy on the sample 1980–2009: a standard VAR model without debt feedback (equation (10)) and one with debt feedback. (In this case the set of regressors in each of the VAR equations is augmented by the lagged debt-to-GDP ratio and the debt dynamics is modeled by the identities in equation (6)).

In practice, we consider the following system of equations for the U.S. economy

$$X_{\text{us},t} = C_{\text{us},0} + C_{\text{us},1}t + C_{i,2}X_{\text{us},t-1} + \varphi_{\text{us}}B_{\text{us},t-1} + \gamma^{g}_{\text{us}}\varepsilon^{g}_{\text{us},t} + \gamma^{\tau}_{\text{us}}\varepsilon^{\tau}_{\text{us},t} + \mu_{\text{us},t}, \tag{11}$$

where, as above, $X_{\text{us},t} \equiv [y_{\text{us},t}, g_{\text{us},t}, \tau_{\text{us},t}, i_{\text{us},t}, p_{\text{us},t}]$. The vector of coefficients φ_{us} describes the feedback from the lagged debt-GDP ratio to the variables included in the system. As in the previous sections, the debt dynamics is endogenized by appending to the system in Equation (11) the identities described in Equation (6).

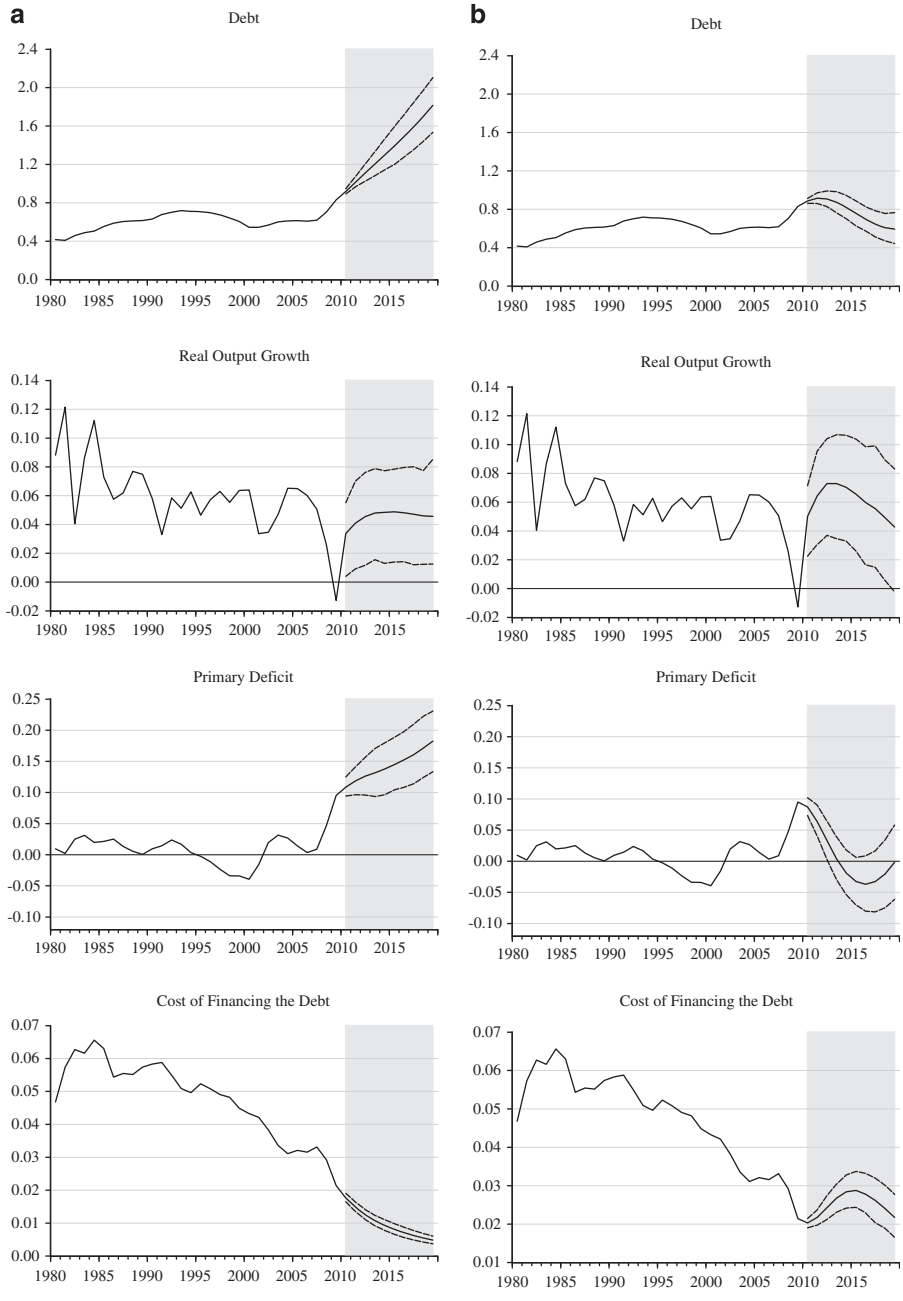
To understand the importance of allowing for a debt feedback in estimating the fiscal multiplier, we shall consider two alternative specifications of this model. First, we analyze Equation (11) *without* feedback, that is, we impose the restriction $\varphi_{\text{us}} = 0$. Next, we relax this assumption and re-estimate the same model allowing for $\varphi_{\text{us}} \neq 0$. When we do this we let $\varphi_{\text{us}} = \{0, \varphi^{g}_{\text{us}}, \varphi^{\tau}_{\text{us}}, \varphi^{i}_{\text{us}}, \varphi^{p}_{\text{us}}\}$, that is we let the feedback affect all variables $X_{\text{us},t}$ except $y_{\text{us},t}$ (because, as mentioned above, debt does not enter significantly the equation for output). We shall refer to this model as the fiscal VAR *with* debt feedback.

