DISCUSSION PAPER SERIES

No. 7930

A MODEL-BASED EVALUATION OF THE DEBATE ON THE SIZE OF THE TAX MULTIPLIER

Ryan Chahrour, Stephanie Schmitt-Grohé
and Martín Uribe

INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

www.cepr.org

A MODEL-BASED EVALUATION OF THE DEBATE ON THE SIZE OF THE TAX MULTIPLIER

Ryan Chahrour, Columbia University
Stephanie Schmitt-Grohé, Columbia University, NBER and CEPR
Martín Uribe, Columbia University and NBER

ABSTRACT

A Model-Based Evaluation of the Debate on the Size of the Tax Multiplier*

The SVAR and narrative approaches to estimating tax multipliers deliver significantly different results. The former yields multipliers of about 1 percent, whereas the latter produces much larger multipliers of about 3 percent. The SVAR and narrative approaches differ along two important dimensions: the identification scheme and the reduced-form transmission mechanism. This paper uses a DSGE-model approach to evaluate the hypothesis that the different tax multipliers stemming from the SVAR and narrative approaches are due to differences in the assumed reduced-form transmission mechanisms. The main finding of the paper is that in the context of the DSGE model employed this hypothesis is rejected. Instead, the observed differences in estimated multipliers are due either to both models failing to identify the same tax shock, or to small-sample uncertainty.

Ryan Chahrour
Department of Economics
Columbia University
420 West 118th Street
New York NY 10027
USA

Email: rc2374@columbia.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=172322

Stephanie Schmitt-Grohé
Department of Economics
Columbia University
420 West 118th Street
New York NY 10027
USA

Email: ss3501@columbia.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=145223

Martín Uribe
Department of Economics
Columbia University
International Affairs Building
New York, NY 10027
USA

Email: martin.uribe@columbia.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=140752

1   Introduction

The recent empirical literature has delivered two main approaches to estimating tax multipliers. The first one, pioneered by the work of Blanchard and Perotti (2002), is based on a structural vector autoregression (SVAR) analysis. The second strand of the empirical literature, initiated by the work of Romer and Romer (2007), estimates tax shocks using a narrative approach and then regresses a measure of aggregate activity, such as GDP, on current and lagged values of the identified tax shock.

The motivation of this paper is that the SVAR and narrative approaches deliver significantly different estimates of the size of the tax multiplier. The former approach yields relatively small values of less than 1 percent, whereas the latter delivers large values of about 3 percent. That is, a cut in taxes equivalent to 1 percent of output generates an increase in output of less than 1 percent according to the SVAR approach but a 3 percent increase in output according to the narrative approach.

The starting point of our analysis is the observation that the two approaches differ along two dimensions. One dimension is the assumed reduced-form transmission mechanism. The transmission mechanism invoked by the SVAR approach consists of a multi-equation, multivariate autoregressive system in which taxes evolve jointly with other endogenous variables. By contrast, the transmission mechanism proposed by the narrative approach involves a single equation expressing output as a linear function of current and past values of the exogenous tax shock.

The second dimension along which the SVAR and narrative approaches differ is, of course, the identification scheme. The SVAR approach imposes a number of restrictions to identify the variance-covariance matrix of the vector of fundamental shocks (one of which is the tax shock), given information on the variance-covariance matrix of the vector of estimated reduced form residuals. By contrast, the identification scheme in Romer and Romer (2007) uses a narrative approach consisting in analyzing written historical records, including presidential speeches, executive-branch documents, and congressional reports, to identify exogenous

changes in tax liabilities.

A natural question that emerges from the above analysis is whether the significant differences in the size of tax multipliers stemming from the Blanchard-Perotti and Romer-Romer empirical models are due to differences in their transmission mechanisms or to fundamental differences in the tax shocks they identify. The goal of the present investigation is to evaluate the hypothesis that the differences in tax multipliers are due to the different transmission mechanisms, taking as given the ability of both models to identify exogenous tax shocks. To this end, we build an optimizing dynamic stochastic general equilibrium (DSGE) model featuring a number of exogenous shocks and real rigidities that have been shown to be important for fitting the U.S. postwar business cycle. We use the DSGE model as our data-generating process to estimate the Blanchard-Perotti and Romer-Romer empirical models under the assumption that the econometrician successfully identifies the structural tax shocks. Fulfillment of this assumption is impossible to guarantee in empirical studies, but trivial to satisfy in our theoretical environment. This, in fact, is our main methodological contribution to the fiscal-multiplier debate.

Our main finding is that the hypothesis posited above is rejected within our data generating process. Conditional on correctly identifying the exogenous tax shock, the Blanchard-Perotti and Romer-Romer models deliver on average remarkably good approximations to the true impulse response of output to an exogenous innovation in factor income tax rates. Consequently, both models also deliver average tax multipliers that are in line with the 'true' one—i.e., the one implied by the DSGE model. This finding suggests that the sharp difference in the size of the tax multiplier implied by the Blanchard-Perotti and Romer-Romer models when estimated on actual data may be due to small sample uncertainty or to the fact that their associated identification schemes uncover fundamentally different fiscal shocks or both. We explore the role of small sample uncertainty conditional on the correct identification of the underlying tax shock in the context of our data generating process. We find that for samples of size similar to the length of the postwar period, small sample errors

are significant. In fact, short sample uncertainty can explain the totality of the observed differences in estimated tax multipliers.

The remainder of the paper is organized in six sections. Section 2 develops the DSGE model that serves as our data generating process. Section 3 contains the main result of the paper. It estimates the Blanchard-Perotti and Romer-Romer models using artificial data and compares the resulting tax multiplier to the true one stemming from the DSGE model. Section 4 studies the effects of introducing anticipated shocks in the DSGE model on the ability of the Blanchard-Perotti and Romer-Romer models to uncover the true tax multiplier. Section 5 analyzes the consequences of finite samples on the variance of the estimated tax multipliers. Section 6 studies the hybrid reduced-form model of Favero and Giavazzi (2010), which combines elements of the VAR and narrative approaches. Section 7 concludes.

2 The Data Generating Process

The DSGE model that we use as our data generating process is an augmented version of the one proposed by Mertens and Ravn (2009). The main difference with the Mertens-Ravn model is that our framework includes a number of additional structural shocks customarily used in the quantitative business-cycle literature. Specifically, our model is driven by four shocks: income-tax shocks, government spending shocks, neutral productivity shocks, and preference shocks. The model distinguishes between durable and nondurable consumption and features four real rigidities: habit formation, adjustment costs in investment, adjustment cost in durable consumption, and variable capacity utilization. This class of model has been shown in several recent studies to fit well the postwar U.S. business cycle along a number of dimensions, including output, consumption, investment, hours worked, and tax revenues (see Mertens and Ravn, 2009; and Schmitt-Grohé and Uribe, 2010a, 2010b).

