American Economic Journal: Economic Policy 2012, 4(2): 214–237
http://dx.doi.org/10.1257/pol.4.2.214

The Effects of Tax Shocks on Output: Not So Large, But Not Small Either†

By Roberto Perotti*

I argue that, on theoretical grounds, the discretionary component of taxation should be allowed to have different effects than the automatic response of tax revenues to macroeconomic variables. Based on a novel dataset, *I show two results. First*, *responses to a tax shock that allow for a distinction between the discretionary and the endogenous components of tax changes are about halfway between the large effects estimated by Romer and Romer* (*2010*) *and the smaller effects estimated, for instance, by Favero and Giavazzi* (*2012*) *or Blanchard and Perotti* (*2002*). *Second*, *there is almost no statistically significant evidence of anticipation effects.* (*JEL* E23, E62, H22, H24, H25, K34)

In a seminal paper, Romer and Romer (2010) (henceforth, RR) construct measures of tax changes based on the original documentation accompanying tax bills, and show that they have large negative effects on output: an exogenous increase in taxes by 1 percentage point of GDP can lead to a decline in GDP by 3 percentage points after three years.

These magnitudes have been criticized as implausibly large. Favero and Giavazzi (2012) (henceforth, FG) challenge the specification used by RR, arguing that it cannot be interpreted as a proper (truncated) moving average representation of the output process. They show that when the system is estimated in its VAR form, or its correct truncated MA representation, a unit realization of the RR shock has much smaller effects on GDP, typically about −0.5 percentage points.

I start by constructing a dataset that extends the RR approach in two main dimensions. First, it breaks down each tax bill into its main items (personal income, corporate income, social security, and indirect taxes) and their sub-items. Second, for each item, I track the effective starting dates, that often differ between items of

* IGIER Universita' Bocconi, Via Roentgen 1, 20136 Milano ITALY; and CEPR and also NBER (e-mail: roberto.perotti@unibocconi.it) I thank an anonymous referee for very helpful suggestions; Carlo Favero for several discussions that helped clarify the issues, and for pointing out a crucial mistake in an earlier version of the paper; Yuriy Gorodnichenko, who discussed an earlier version at the NBER-IGIER TAPES conference held in Varenna, Italy in June 2010; and Morten Ravn for very useful comments. Elia Boe provided excellent research assistance. This paper was 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 Programme, contract number 244725.

† To comment on this article in the online discussion forum, or to view additional materials, visit the article page at http://dx.doi.org/10.1257/pol.4.2.214.

each tax bill, and the pattern of exogenous quarterly changes over time. Doing so is important in order to track correctly the news process of the private sector (see Leeper, Richter, and Walker 2012).

On theoretical grounds, one should expect the discretionary component of tax changes to have stronger effects on output than the component capturing the endogenous response of taxes to, say, output fluctuations. Under this assumption, I show that the approaches of FG and of Blanchard and Perotti (2002) generate impulse responses that are biased towards zero. By adopting a VAR approach that allows for a different impact of the discretionary and endogenous components of taxation, I show that the estimated effects of a discretionary tax change is larger (in absolute value) than that estimated by FG, although smaller than that estimated by RR. Now a 1 percentage point of GDP increase in taxation is typically associated with a decline in GDP by about 1.3 percentage point after three years.

Mertens and Ravn (2012) (henceforth, MR) decompose changes of current taxation into anticipated and unanticipated changes, and also estimate the response of output to changes in future taxation. If individuals are liquidity constrained, the responses to anticipated and unanticipated changes in current taxation should be approximately the same; in this case, one would expect also a limited response to changes to future expected taxation. MR find contradictory evidence on these two points: the responses to anticipated and unanticipated changes in current taxation are similar, but there is evidence of a *positive* response of output to positive changes to future expected taxation, pointing to the presence of strong anticipation effects.

I too find similar responses to anticipated and unanticipated changes in current taxation; but, unlike MR, I find little evidence of a response to changes in future expected taxation. Thus, using a dataset that tracks explicitly the starting dates and the exogenous changes over time of the different items of each tax bill leads one to different conclusions about anticipation effects.

I then argue that the same logic one applies to current taxation should be applied to future expected taxation, i.e., one should distinguish anticipated and unanticipated changes in future expected taxation; in fact, one should allow for a different dynamics for each shock to future taxation, at different horizons. When I do this, again I find no robust evidence of positive responses of output to shocks to future expected taxation.