The two alternative specifications, with and without debt feedback, have strikingly different effects on the dynamics of the endogenous variables following a fiscal shock—and this plays an important role when computing fiscal multipliers. To illustrate this point, we report in Figure 6 the simulated out-of sample dynamics of output growth, of the debt-to-GDP ratio, the primary-deficit-to-GDP ratio, and the cost of financing the debt, as generated by the VAR without feedback (left column) and with a debt feedback (right column). The simulated series are generated by taking, as initial conditions for all variables, their value in 2009 and then projecting each future path up to 2020 by solving the nonlinear model forward. Thus, the initial conditions for the out-of-sample simulations reported in Figure 6 reflect the U.S. fiscal expansion of 2008–09.

Figure 6 shows that the dynamics implied by the VAR model with no debt feedback is unstable for all fiscal variables, although real GDP growth converges to a long-run value of about 4 percent. The same long-run steady state for growth is obtained by the model with debt feedback, but with a very different path for the fiscal variables.

The out-of-sample simulation of the model *without* feedback shows that following the fiscal expansion of 2008–09 the debt-GDP ratio diverges Along this path: (i) the debt-to-GDP ratio reaches 1.75 in 2020, (ii) an unsustainable fiscal policy cumulates yearly primary deficits in the range of 10–20 percent of GDP, (iii) the rapid increase in the debt ratio has no effect on interest rates—in effect, following the historical trend, the cost of debt service falls to zero, (iv) despite the divergence of the debt ratio, real growth converges rapidly toward its steady state value estimated at 4 percent. The results from the model with a debt feedback are very different. In the fiscal VAR *with*

Figure 6. The United States:Out-of Sample Simulated Paths with and Without a Debt Feedback, Starting from 2008–09 Macro Conditions: (a)Without Debt Feedback; (b)With Debt Feedback



feedback debt stabilization is achieved because the initial fiscal expansion, occurred in 2008–09, is eventually reversed, and the dynamics of the cost of financing switches form an increasing path to a converging one. The projected dynamics of the model with feedback reveals all the features of a sustainable debt dynamics: (i) the debt-to-GDP ratio converges quickly toward its steady state value, (ii) the primary deficit after its peak at 10 percent of GDP in 2009 is progressively reduced and turns into a surplus by 2014–20, (iii) interest rates respond positively to the fiscal expansion, but also to the inversion in the path of the deficit, and eventually converge toward a level between 2 and 3 percent, (iv) output growth converges to its steady state level of 4 percent.

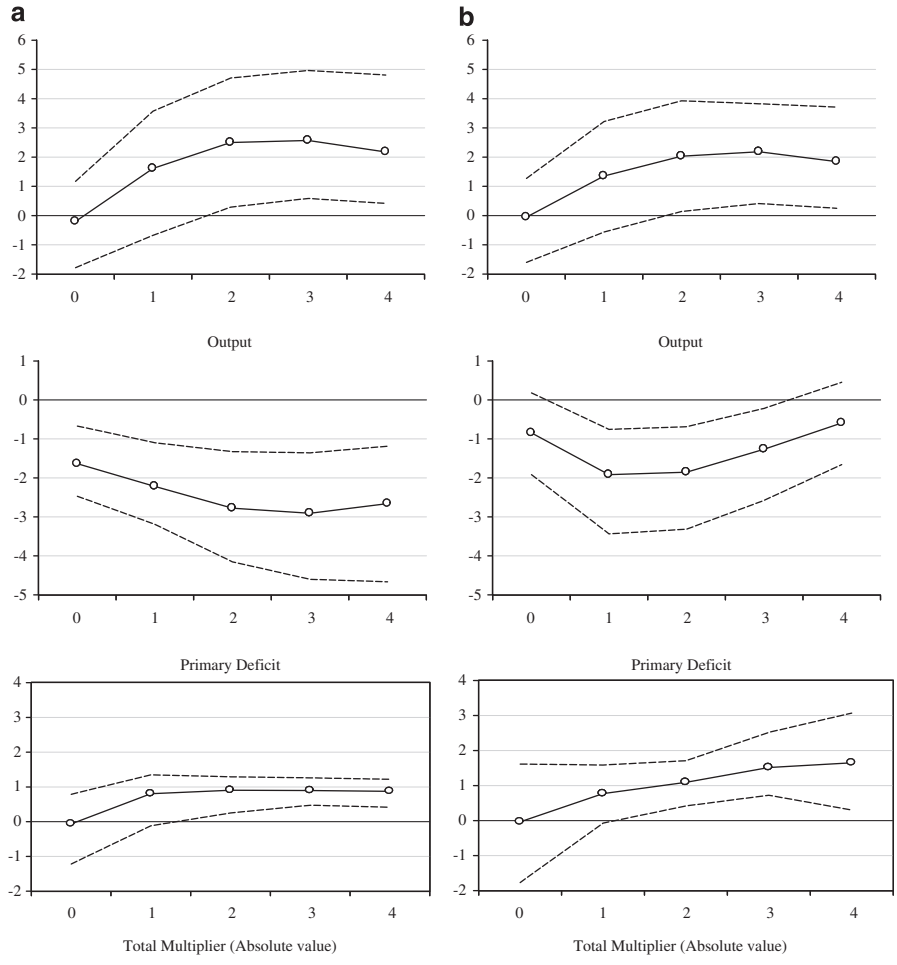
This evidence shows that impulses responses computed on the two models should be interpreted very differently. In the case of the model without feedback the initial shock lands the economy on an unsustainable fiscal path, while in the case of the model with feedback this does not happen.