The economy is populated by a large number of identical households that seek to maxi-

mize the lifetime utility function

$$E_0 \sum_{t=0}^{\infty} \beta^t \mu_t \left[\frac{X_t^{1-\sigma} - 1}{1 - \sigma} - \frac{\omega n_t^{1+\kappa}}{1 + \kappa} Z_t^{1-\sigma} \right]$$

subject to the following sequential budget constraints

$$X_t = C_t^\nu V_t^{1-\nu} - b C_{t-1}^\nu V_{t-1}^{1-\nu},$$

$$V_{t+1} = (1 - \delta_v) V_t + D_t \left[1 - \Phi_v \left(\frac{D_t}{D_{t-1}} \right) \right],$$

$$K_{t+1} = [1 - \delta_k - \Psi_k(u_t)] K_t + I_t \left[1 - \Phi_k \left(\frac{I_t}{I_{t-1}} \right) \right],$$

$$C_t + I_t + D_t + G_t = W_t n_t (1 - \tau_t^n) + r_t u_t K_t (1 - \tau_t^k) + \tau_t^k \delta_\tau K_{\tau t} + F_t,$$

and

$$K_{\tau t+1} = (1 - \delta_\tau) K_{\tau t} + I_t,$$

where X_t is a composite good made of nondurable consumption and services derived from a stock of durable consumption goods, C_t denotes consumption of nondurables, V_t denotes the stock of durables, D_t denotes purchases of durable goods, n_t denotes hours worked, Z_t is a deterministic log-linear trend growing at the (gross) rate γ_z, μ_t denotes a stochastic preference shock, K_t is the capital stock, I_t denotes gross investment, u_t denotes capital capacity utilization, W_t denotes the real wage rate, r_t denotes the rental rate of capital, G_t denotes government spending, τ_t^n and τ_t^k denote, respectively, labor and capital income tax rates, F_t are lump-sum transfers received from the government, and $K_{\tau t}$ denotes a measure of the capital stock used by the fiscal authority to calculate the depreciation allowance. Note that the depreciation rate used for tax purposes may not equal the economic rate of depreciation ($\delta_\tau \neq \delta_k$).

Firms purchase labor and capital services to produce a single perishable good, Y_t, by

4

means of the production technology

$$Y_t = a_t(u_t K_t)^\alpha (Z_t n_t)^{1-\alpha}.$$

Firms are assumed to be perfectly competitive in product and factor markets. They choose input quantities to maximize profits, given by $Y_t - W_t n_t - r_t u_t K_t$, subject to the production technology given above.

The government is assumed to hold no debts or assets at any time. Lump-sum transfers are assumed to adjust endogenously each period to ensure budget balance. The government budget constraint is therefore given by

$$G_t + F_t = \tau_t^n W_t n_t + \tau_t^k (r_t u_t K_t - \delta_\tau K_{\tau t}).$$

The laws of motion of the factor income tax rates are assumed to be of the form

$$\tau_t^n - \tau^n = \rho_1^n(\tau_{t-1}^n - \tau^n) + \rho_2^n(\tau_{t-2}^n - \tau^n) + \epsilon_t^\tau$$

and

$$\tau_t^k - \tau^k = \rho_1^k(\tau_{t-1}^k - \tau^k) + \rho_2^k(\tau_{t-2}^k - \tau^k) + \epsilon_t^\tau,$$

where ϵ_t^τ is an i.i.d. innovation with mean zero and standard deviation σ_τ, and τ^n and τ^k denote, respectively, the deterministic-steady-state values of τ_t^n and τ_t^k. Note that the innovation ϵ_t^τ is common to both processes.

The laws of motion of the remaining three driving forces are as follows:

$$\ln \mu_t = \rho_\mu \ln \mu_{t-1} + \epsilon_t^\mu$$

$$\ln(g_t/g) = \rho_g \ln(g_{t-1}/g) + \epsilon_t^g,$$

and

$$\ln a_t = \rho_a \ln a_{t-1} + \epsilon_t^a,$$

where $g_t \equiv G_t/Z_t$ denotes the detrended level of government spending, g denotes the deterministic-steady-state of g_t, and $\rho_\mu, \rho_g, \rho_a \in (0,1)$ are parameters. The disturbances ϵ_t^μ, ϵ_t^g and ϵ_t^a are i.i.d. with mean zero and standard deviations σ_μ, σ_g, and σ_a, respectively.

The time unit in the model is meant to be one quarter. The calibration of the model follows closely Mertens and Ravn (2009), who estimate the structural parameters of the model to match observed impulse responses of a number of macroeconomic variables to tax shocks. For more details regarding the parameterization of the model, we refer the reader to the work of Mertens and Ravn (2009). Our model requires the calibration of a number of parameters that are not present in the Mertens-Ravn model, namely, the parameters defining the stochastic processes of preference shocks, government spending shocks, and productivity shocks, and the volatility of tax shocks. We set the serial correlations of preference and government spending shocks at 0.9 and the serial correlation of the technology shock at 0.95, which are values in the range used in business-cycle analysis. All of the results of this paper are invariant to proportional changes in the standard deviations of all shocks. We therefore arbitrarily normalize the standard deviation of the technology shock at one percent and set the standard deviations of the remaining three shocks to ensure that the share of the variance of output explained by tax shocks, government spending shocks, productivity shocks, and preference shocks be, respectively, 20, 10, 35, and 35 percent. Table 1 summarizes the calibration of the model.

Table 1: Calibration

Parameter	Value	Memo
α	0.36	technology parameter
γ_z	1.005	gross growth rate
$\beta\gamma_z^{1-\sigma}$	$1.03^{-1/4}$	growth-adjusted discount factor
δ_k	0.025	depreciation rate of capital
δ_v	0.025	depreciation rate of durables
$\Psi'_k(1)$	0.0375	marginal cost of capacity utilization
ν	0.8557	preference parameter
ω	249.35	preference parameter
s_g	0.201	share of government spending in GDP
δ_τ	0.05	fiscal depreciation rate
τ_k	0.42	capital income tax rate
τ_n	0.26	labor income tax rate
σ	2.572	preference parameter
b	0.822	habit formation
κ	0.355	preference parameter
$\Phi''_v(\gamma_z)$	4.444	adjustment costs durables
$\Phi''_k(\gamma_z)$	6.581	adjustment costs investment
$\Psi''_k(1)/\Psi'_k(1)$	0.367	capacity utilization
ρ_1^n	0.999	labor-income tax rate
ρ_2^n	0	labor-income tax rate
ρ_1^k	1.629	capital income tax rate
ρ_2^k	-0.652	capital income tax rate
ρ_μ	0.9	preference shock
ρ_g	0.9	government spending shock
ρ_a	0.95	technology shock
σ_τ	0.0031	std. dev. of tax shock
σ_μ	0.0308	std. dev. of preference shock
σ_g	0.0565	std. dev. of government spending shock
σ_a	0.01	std. dev. of technology shock

3 A Model-Based Test of the Transmission-Mechanism Hypothesis

We wish to evaluate the hypothesis that, assuming the correct identification of tax shocks, the Blanchard-Perotti and Romer-Romer models identify different transmission mechanism of tax shocks on output. Using the DSGE model of the previous section as the data generating process, we estimate the transmission mechanisms associated with the Blanchard-Perotti and Romer-Romer models. In so doing, we use our knowledge of the shocks driving the model economy to leave completely aside the issue of identification.