The outline of the paper is as follows. Section I discusses briefly which and how tax shocks matter in a stylized economy. Section II discusses how measures of changes to discretionary taxation are constructed. Section III presents an overview of the dataset I use. Section IV introduces the specification I use to accommodate differences in the effects of the discretionary and endogenous changes in tax revenues. Section V presents alternative specifications, including those by RR and FG, and maps them into the coefficients of the benchmark model in which discretionary taxation has different effects than the endogenous component of taxation. Section VI summarizes these specifications. Estimation results are discussed in Section VII. Section VIII discusses the results of decomposing changes to current taxation into anticipated and surprise changes. Section IX presents some robustness checks. Section X concludes.

I. Why Tax Shocks Matter

Consider the benchmark representative agent neoclassical model, with infinite horizon, rational expectations, and no borrowing or lending constraints. Government spending consists of purchases of goods and services which do not enter the household's utility function nor the production function; taxes are lump sum. In this economy, tax shocks matter only if they change the present value of government spending, hence, from the intertemporal government budget constraint, the present value of taxes the household expects to pay: when the present value of taxes increases, human wealth falls, and so do consumption and leisure; the ensuing increase in labor supply increases income, but welfare falls (see Baxter and King 1993). For a given change in the present value of government spending, the timing of taxes does not matter;[1] what matters is the distinction between **surprise** and **anticipated** changes to the present value of taxation. Since taxation has only a wealth effect, changes in taxation matter when they enter the information set of the individual, which could be before they are actually implemented.

Now remove the assumption of lump-sum taxation, and consider a proportional tax on income. In this case changes to taxation can affect the economy also via intra- and inter-temporal substitution effects. These effects can be complicated, as they affect various margins, and depend on the expected time path of the tax rate. But the key intuition is simple. Suppose the public anticipates higher taxes on income in $t + 1$; by the inter-temporal substitution effect, individuals will prefer to work in t rather than in $t + 1$, because the relative real after tax wage rises in t; they will also prefer to consume in t rather than in $t + 1$, because the real interest rate falls. Hence, output might even increase in t in response to an anticipated future increase in distortionary taxation (see Mertens and Ravn 2011a for an example). When the tax rate actually increases at time $t + 1$, labor supply falls: as the real after-tax wage falls, the incentives to work decline.[2] This suggests that it is important to distinguish between **surprises** in **current** taxation (that is, implemented at t) and in **future** taxation.[3]

Consider now removing a different assumption of the benchmark model, that of no liquidity constraints. For simplicity, assume that individual cannot borrow nor lend.[4] In this case, changes to expected future taxation have no effects on current variables: tax changes have effects only when implemented. In this case, what matters is the distinction between **current** (regardless of whether anticipated or unanticipated) and **future** (regardless of whether anticipated or unanticipated) changes in taxation. It also suggests an obvious test of liquidity constraints: surprise and

[1] Of course, this begs the question of how to control for the change in the present value of expected government spending. This is difficult: including government spending in the VAR does not address the issue entirely, since it controls at most for the "average" response of government spending to a tax shock. See below, Section IX for an attempt to control for the *expectation* of the present value of future government spending.

[2] Obviously one can imagine much more complicated scenarios, with changes in taxation of positive and negative sign expected to occur in periods $t + 2$, $t + 3$, $t + 4$ etc. One can obtain virtually any effect on current variables depending on the pattern of future announced tax changes.

[3] The importance of news about future taxes has been emphasized by Mertens and Ravn (2011a, 2012) and by Leeper, Richter, and Walker (2012).

[4] See Galí, López-Salido, and Vallés (2007) for a model that combines such agents with forward looking, unconstrained agents.

anticipated changes in current taxation should have the same effects. Similar arguments apply to the case where individuals are not liquidity constrained, but have finite lifetimes and are not linked by intergenerational altruism: future taxation falls only or mainly on future generations, hence it tends to have weaker effects on today's behavior.

II. Estimates of Discretionary Taxation

Narrative estimates of tax changes are based on "discretionary" changes to taxation (also called changes in "cyclically adjusted" revenues, or "fiscal impulse"). These are meant to capture the intentional actions of policymakers, as opposed to the automatic effects of the business cycle on revenues. In this section, I introduce the notation and describe the construction of measures of discretionary taxation.

"Discretionary" taxation is a convention. One starts by defining a reference level of output, such as the natural level of output, or trend output, or last year's output. Given existing legislation, one can then define the reference level of tax revenues, i.e., the level of revenues that would obtain if output were at its reference level. Deviations of revenues from this notional reference level can be due to three sources: the deviation of output from its own reference level, the discretionary action by policymakers (changes in tax rates, depreciation rules etc.), and other random causes.[5]

To simplify the exposition, and following Blanchard (1990), from now on I will assume that the reference level of output is output in the previous period. Letting s_t and y_t denote the log change in tax revenues and output, one can write

$$s_t = d_{t/t} + \eta y_t + \mu_t, \tag{1}$$

where $d_{t/t}$ is a suitable measure of the discretionary change in taxation. η is the automatic elasticity of revenues to output, and μ_t is a random term, which I take to be i.i.d. with zero mean. Following Blanchard and Perotti (2002), I will assume that $d_{t/t}$ is uncorrelated with μ_t: because of decision and implementation lags, policymakers cannot react to exogenous shocks within a quarter. Instead, as we will see below μ_t in general is correlated with y_t.