To further elaborate on this point, for each of the two different specifications of Equation (11), with and without the debt feedback, we simulated the effect of a fiscal consolidation shock corresponding to 1 percent of GDP, respecting the historical policy style, that is the correlation between tax and spending shocks that characterizes the U.S. experience. In Figure 7 we show the responses of output and of the primary deficit.

Consider first the response of output to the fiscal adjustment under the two models: there is no difference between the two specifications, with and without the debt feedback. A clear difference, instead, emerges when we compare the effect of the fiscal adjustment on the primary deficit. In the model without feedback, the fiscal contraction has a permanent effect on the primary deficit. The deficit falls and then remains permanently negative. This explains what we observed in Figure 6 where the debt-to-GDP ratio landed on a diverging path. Instead, in the model with feedback, the effect of the initial shock on the primary deficit is eventually reversed, and the debt ratio converges toward its long run mean.

The lesson from Figure 7 is that fiscal multipliers cannot be inferred by simply analyzing the impulse response of output to a fiscal shock because the same impulse response can correspond to very different fiscal multipliers. In our case, in the model without feedback, an initial fiscal retrenchment of 1 percent of GDP determines, after five years, a total fiscal retrenchment of 11 percent of GDP. In the model with feedback the total fiscal retrenchment generated by the same initial shock is instead 8 percent of GDP. The same total effect on output—namely a marginally significant expansion of about 2 percent over a five-year period—is therefore obtained with a change in the deficit/GDP ratio which differs by 3 percent between the two simulated fiscal manoeuvres. The last panel of Figure 7 illustrates this point by reporting total fiscal multipliers, defined as the ratio of GDP change to the cumulative change in the primary deficit over the period. The total multiplier after four years reaches a value slightly less than one in the model without feedback while it becomes close to two in the model with feedback.

Figure 7. Impulse Responses to Aggregate Fiscal Consolidation Shocks, the United States Only: (a)Without Debt Feedback; (b)With Debt Feedback



To understand the difference between the two simulations it is important to realize that whenever a fiscal shock is introduced there is a relevant related question on the future "exit strategy" from the policy that is being implemented. In the model with feedback the "exit strategy" is embedded by construction in the simulation, while in a model without debt feedback the question on the exit strategy goes totally unanswered. Of course, as the debt-to-GDP ratio is a slow moving variable, the difference between the two simulations starts becoming relevant 2–3 years after the fiscal shocks hits the system. This difference, however, does eventually become sizable because the simulation of the effect of current fiscal policy taking into account the constraint it will impose on the future policy is different from an unconditional simulation of the effects of current fiscal policy.

Computing the Effects of Fiscal Policy Allowing for Heterogeneity

We now come to the central point of our paper. We estimate fiscal multipliers in a model that allows for a debt feedback, international comovements, and cross-country heterogeneity. We do this using the full model presented in Equation (2) to compute the effects of a global fiscal contraction of 1 percent of GDP obtained with a mix of tax increase and expenditure reduction that reflects, country by country, the historical pattern of fiscal policy. The model allows for different policy styles across countries, different debt-deficit dynamics, and different degrees of exposure to the international cycle. Table 1 illustrates the significance of the debt feedback by reporting the estimated coefficients on the debt in the fiscal reaction function of the different countries. Note that debt stabilization plays a role for most countries: the difference between the feedback coefficients implies a positive feedback of the primary surplus to the debt-to-GDP ratio for all countries with the only exception of Japan and the United Kingdom, where following an increase in the debt ratio, the primary deficit increases. The style of stabilization is however heterogenous across countries: Canada and the United States typically respond to an increase in debt cutting spending, while Italy and France rely on tax increases. Note however that the coefficients in Table 1 track only the response of the primary deficit to the debt. Whether this is enough to stabilize debt will also depend on the dynamics of the interest payments.

The output multipliers for the eight countries, reported in Figure 8, document a very high level of heterogeneity, suggesting that an aggregate homogeneous fiscal multiplier, such as the one reported in Figure 4, would be difficult to interpret. The fact that the impulse responses to fiscal shocks of some countries lie frequently outside of the confidence intervals of the same impulse responses for other countries is a statistical evidence of such heterogeneity.[10] The output response to a global fiscal retrenchment ranges from significantly contractionary in Belgium and France, to not significantly different from zero in the United Kingdom, and Italy, to initially zero and then slightly expansionary in Canada, and the United Sttaes, to significantly expansionary in Japan and in Sweden, at least on impact.