3.1 Estimating the Blanchard-Perotti Transmission Mechanism on Artificial Data

To estimate the Blanchard-Perotti reduced-form transmission mechanism, we draw a sample of 1,000 quarters of the four disturbances of the model to produce time series for $\hat{y}_t \equiv \ln(y_t/y)$, $\hat{\tau}_t \equiv \ln(\tau_t/\tau)$, and $\hat{g}_t \equiv \ln(g_t/g)$, denoting, respectively, the log-deviations from steady state of detrended output, $y_t \equiv Y_t/Z_t$, detrended tax revenues, $\tau_t \equiv T_t/Z_t$, and detrended government spending, $g_t \equiv G_t/Z_t$. Tax revenues are given by $T_t \equiv \tau_t^n W_t n_t + \tau_t^k(r_t u_t K_t - \delta_\tau K_{\tau t})$. The parameters y, τ, and g denote the steady-state values of y_t, τ_t, and g_t. We keep only the last 250 observations of each artificial time series, which roughly corresponds to the length of the postwar period, and discard the initial 750 observations. Our choice of variables used for estimation is guided by the observation that in their empirical model Blanchard and Perotti (2002) include output, tax revenues, and government spending.

The next step in our simulation exercise is to estimate the VAR system

$$X_t = \sum_{i=1}^{4} A_i X_{t-i} + u_t,$$

where

$$X_t \equiv \begin{bmatrix} \hat{y}_t \\ \hat{\tau}_t \\ \hat{g}_t \end{bmatrix}.$$

Following Blanchard and Perotti (2002), we posit that the reduced form shock u_t is related to a vector of 'fundamental' shocks, ϵ_t, as

$$u_t = B\epsilon_t,$$

where

$$\epsilon_t \equiv \begin{bmatrix} \epsilon_t^1 \\ \epsilon_t^2 \\ \epsilon_t^3 \end{bmatrix}.$$

We note that because the DSGE model features four structural innovations and the size of the VAR is three, the elements of ϵ_t cannot be interpreted as structural. Nevertheless, as stated earlier, our exercise consists in taking for granted the identification of tax shocks and examines instead the ability of the estimated VAR system to propagate that type of shock. To this end, we identify ϵ_t^1 with the tax shock and set the first column of the matrix B equal to the impact effect of a unit increase in ϵ_t^τ on the vector X_t in the DSGE model. In other words, the restriction we impose on the VAR system imply that the impulse responses of output, tax revenues, and government spending implied by the DSGE and VAR models are identical in the initial period. In subsequent periods, the true and estimated impulse responses will in general be different.

We replicate this exercise 1,000 times and report the average impulse responses of tax revenues and output to an innovation in ϵ_t^τ that raises tax revenues by one percent of steady-state output on impact.

Figure 1 displays with blue solid lines the impulse responses of tax revenue and output to a tax innovation implied by the DSGE model. The figure shows with red broken lines

Figure 1: Impulse Response to a Tax Innovation in the DSGE and Blanchard-Perotti Models
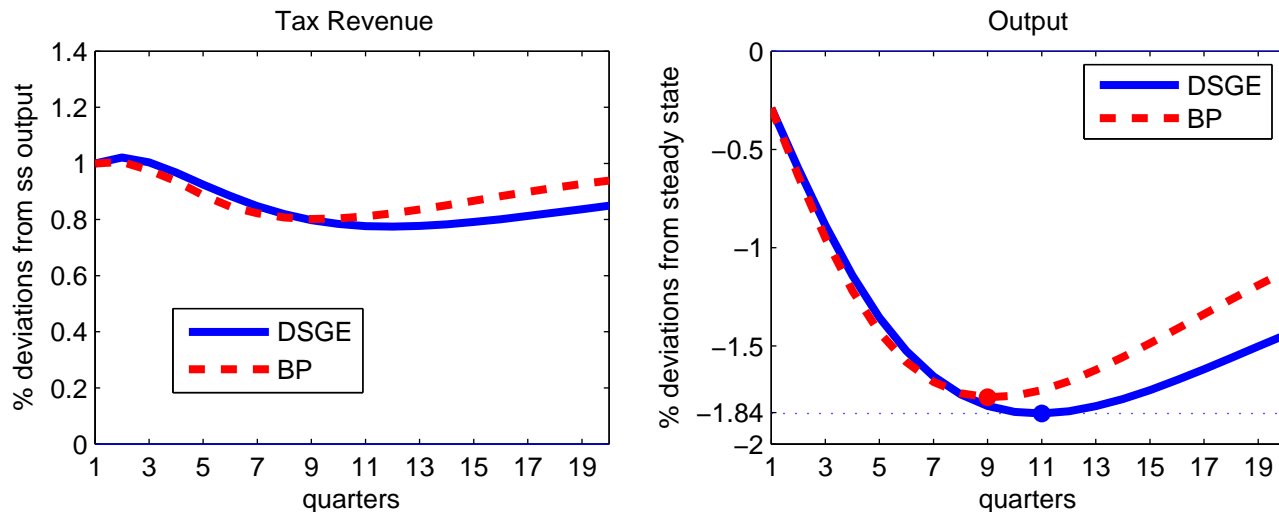
the corresponding mean impulse response functions implied by the Blanchard-Perotti model estimated using artificial data. By construction, the impulse responses implied by the DSGE and the Blanchard-Perotti models are identical on impact. The Blanchard-Perotti empirical model does a good job at tracing the propagation of the tax shock on output. Both in the DSGE and in the Blanchard-Perotti models output displays an inverted hump-shaped contraction. The trough in output is slightly less pronounced and occurs about a two quarters earlier (quarter 9 versus quarter 11) in the Blanchard-Perotti model. We measure the tax multiplier as the maximum drop in output divided by the initial increase in taxes. In the figure, the response of tax revenue is measured as a percentage of steady-state output, and the response of output is measured in percent deviations from its steady-state value. This means that the tax multiplier coincides with the magnitude of the trough in the output response. The tax multiplier predicted by the both the DSGE and the Blanchard-Perotti models is about 1.8 percent.