With an outside estimate of the elasticity η, the standard procedure to estimate $d_{t/t}$ is precisely to subtract ηy_t from s_t. Conceptually, this is what Blanchard and Perotti (2002) do, except that they use innovations estimated from a VAR rather than actual changes in revenues and output (note, though, that by doing this one cannot separate $d_{t/t}$ from μ_t).

RR turn this procedure around by starting from estimates of $d_{t/t}$ as provided by official documents, and based on the specific provisions of each tax bill enacted by Congress. Thus, they decompose s_t into the discretionary component $d_{t/t}$ and the term $\eta y_t + \mu_t$, which with a slight abuse of terminology I will call the "endogenous component" of revenue changes.

[5] Of course, the distinction is not so clear-cut as it might appear: one could object that the policymaker could always have prevented, by a suitable change in rules, the automatic effect of the deviation of output from its reference level.

How is $d_{t/t}$ constructed in practice? Let $D_{t+i/t}$, with $i \geq 0$, denote discretionary taxation at time $t + i$, expected at (i.e., based on all laws enacted up to) time t. A law enacted at time t specifies a path of revisions to expected discretionary taxation, from time t onward: $D_{t/t} - D_{t/t-1}$, $D_{t+1/t} - D_{t+1/t-1}$, $D_{t+2/t} - D_{t+2/t-1}$ up to time $t + M$, where M is the maximum horizon for a law (of course, many of these revisions will be 0).[6]

Let $u_{t/t-j}$ be the revision, caused by a law enacted at time $t - j$, of the expected *change* in discretionary taxation at time t:[7]

$$u_{t/t-j} \equiv (D_{t/t-j} - D_{t-1/t-j}) - (D_{t-1/t-j-1} - D_{t-1/t-j-1}) \tag{2}$$

$d_{t/t}$ is equal to the sum of all such revisions, caused by all laws between t and $t - M$:

$$\begin{aligned} D_{t/t} - D_{t-1/t-1} &= \underbrace{[D_{t/t} - D_{t/t-1}]}_{\mathbf{u}_{t/t}} + \underbrace{[(D_{t/t-1} - D_{t-1/t-1}) - (D_{t/t-2} - D_{t-1/t-2})]}_{\mathbf{u}_{t/t-1}} \\ &\quad + \underbrace{[(D_{t/t-2} - D_{t-1/t-2}) - (D_{t/t-3} - D_{t-1/t-3})]}_{\mathbf{u}_{t/t-2}} + \dots . \\ &\quad + \underbrace{[(D_{t/t-M} - D_{t-1/t-M}) - (D_{t/t-M-1} - D_{t-1/t-M-1})]}_{\mathbf{u}_{t/t-M}} \\ &= \sum_{j=0}^{M} u_{t/t-j}. \end{aligned} \tag{3}$$

(Note that, if M is the maximum horizon, $D_{t/t-M-1} = D_{t-1/t-M-1}$.)[8]

In expression (3), the first component, $u_{t/t}$, is unanticipated, the rest is known at date t:

$$\begin{aligned} d_{t/t} &= \underbrace{u_{t/t}}_{\textbf{contemp. revision}} + \underbrace{\sum_{j=0}^{M-1} u_{t/t-j-1}}_{\textbf{sum of all past revisions}} \\ &= u_{t/t} + d_{t/t-1}. \end{aligned} \tag{4}$$

[6] In practice, I set $M = 20$; this leaves out only a few tax changes that occurred more than 20 quarters after enactment of a tax bill. Specifically, five reductions of the telephone excise tax set by the P.L. 91-614 of 1970 for 1976, 1977, 1978, and 1979, each for \$160mn; the end of the repeal of the 1985 and 1986 increases in accelerated cost recovery deductions, set by the 1982 TEFRA for 1988, for a total of \$15.9bn; and the increase in the social security tax rate decided in 1972:III for 1978:I, for \$3.3bn.

[7] Note that when $j = 0$, $u_{t/t} = D_{t/t} - D_{t/t-1}$ because $D_{t-1/t-1} = D_{t-1/t}$.