Interestingly, one of the countries that shows an expansionary effect of a fiscal policy retrenchment on output is Japan, suggesting that bolder fiscal stabilization could be successful.

IV. Robustness

In this section we check the robustness of our findings of heterogeneity in fiscal multipliers to two modifications of the fiscal shocks included in our

[10]This criterion allows to judge heterogeneity of impulse responses at all horizons and to assess the importance of deviations from the null of homogeneity. The evidence implies that both the impact effect and the dynamic effect of narrative shocks on output is statistically different across different countries.

Table 1. Coefficients of b_{t-1} (t-Stats in Parenthesis)

Countries	Expenditures g_t	Revenues τ_t
Belgium	−0.038 (−0.27)	0.030 (0.35)
Canada	−0.149 (−2.25)	−0.072 (0.27)
France	0.036 (0.57)	0.144 (1.56)
Italy	−0.110 (−2.12)	0.218 (3.18)
Japan	0.015 (0.10)	0.180 (1.52)
Sweden	−0.072581 (−0.64)	0.045513 (0.38)
United Kingdom	0.183 (2.02)	0.086 (0.47)
United States	−0.292 (−2.23)	0.47 (1.62)

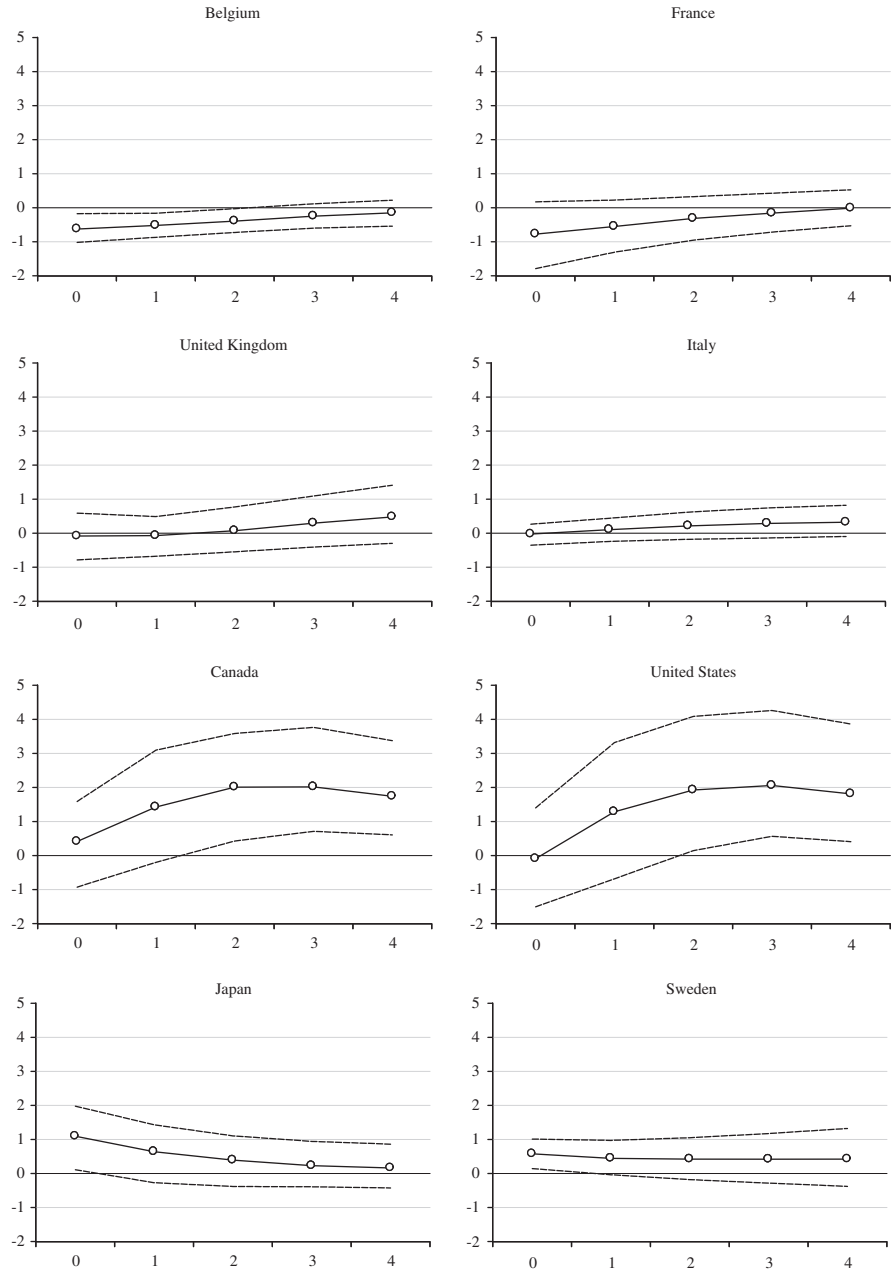
model. We first analyze the impact of local rather than international fiscal shocks, then we address explicitly the potential endogeneity of narrative fiscal shocks.