We conclude that, conditional on the correct identification of the tax shock, the Blanchard-Perotti reduced-form model does a good job at replicating the propagation of such shocks.

3.2 Estimating the Romer-Romer Transmission Mechanism on Artificial Data

Following the same logic as in the previous subsection, we continue to assume that the econometrician has correctly identified exogenous innovations in taxes. We therefore use the artificial time series generated from the DSGE model (and used in the previous section) to estimate the Romer-Romer model. That is, we estimate the single equation

$$\hat{y}_t = \sum_{i=0}^{12} C_i \epsilon_{t-i}^\tau + u_t, \tag{1}$$

where both $\hat{y}_t$ and ϵ_t^τ are equilibrium realizations of output and tax disturbances generated from the DSGE model and u_t is the residual of the regression.

Before proceeding to the estimation of the Romer-Romer equation on artificial data, we wish to show that, conditional on the correct identification of tax innovations, the OLS estimates of the parameters C_i for $i = 0, \ldots, 12$, deliver unbiased estimates of the impact and subsequent 12 points of the true impulse response function of output to a tax innovation. To see this, recall that the DSGE model admits a MA(∞) representation of output of the form

$$\hat{y}_t = \sum_{i=0}^{\infty} a_i \epsilon_{t-i}^\tau + \sum_{i=0}^{\infty} b_i \epsilon_{t-i}^a + \sum_{i=0}^{\infty} c_i \epsilon_{t-i}^\mu + \sum_{i=0}^{\infty} d_i \epsilon_{t-i}^g.$$

All of the innovations on the right-hand side of this expression are independent of one another. Comparing this expression with the Romer-Romer model, equation (1), we have that
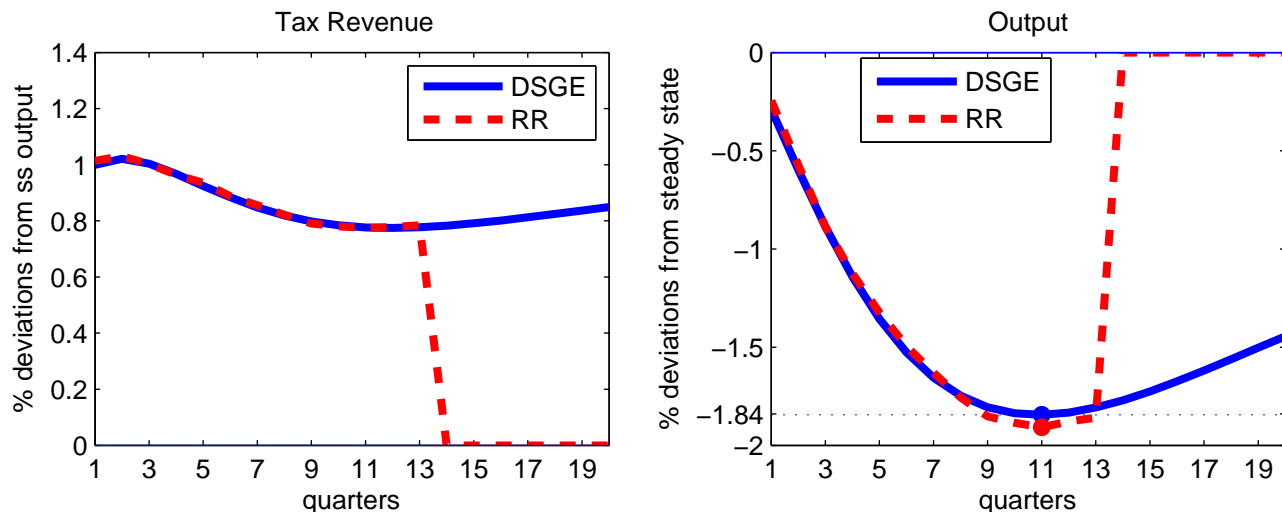
$$C_i = a_i, \quad i = 0, 1, \ldots, 12$$

and

$$u_t = \sum_{i=13}^{\infty} a_i \epsilon_{t-i}^\tau + \sum_{i=0}^{\infty} b_i \epsilon_{t-i}^a + \sum_{i=0}^{\infty} c_i \epsilon_{t-i}^\mu + \sum_{i=0}^{\infty} d_i \epsilon_{t-i}^g.$$

It follows that u_t is uncorrelated with ϵ_{t-i}^τ for $i = 0, 1, \ldots, 12$. As a result, the OLS estimate of C_i in the Romer-Romer model of equation (1) is unbiased for $i = 0, 1, \ldots, 12$. Because in

Figure 2: Impulse Response to a Tax Innovation in the DSGE and Romer-Romer Models


our DSGE model the trough of the impulse response function of output to a tax innovation occurs before period 12, it follows that the estimates of the Romer-Romer reduced-form transmission mechanism of equation (1) deliver unbiased estimates of the tax multiplier. The OLS estimate of C_i is, however, not efficient because the vector u_t is serially correlated. We explore this issue further in section 5.

Figure 2 displays the responses of tax revenues and output to an innovation in taxes in the DSGE and Romer-Romer models. The Romer-Romer impulse responses correspond to the average of 1,000 estimates of equation (1). The Romer-Romer reduced-form model captures almost perfectly the true impulse response functions. In particular, it replicates the true tax multiplier of 1.8 in period 11.

To preserve comparison with the Blanchard-Perotti model, we have estimated a version of the Romer-Romer model in which the explained variable is the level of output. The original Romer-Romer model, however, features the growth rate of output as the independent variable in an equation of the form

$$\Delta \hat{y}_t = \sum_{i=0}^{12} D_i \epsilon_{t-i}^{\tau} + u_t.$$

12

Figure 3: Impulse Response to a Tax Innovation in the DSGE and Difference-Romer-Romer Models
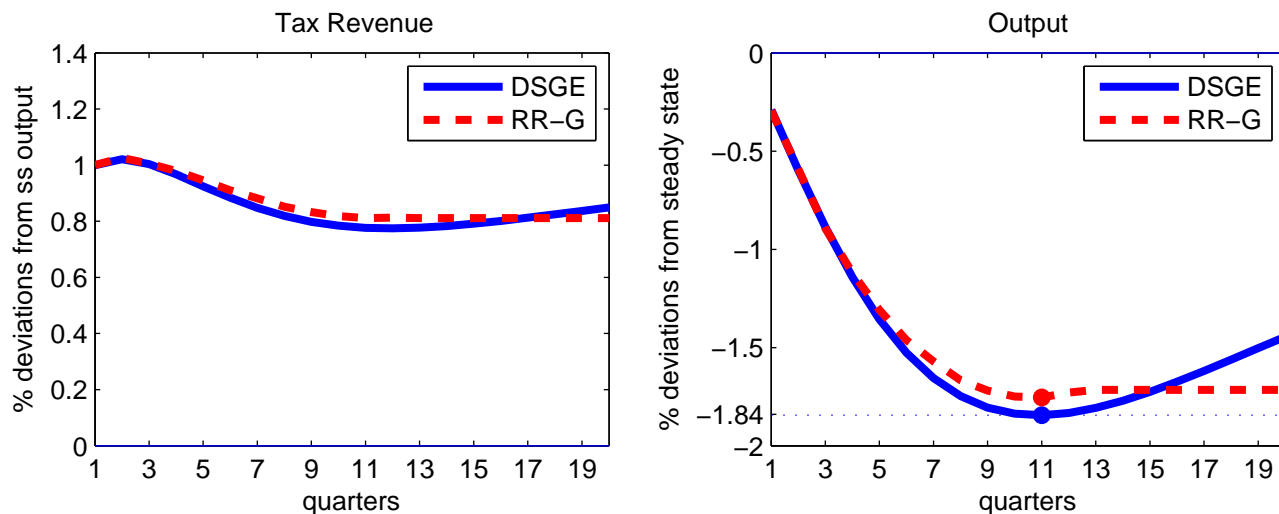
We note that our DSGE model implies that the growth rate of output possesses an MA(∞) representation in the four structural shocks. Consequently, by the same argument given in discussing the properties of the Romer-Romer model in levels, we have that an OLS estimation of the Romer-Romer model in growth rates delivers unbiased estimates of the coefficients D_i for $i = 0, 1, \ldots, 12$. As a corollary, the OLS estimator also delivers an unbiased estimate of the tax multiplier at any horizon below 12 quarters.

Figure 3 displays the impulse responses of tax revenues and the level of output to a tax innovation implied by the DSGE model and by the Romer-Romer model estimated using output growth as the independent variable. The figure shows that the Romer-Romer model in growth rates, like its counterpart in levels, captures nearly perfectly the transmission of tax shocks to output. In particular, the Romer-Romer model estimated using the growth rate of output uncovers the correct tax multiplier of 1.8 percent. For the remainder of the paper, we focus on the Romer-Romer model featuring the level of output as the dependent variable.

13

3.3 Evaluation of the Transmission-Mechanism Hypothesis

We have shown that, conditional on the correct identification of the exogenous tax disturbances, the average transmission mechanisms invoked by the Blanchard-Perotti and Romer-Romer reduced-form models yield virtually identical tax multipliers, which, in turn, are in line with the true multiplier associated with the DSGE data generating process. We take this result as suggesting two alternative explanations for the fact that empirical estimates of the Blanchard-Perotti and Romer-Romer models deliver significantly different tax multipliers. One possible explanation is that actual estimated tax multipliers are different because of small-sample undertainty. We explore this hypothesis in detail in section 5. A second possible explanation is that the Blanchard-Perotti and Romer-Romer regression models identify fundamentally different tax disturbances.

4 Anticipation

We have established that conditional on the correct identification of tax innovations, both the Blanchard-Perotti and Romer-Romer models satisfactorily capture the transmission mechanism of tax disturbances. We now explore whether this continues to be the case when the DSGE model is assumed to be driven by anticipated and unanticipated shocks. Schmitt-Grohé and Uribe (2010a) argue that at least half of the variance of output and other macroeconomic aggregates are driven by anticipated shocks. Mertens and Ravn (2009) argue that 37 out of the 70 exogenous tax liability changes identified by Romer and Romer (2007) are indeed anticipated, with a median anticipation horizon of six quarters. Anticipation can potentially affect the ability of both the Blanchard-Perotti model and the Romer and Romer model to capture the transmission mechanism of fiscal shocks. In the case of the Blanchard-Perotti model, the VAR system is estimated using data driven by both anticipated and unanticipated shocks. In the case of the Romer-Romer model, the econometrician regresses output onto a tax shock that is the sum of a purely unanticipated component and

14

a component that was announced in the past.

To introduce anticipation into the DSGE model, we assume the following specification for the four structural disturbances:

$$\epsilon_t^x = \nu_t^{x0} + \nu_{t-6}^{x6},$$

for $x = \tau, \mu, g, a$. We assume that ν_t^{x0} and ν_t^{x6} are distributed independently of each other and across time with mean 0 and standard deviation σ_{x0} and σ_{x6}, respectively. The innovation ν_t^{x0} is announced in period t and materializes in period t. That is, ν_t^{x0} is a purely unanticipated shock. The innovation ν_t^{x6} is announced in period t and materializes in period $t + 6$. That is, ν_t^{x6} is a disturbance anticipated six quarters. We pick six quarters of anticipation for tax shocks based on the finding of Mertens and Ravn (2009) referred to above. Schmitt-Grohé and Uribe (2010b) present econometric evidence of anticipation in technology, government spending, and preference shocks at horizons 4 and 8 quarters. For simplicity, we arbitrarily assume anticipation horizons of six quarters for these three shocks.

The calibration of the model is as before. In particular, we assume that $\sigma_{a0}^2 + \sigma_{a6}^2 = \sigma_a^2 (= 0.01^2)$. We also assume that tax shocks, government spending shocks, and preference shocks explain, respectively 20, 10, and 35 percent of the variance of output. Finally, we assume that the variance of each shock is explained in equal parts by its anticipated and its unanticipated components, that is, $\sigma_{x0}^2 = \sigma_{x6}^2$ for $x = \tau, g, \mu, a$.

Figure 4 displays with solid lines the impulse responses to a surprise tax shock in the DSGE model. Given that the model is approximated up to first order, these responses are identical to those corresponding to the DSGE model featuring only unanticipated shocks. The figure displays with broken lines the responses of the Blanchard-Perotti (top panels) and Romer-Romer (bottom panels) models. As before, each model is estimated 1,000 times on artificial data of length 250 quarter generated by the DSGE model, with 750 burn-in periods.