Country-Specific vs. Global Shocks

So far the type of fiscal policy experiment considered is a simultaneous fiscal consolidation across all countries in our sample. This is the type of shocks that is analyzed in Devries and others (2011) when computing the fiscal multiplier based on the pooled output growth equation and reported in Figure 4a. In the nonlinear GVAR specification there is no need for imposing the restriction that fiscal policy is synchronized across all countries as the extended specification of the model allows for heterogeneity across countries and for experimenting with local policy shocks. We have implemented such an experiment by running eight different simulations, one for each country in the sample, in which only one country at a time is subject to a fiscal shock while the others are inactive. We collected the impulse responses in Figure 9.

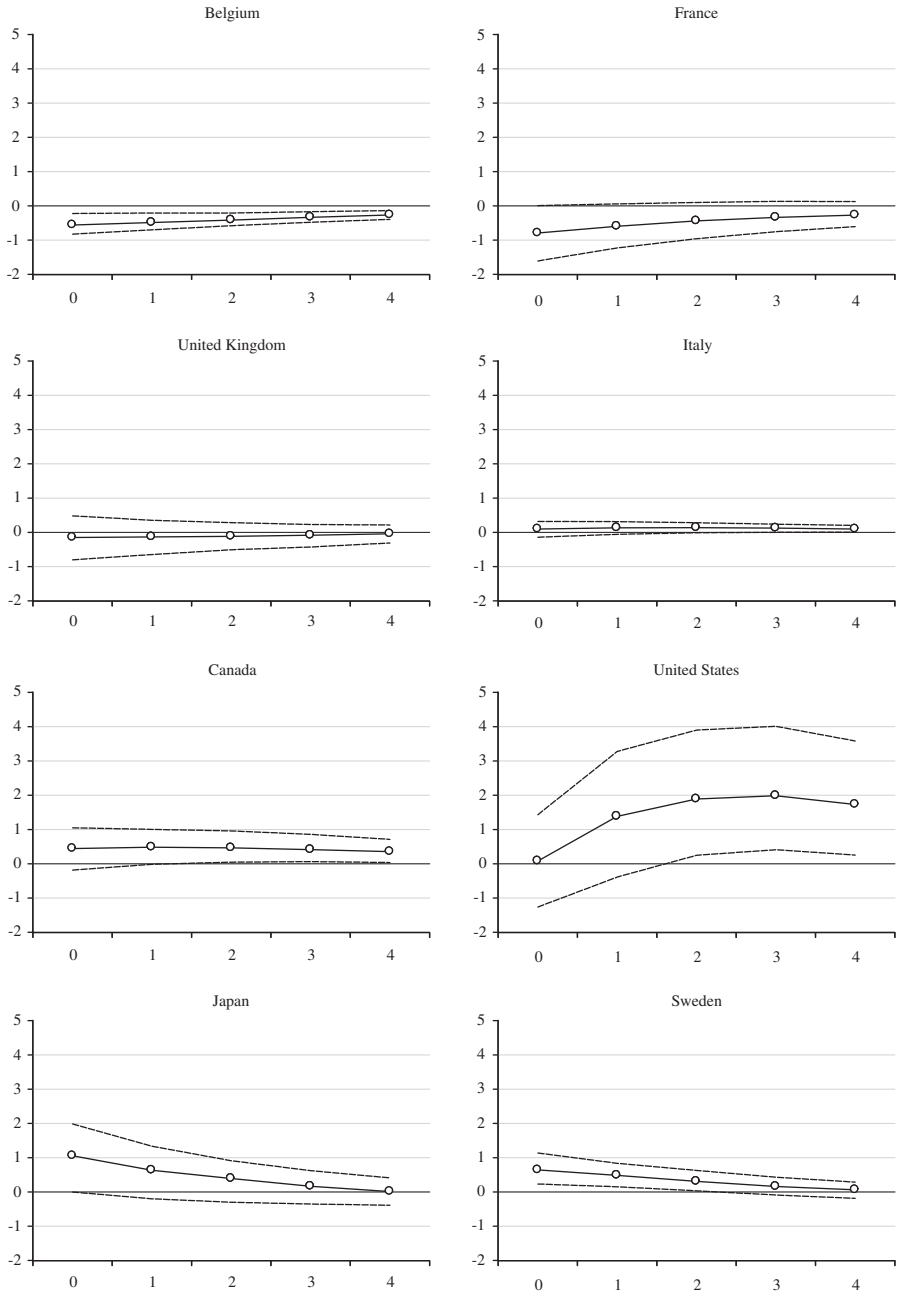
The results strengthen our general point on the importance of allowing for heterogeneity. Consider for example the similarity of the output response to a shift in fiscal policy in Canada and the United States that we reported in Figure 8 while analyzing the effect of global fiscal shocks. Such similarity totally disappears in Figure 9 where we consider the response to a local shock. The U.S. response is virtually unaltered while the response in Canada is practically flat at zero. This evidence shows that the exercise based on a synchronized shock was basically measuring the response of the Canadian economy to a U.S. policy shock rather than the domestic fiscal multiplier in Canada. In other words, while in the United States a fiscal consolidation is expansionary (in the medium run), the same shift in fiscal policy, if carried out only in Canada, has no effect on output in Canada. However when the United States cut spending and raise taxes, output rises not only in the United States but also in Canada. The same does not happen symmetrically for the United States when Canada consolidates, most likely because of the different size of the two economies. The identification of the two separate

Figure 8. Output Responses to Global (Simultaneously Implemented in All Countries) Fiscal Consolidation Shocks



effects requires the higher degree of heterogeneity of our nonlinear GVAR model. In general, although all responses to local shocks are smaller than those based on global shocks, Canada is the only case in which there is a dramatic difference between these two experiments.