Figure 4: Impulse Response to a Tax Innovation in the Blanchard-Perotti and Romer-Romer Models when the DSGE Model is Driven by Anticipated and Unanticipated Shocks
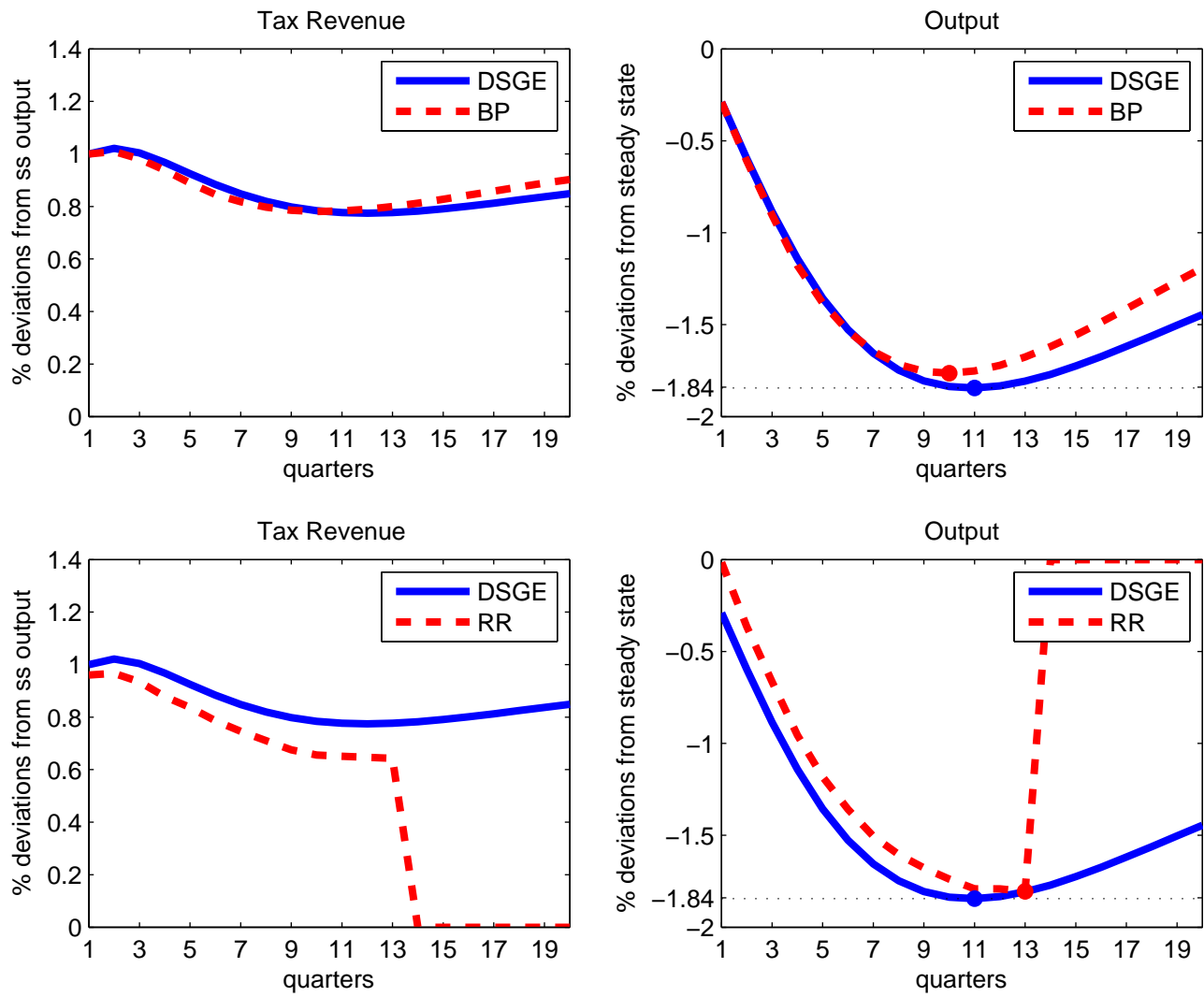
In the case of the Blanchard-Perotti model, we continue to assume that the econometrician is able to identify the impact response of output and tax revenues to an unanticipated tax shock. The coefficients of the VAR, however, are estimated using data from the DSGE model driven by anticipated and unanticipated disturbances in taxes, preferences, technology, and government spending. The figure shows that on average the Blanchard-Perotti model is able to capture quite well the true impulse response functions. In particular, it delivers an average tax multiplier of about 1.8, which is in line with its theoretical counterpart.

In the case of the Romer-Romer model, we run the following regression:

$$\hat{y}_t = \sum_{i=0}^{12} C_i(\nu_{t-i}^{\tau 0} + \nu_{t-i-6}^{\tau 6}) + u_t.$$

Notice that $\nu_t^{\tau 0} + \nu_{t-6}^{\tau 6}$ equals ϵ_t^{τ}, which is the total innovation in taxes materialized in period t. This is the correct regression for the Romer-Romer model, because the econometrician is not assumed to distinguish between anticipated and unanticipated tax disturbances. In spite of the assumed inability of the econometrician to isolate the unanticipated tax innovation, the regression recovers remarkably well the true impulse response to an unanticipated tax shock. In particular, the estimated Romer-Romer model correctly predicts an average tax multiplier of 1.8 percent.

We note that, unlike in the economy driven only by unanticipated shocks, in the economy under study here, the Romer-Romer regression does not have a theoretical underpinning. To see this, note that the $MA(\infty)$ representation of $\hat{y}_t$ implied by the DSGE model with anticipation is of the form

$$\hat{y}_t = \sum_{i=0}^{\infty} C_i^0 \nu_{t-i}^{\tau 0} + \sum_{i=0}^{\infty} C_i^6 \nu_{t-i}^{\tau 6} + \text{rest},$$

where in this equation the term labeled 'rest' is orthogonal to the anticipated and unanticipated tax disturbances that appear on the right hand side.

By regressing $\hat{y}_t$ onto $\nu_{t-i}^{\tau 0} + \nu_{t-6-i}^{\tau 6}$, the Romer-Romer regression incorrectly imposes the

17

Figure 5: Tax Coefficients in the MA(∞) Representation of Output
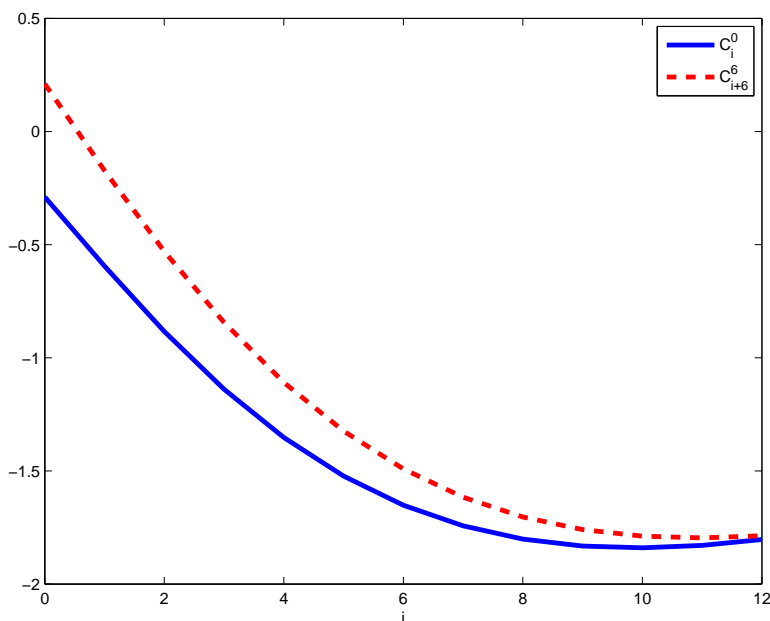
restriction that $C_i^0 = C_{i+6}^6$ for $i = 0, 1, \ldots, 12$. Theoretically, these two coefficients are not identical. However, as can be seen in figure 5, numerically in the DSGE model under study these two coefficients are quite similar. As a result, the Romer-Romer model does not miss by much the magnitude of the multiplier to an unanticipated tax shock (see the bottom-right panel of figure 4). The economic reason for the similitude of the MA coefficients C_i^0 and C_{i+6}^6 is that, given the state of the economy (i.e., given the levels of capital, habits, durable goods, past investment, and the levels of the exogenous shocks) the economy responds equally to a purely unanticipated shock that occurs in the current period and to a shock that materializes in the current period but was announced in the past.

The results presented in this section, suggest that, conditional on the correct identification of tax innovations, both the Blanchard-Perotti and Romer-Romer models provide on average an adequate reduced-form description of the underlying transmission mechanism of fiscal shocks and a fairly precise estimate of the tax multiplier. The success of both empirical models occurs in spite of the fact that half of the variance of the data generated by the DSGE model is explained by six-quarter anticipated shocks. This result further strengthens

18

the main thesis of this paper, namely, that the differences in tax multipliers implied by estimates of the Blanchard-Perotti and Romer-Romer models on actual data might be due to factors other than the transmission mechanisms that these models invoke.

5   Small Sample Uncertainty and Tax Multipliers

Thus far we have concentrated attention on the average responses of the Blanchard-Perotti and Romer-Romer models to tax shocks. These averages were taken over 1,000 samples of artificial data, each 250 quarters long. As explained earlier, the sample size of 250 quarters is meant to roughly capture the length of the postwar period. We now address the issue of small sample uncertainty. This analysis will allow us to answer questions such as how likely it is to observe within a finite sample of 250 quarters an estimate of the Romer-Romer tax multiplier that exceeds the Blanchard-Perotti multiplier by two percentage points, conditional on the correct identification of the tax shock. The difference of two percentage points reflects the difference between the two multipliers when estimated on actual U.S. postwar data.

We characterize the distribution of tax multipliers at horizons 10, 11, and 12 quarters. The tax multiplier at horizon 11, for instance, is defined as the percentage deviation of output from steady state in period 11 triggered by an increase in tax revenues in period 1 equivalent to 1 percent of output. We focus on horizons 10, 11, and 12 quarters because in our assumed data generating process the maximum contraction in output in response to an exogenous unanticipated tax innovation occurs in quarter 11.

The top panel of table 2 displays summary statistics of tax multipliers obtained from 1000 samples of 250 quarters of artificial data generated using the DSGE model. As discussed in previous sections, the mean of the tax multiplier implied by the Blanchard-Perotti and Romer-Romer models is quite close to the true multiplier of 1.8. But, as the top panel of the table shows, the short sample uncertainty surrounding both estimates is quite large, as reflected in the standard deviation of the multiplier estimates over the 1,000 samples.

Table 2: Small Sample Properties of the Blanchard-Perotti and Romer-Romer Tax Multiplier Conditional on the Correct Identification of Tax Shocks

Horizon	m^{BP}			m^{RR}			Probability	
(qrt.)	Mean	Median	Std.Dev.	Mean	Median	Std.Dev.	$m^{RR} > m^{BP}$	$m^{RR} - m^{BP} > 2$
	Sample Size 250 Quarters							
10	1.71	1.64	0.96	1.88	1.90	2.04	0.53	0.17
11	1.69	1.62	0.97	1.90	1.94	2.04	0.54	0.17
12	1.65	1.59	0.97	1.90	1.85	2.04	0.55	0.18
	Sample Size 1000 Quarters							
10	1.80	1.79	0.47	1.82	1.74	0.95	0.48	0.02
11	1.79	1.78	0.47	1.83	1.77	0.94	0.50	0.01
12	1.77	1.75	0.47	1.83	1.75	0.94	0.50	0.02

Note: m^{BP} and m^{RR} stand for the tax multipliers implied by the Blanchard-Perotti and Romer-Romer models, respectively. All statistics are computed from 1000 samples of artificial data generated from the DSGE model.

The Romer-Romer multiplier estimate appears to be substantially more vulnerable to small-sample uncertainty. Its associated standard deviation is twice as large as the one associated with the Blanchard-Perotti estimate. This suggests that, conditional on the ability of both models to correctly identify exogenous tax shocks, the Blanchard-Perotti model delivers a more efficient estimate of the tax multiplier.

The penultimate column of table 2 shows that the estimated Romer-Romer tax multiplier can be larger or smaller than the Blanchard-Perotti multiplier with almost equal probability. The last column shows that the probability that in a sample of 250 quarters the Romer-Romer multiplier is two percentage points larger than the Blanchard-Perotti multiplier (as estimated in actual data) is 17 percent. This means that one cannot reject the hypothesis that, conditional on correct identification, the observed differences in estimated tax multipliers is due to small sample uncertainty.

The bottom panel of table 2 suggests that, conditional on the correct identification of tax shocks, both the Blanchard-Perotti and Romer-Romer reduced-form models produce consistent estimates of the tax multiplier. As the number of observations increases from 250 to 1000 quarters, the standard deviations of both estimates fall by half, and the probability

of observing a Romer-Romer multiplier that exceeds the Blanchard-Perotti multiplier by two percentage points falls from 17 percent to 2 percent.

6   Hybrid Specifications

In a recent contribution, Favero and Giavazzi (2010) augment the Blanchard-Perotti model by including the Romer-Romer shock as a regressor. They then compute the tax multiplier induced by an innovation in the Romer-Romer shock. Favero and Giavazzi find that the size of the multiplier is around unity, in line with the results of Blanchard and Perotti (2002).[1] They argue that their combined model is the best approach to measure tax multipliers. The key premise of their paper is that both the Blanchard-Perotti and the Romer-Romer models correctly identify exogenous tax innovations. They interpret their findings, therefore, as suggesting that the Romer-Romer single-equation model fails to capture the transmission of tax shocks onto output.

We have shown that conditional on the correct identification of tax shocks, both the Blanchard-Perotti and the Romer-Romer models produce the correct transmission mechanism and tax multipliers on average. We have also shown, again conditional on correct identification, that the Blanchard-Perotti model yields more efficient estimates of the tax multiplier than does the Romer-Romer model when small sample uncertainty is taken into account. In light of these findings, we analyze the Favero-Giavazzi specification along two dimensions, namely, bias and efficiency of the estimated tax multiplier.