Figure 9. Output Responses to Local (Implemented in Only One Country) Fiscal Consolidation Shocks



Are the Devries and others (2011) Narrative Shocks Exogenous?

Our second robustness check is based on a closer analysis of the exogeneity of the narrative shocks. Our specification strategy, based on the direct inclusion

of narrative shocks in the nonlinear VAR model, takes for granted the exogeneity of the tax hikes and the expenditure cuts constructed by Devries and others (2011). Such fiscal actions represent a response to past decisions and past economic conditions rather than to prospective conditions. As it is clearly stated by Devries and others, "they are unlikely to be systematically correlated with other developments affecting output in the short-term and are thus valid for estimating the macroeconomic effects of fiscal consolidations." It is possible, however, that these narrative shocks, rather than being truly exogenous, capture the systematic response of fiscal policy to the debt dynamics. To fix ideas consider this simplified version of a fiscal reaction function which decomposes the fiscal reaction function into three components: an *automatic* output stabilization component, ($\beta_1 y_t$), a *discretionary* output stabilization component, ($\beta_2 y_{t-1}$), and a *discretionary* debt stabilization component, ($\beta_3 d_{t-1}$). u_t^{fp} denotes a shock to the fiscal reaction function:

$$(fiscal\,reaction\,function)_t = \beta_0 + \beta_1 y_t + \beta_2 y_{t-1} + \beta_3 d_{t-1} + u_t^{fp}.$$

The narrative approach filters out of the fiscal reaction function the automatic and the discretionary responses to fluctuations in output (β_1 and β_2), but shifts in fiscal policy motivated by a rising debt level may reflect the true fiscal shock u_t^{fp} as well as the discretionary debt stabilization component (β_3). In other words, the narrative approach might lead to identify as fiscal shocks ($fp_t - \beta_0 - \beta_1 y_t - \beta_2 y_{t-1}$) rather than u_t^{fp}, generating a potential endogeneity bias in our GVAR specification augmented with debt dynamics.[11]

To address this potential source of endogeneity we have first regressed the narrative shocks on lags of output only, and then on the full information set used in our nonlinear GVAR. The results reported in Table 2 show that indeed the shocks are orthogonal to lags of output, but that they are not always orthogonal to the full information set included in the nonlinear GVAR model. To evaluate the consequences of this fact we have re-run our model by keeping the original narrative shocks for the countries where they appear to be orthogonal to the full information set, and by substituting to the original narrative shocks the residuals of the regression of the narrative shocks on the relevant information set for the countries where the orthogonality hypothesis has been clearly rejected. The results are reported in Figure 10 that is directly comparable with Figure 9 because we consider the output responses to local shocks. Figure 10 shows that our finding of an heterogenous response of output to fiscal shocks is robust to the orthogonalization of the shocks, although for some of the countries, *the*

[11]Note that this problem would not arise in the case of shocks identified in a VAR, because the impact of shocks to the policy reaction function implicitly estimated in a VAR is implemented by keeping all estimated parameters constant. Therefore the Lucas' critique (Lucas, 1976) does not apply despite the nature of a reduced form model of any VAR. This argument, however, cannot be applied to the narrative identified shocks inserted in a VAR if they are not orthogonal to the relevant information set.