We consider the following version of the Favero-Giavazzi model:

$$X_t = \sum_{i=1}^{4} A_i X_{t-i} + C \epsilon_t^\tau + u_t,$$

---

[1]Mertens and Ravn (2009) also estimate a hybrid specification that combines a VAR system with the Romer-Romer tax shock as a regressor. In their specificaiton, these auhors distinguish between anticipated and unanticipated Romer-Romer shocks. Unlike Favero and Giavazzi (2010), Mertens and Ravn estimate a tax multiplier of around 2 percent. This difference deserves study.

Figure 6: Impulse Response to a Tax Innovation in the DSGE and Favero-Giavazzi Models
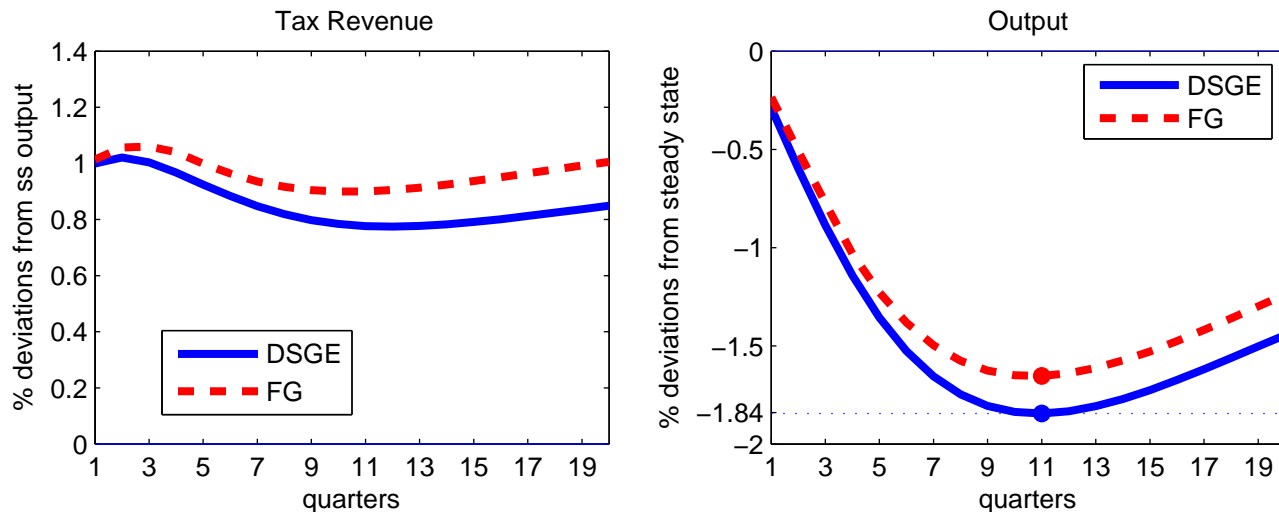
where the notation is as in earlier sections. As in our previous Montecarlo exercises, we estimate this version of the Favero-Giavazzi model 1,000 times. Each estimation uses a sample of 250 quarters generated using the DSGE model. Figure 6 displays with solid lines the response to a tax innovation implied by the DSGE model and with broken lines the average response implied by the Favero-Giavazzi model. Like the Blanchard-Perotti and Romer-Romer models, on average the Favero-Giavazzi model does a good job at uncovering the transmission mechanism of tax innovations. We conclude that, conditional on the correct identification of the tax shock, the Favero-Giavazzi model produces an unbiased estimate of the tax multiplier. In this respect, therefore, the reduced-form transmission mechanisms invoked by the Blanchard-Perotti, Romer-Romer, and Favero-Giavazzi models are on equal footing.

But is the Favero-Giavazzi estimate of the tax multiplier more efficient than the one produced by the Blanchard-Perotti or Romer-Romer models? We find that for a sample size of 250 quarters, the estimate of the tax multiplier at a horizon of 11 quarters implied by the Favero-Giavazzi model has a standard deviation across the 1,000 samples of 1.1. This figure is slightly larger than the standard deviation of 0.97 found for the Blanchard-Perotti estimate and significantly smaller than the standard deviation of the Romer-Romer estimate

(see table 2). We conclude that, conditional on the correct identification of the tax shock, the Favero-Giavazzi model produces a more efficient estimate than the Romer-Romer model, but offers no efficiency gains with respect to the Blanchard-Perotti model.

7   Conclusion

Since the revolutionary ideas of Keynes, governments have been fighting recessions with spending increases and tax cuts. The justification of these policy measures often references estimates offiscal multipliers. But the literature on the size of fiscal multipliers, be it tax or government spending multipliers, does not speak with one voice. The VAR literature delivers tax multipliers of about one percent, whereas the narrative literature produces tax multipliers of about three percent. These differences are sizable enough to leave policymakers without a clear guidance on the power of tax cuts to stimulate the economy.

The VAR and narrative approaches differ along two important dimensions. One is the assumed transmission mechanism. The second is the methodology for identifying tax shocks. This paper uses a micro-founded data-generating process to evaluate the hypothesis that differences in estimated tax multipliers are due to differences in the assumed transmission mechanism. In testing this hypothesis, it is assumed that both methodologies identify the same tax shock. The main finding of this paper is that this hypothesis is rejected. Both reduced-from models correctly uncover the size of the underlying tax multiplier.

Our results leave open two alternative explanations for the observed differences in estimated tax multipliers. One is small sample uncertainty. Accordingly, we explore the small-sample properties of the estimated tax multipliers stemming from the VAR and narrative models. We find that, conditional on both models identifying the same tax shock, small sample uncertainty is large. In fact, small sample uncertainty accounts for all of the observed differences in estimated tax multipliers according to our data generating process.

All of the results reported in this investigation are conditional on the assumption that

the VAR and narrative approaches successfully identify exogenous innovations in taxes. An alternative explanation of the observed differences in estimates of tax multipliers is, of course, that the two methodologies fail to identify the same tax shock. We believe that this alternative warrants future investigation.

References

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

Favero, Carlo, and Francesco Giavazzi, "VAR-BAsed and Narrative Measures of the Tax Multiplier," manuscript, Universita Bocconi, May 2010.

Mertens, Karel, and Morten O. Ravn, "Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated U.S. Tax Policy Shocks," Cornell University, 2009.

Romer, Christina, and David H. Romer, "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," NBER Working Paper No. 13264, July 2007.

Schmitt-Grohé, Stephanie and Martín Uribe, "What's News in Business Cycles," Columbia University, April 2010a.

Schmitt-Grohé, Stephanie and Martin Uribe, "Business Cycles With A Common Trend in Neutral and Investment-Specific Productivity," NBER Working Paper No. 16071, June 2010b.