Table 2. Exogeneity Test for IMF Shocks

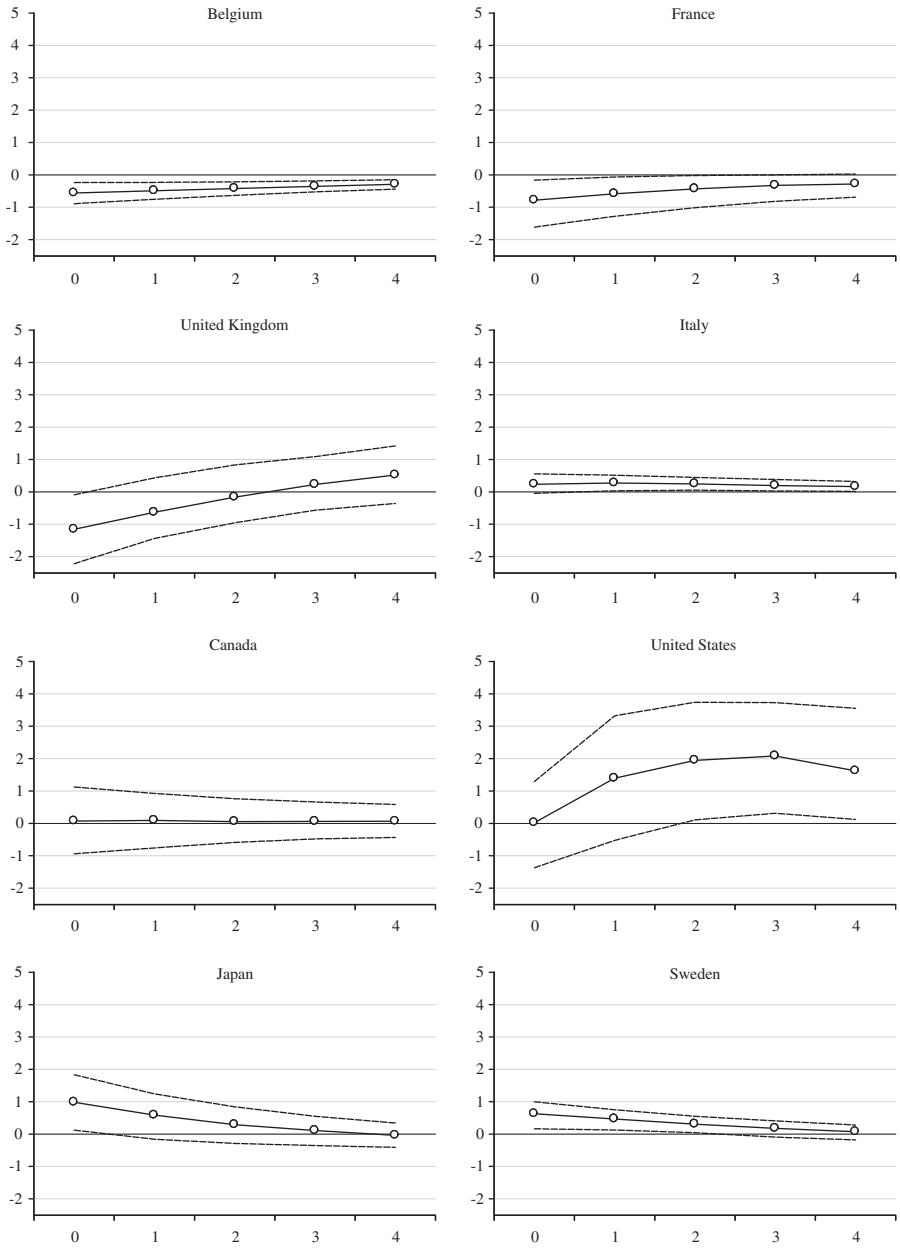
	Belgium	Canada	France	Italy	Japan	Sweden	United Kingdom	United States
Exogeneity of spending shocks $\varepsilon^{g}_{i,t}$								
$\mathcal{I}_{i,t} = \{y_{i,t-1}, y_{i,t-2}\}$								
F-Stat	4.41	0.15	0.14	1.71	0.24	0.12	4.17	0.04
F-Prob	0.021	0.85	0.86	0.19	0.78	0.88	0.025	0.95
$\mathcal{I}_{i,t} = \{X_{i,t-1}, X_{i,t-2}, b_{i,t-1}, b_{i,t-2}, y^{*}_{i,t}, y^{*}_{i,t-1}, s^{*}_{i,t}, s^{*}_{i,t-1}\}$								
F-Stat	0.98	4.03	1.22	5.48	1.00	3.14	4.93	2.29
F-Prob	0.51	0.015	0.35	0.001	0.49	0.029	0.006	0.06
Exogeneity of tax shocks $\varepsilon^{\tau}_{i,t}$								
$\mathcal{I}_{i,t} = \{y_{i,t-1}, y_{i,t-2}\}$								
F-Stat	1.78	3.54	0.11	0.70	0.01	0.23	4.10	1.68
F-Prob	0.18	0.04	0.89	0.50	0.98	0.79	0.026	0.20
$\mathcal{I}_{i,t} = \{X_{i,t-1}, X_{i,t-2}, b_{i,t-1}, b_{i,t-2}, y^{*}_{i,t}, y^{*}_{i,t-1}, s^{*}_{i,t}, s^{*}_{i,t-1}\}$								
F-Stat	2.77	10.4	1.94	1.30	0.67	3.40	4.34	1.17
F-Prob	0.03	0.0003	0.11	0.31	0.77	0.022	0.004	0.38

shape of the impulse responses is somewhat affected (although *the impact effect* is not).

V. Conclusions

The main conclusion of our empirical analysis is that the question "what is the fiscal policy multiplier" asked unconditionally is impossible to answer empirically and makes little sense theoretically. There is no unconditional fiscal policy multiplier. The effect of fiscal policy on output is different according to the different debt dynamics, the different degree of openness, and the different fiscal reaction functions in different countries. Pooling together the evidence for different countries to derive a single measure of the effect of fiscal retrenchments on output is therefore of very little use to describe the effect of exogenous shifts in fiscal policy on output. In this paper we have derived empirical evidence on fiscal multipliers by specifying a nonlinear VAR that allows for the three sources of heterogeneity mentioned above. Our model generates fiscal multipliers computed along paths that are

Figure 10. Output Responses to Local Fiscal Consolidation Shocks Orthogonalized to $\mathcal{I}_{t-1}$



not inconsistent with a positive value of government bonds, interestingly with the exception of Japan. We find that: (i) international spillovers are important. In the case of Canada, for instance, it makes a big difference whether a fiscal consolidation happens contemporaneously also in the United

States, or only in Canada; (ii) there is significant heterogeneity in the response of output to fiscal consolidations across countries, ranging from expansionary to contractionary. As a consequence there are many fiscal multipliers and an average, unconditional, multiplier does not exist.

We close with a note of caution: our results should not be used to answer policy questions such as "How should a government respond to a particular macro shock?" These questions need to be addressed within the framework of quantitative general equilibrium models of the business cycle—that is, within the context of a theoretical macro model rather than on an empirical reduced form econometric model. Empirical results like those presented in this paper are however useful in the specification of a DSGE model relevant for policy simulation analysis.

REFERENCES

Alesina, A, and S. Ardagna, 2010, "Large Changes in Fiscal Policy: Taxes Versus Spending," NBER Chapters, in *Tax Policy and the Economy*, Vol. 24, pp. 35–68 (National Bureau of Economic Research, Inc.).

Auerbach, A, and Y. Gorodnichenko, 2011, "Measuring the Output Response to Fiscal Policy," *American Economic Journal: Economic Policy*.

Bachmann, R., and E.R. Sims, 2011, "Confidence and the Transmission of Government Spending Shocks," NBER Working Paper 17063.

Bohn, Henning, 1998, "The Behaviour of U.S. Public Debt and Deficits," *Quarterly Journal of Economics*, Vol. 113, pp. 949–63.

Cloyne, J., 2011, "What are the Macroeconomic Effects of Tax Changes in the United Kingdom," mimeo, University College London.

Corsetti, G., A. Meier, and G. Mueller, 2009, "Fiscal Stimulus with Spending Reversals," CEPR DP 7302.

Dees, S., F. di Mauro, M.H. Pesaran, and L.V. Smith, 2007, "Exploring the International Linkages of the Euro Area: A Global VAR Analysis," *Journal of Applied Econometrics*, Vol. 22, No. 1, pp. 1–38.

Devries, P., J. Guajardo, D. Leigh, and A. Pescatori, 2011, "A New Action-Based Dataset of Fiscal Consolidation," IMF Working Paper No. 11/128. The data set is available on the IMF's website (www.imf.org/external/pubs/cat/longres.aspx?sk= 24892.0).

Favero, C., and F. Giavazzi, 2007, "Debt and the Effects of Fiscal Policy," NBER Working Paper No. 12822, April.

Ghosh, A., others 2011, "Fiscal Fatigue, Fiscal Space and Debt Sustainability in Advanced Economies," NBER Working Paper 16782.

Hansen, L.P., and T.J. Sargent, eds., 1991, "Two Difficulties in Interpreting Vector Autoregressions," in *Rational Expectations Econometrics*, (Boulder, CO: Westview Press), pp. 77–19.

Ilzetzki, E., 2011, "Fiscal Policy and Debt Dynamics in Developing Countries," mimeo, London School of Economics.

Ilzetzki, E., E.G. Mendoza, and C. Vegh, 2011, "How Big (Small) are Fiscal Multipliers?," paper presented at the IMF-EUI Conference on Fiscal Policy, Stabilization and Sustainability, Florence June 6–7, NBER Working Paper 16479.

International Monetary Fund (IMF), 2010, "Will it Hurt? Macroeconomic Effects of Fiscal Consolidation," *World Economic Outlook*, Chapter 3, October).

Killian, L., and R. Vigfusson, 2010, *Are the Responses of the U.S. Economy Asymmetric in Energy Prices Increases and Decreases?* Available via the Internet: www-personal .umich.edu/∼lkilian/kv032310.pdf.

Kilian, Lutz, and Robert J. Vigfusson, 2011, "Nonlinearities in the Oil Price-Output Relationship," CEPR Discussion Papers 8174, C.E.P.R.

Leeper, E.M., 2010, "Monetary Science, Fiscal Alchemy," Federal Reserve Bank of Kansas City's Jackson Hole Symposium *Macroeconomic Policy: Post-Crisis and Risks Ahead*, August 26–28.

Leeper, E.M., T. Walker, and S.C. Yang, 2008, "Fiscal Foresight: Analytics and Econometrics," mimeo.

Lucas, R.J., 1976, "Econometric Policy Evaluation. A Critique," in *Carnegie-Rochester Conference Conference Series on Public Policy*, ed. by K. Brunner and A. Meltzer (Elsevier) Vol. 1, No. 1, pp. 19–46.

Mertens, K., and M.O. Ravn, 2011, "Measuring Fiscal Shocks in Structural VARs using Narrative Data," mimeo.

Perotti, R., 2011, "Expectations and Fiscal Policy. An Empirical Investigation," mimeo, IGIER, Bocconi University. Available via the Internet: www.igier.unibocconi.it/perotti.

Pesaran, M.H., T. Schuermann, and S.M. Weiner, 2004, "Modelling Regional Interdependencies Using a Global Error-Correcting Macro-Econometric Model," *Journal of Business and Economic Statistics*, Vol. 22, No. 2, pp. 129–62.

Ramey, Valerie A., 2011, "Identifying Goverment Spending Shocks: It's all in the Timing," *The Quarterly Journal of Economics*, Oxford University Press, Vol. 126, No. 1, pp. 1–50.

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

Zubairy, S., 2011, "On Fiscal Multipliers: Estimates from a Medium Scale DSGE Model," Banque of Canada Working Paper 2010–30.