[8] More generally, the discretionary change in date $t + i$'s taxation, expected at date $t - s$, is the the sum of all revisions, to the expected change in discretionary taxation at time $t + i$, decided up to date $t - s$

$$\begin{aligned} d_{t+i/t-s} &\equiv D_{t+i/t-s} - D_{t+i-1/t-s} \\ &= \sum_{j=0}^{M-s-i} u_{t+i/t-j-s} \quad i + s \leq M. \end{aligned}$$

Obviously, when $s = i = 0$ we have expression (3), given that $D_{t-1/t} = D_{t-1/t-1}$.

Table 1—Breakdown of Total Taxes into Different Items

Individual	Corporate	Indirect	Soc. Sec.
1. Tax rates	Tax rates	Indirect taxes	Tax rates
2. Deductions, allowances	Employment credit		Earnings base
3. Tax credits	Investment tax credit		Others
4. Capital gains	Depreciation		
5. Depreciation	Others		
6. Earned income tax credit			
7. Rebates			
8. Estate and gift			
9. Others			

In fact, the second term on the r.h.s. of (4), $d_{t/t-1}$, is the sum of all past revisions, from $t - 1$ backward, to the expected change in discretionary taxation at time t.[9]

III. The Data

To estimate the models of this paper, I use the original RR data on discretionary tax changes, and also a new dataset that I have assembled starting from the individual items detailed in Table 1 above.[10]

Although in this paper, which focuses on methodological issues, I aggregate all tax items into one measure of total tax revenues, starting from each item separately is important because it allows one to better track the effective starting dates and the exogenous quarterly changes of each item of a given tax bill.

RR typically report the effect of a tax bill as a single number, the first full year effect after enactment.[11] For instance, for the 1982 Tax Equity and Fiscal Responsibility Act signed on September 2, 1982, RR give a single number, \$26.4bn, in 1983:I; I record 16 different items, which are legislated to take effect at different dates, from 1982:IV to 1984:I. Examples of these items are: a medical expense deduction, effective 1983:I, a ten percent casualty deduction floor, effective 1984:I, and an increase in airport and airway taxes, effective 1982:IV. Also, by breaking down each tax bill into its items, it is possible to take into account exogenous changes over time in the effects of a tax bill, since often different items have different time paths. For instance, legislation that allows for accelerated depreciation causes a large change in the time profile of receipts and liabilities,[12] but a small change in their present discounted value: receipts and liabilities decline initially, only to increase later. Using the first full-year effect would therefore provide a distorted picture of the effects of the tax bill. An

[9] Thus, the RR observations are not strictly speaking tax "shocks" in the usual sense, because they contain an anticipated component, a fact pointed out by Mertens and Ravn (2011a) that will be exploited below.

[10] The dataset is available on my Web site. Since it is based on budget documents, obviously it covers, like the RR dataset, only federal tax revenues.

[11] In some cases, RR reports more than one number, when important measures are legislated to take effect on different dates.

[12] RR refer to their tax measures as changes in "liabilities;" however, in some cases the sources do not distinguish between receipts and liabilities, and in others they display receipts. I construct two separate datasets on receipts and liabilities. In this paper, in order not to clutter the exposition I report results using receipts. In Perotti (2011a) I use data also on liabilities, and show that for the purposes of this paper the differences with data on receipts are not large.

example is still 1982 TEFRA, whose largest item was the repeal of increases in accelerated cost recovery deductions for 1985 and 1986, implying an anticipated increase in revenues in 1985 and an even larger increase in 1986–87, and then a large decline from 1988:I.[13]

Even aside from these cases, the effect on receipts of a tax measure in the first fiscal or calendar year after enactment is often smaller than the effect in later years. On the other hand, there is a normal trend increase in the effect due to the assumed exogenous increase in GDP over time. Hence, I adopt the rule that, if in fiscal year $x + 1$ receipts are different from year x by a factor of more than 30 percent, I display a change in $x + 1$:1.

Thus, I track the exogenous quarterly changes of each item included in a tax bill. In terms of the notation of Section II, for each tax bill signed at time t I compute the revision of the change in discretionary taxation $u_{t+i/t}$, for i up to 20 quarters, and for each item included in the tax bill. Some sources not used by RR, most notably the *Survey of Current Business*, display the quarterly effects (as opposed to the annual effects used by RR) of the different tax bills. I use these sources, and I complement them with a specific methodology to infer quarterly changes from annual data.

First, I keep track of changes in withholding rates for individuals. If these change at the time of enactment, I assume that receipts track liabilities from the time of enactment, unless receipt data indicate otherwise. If withholding rates are not changed immediately, some or all of the liabilities are paid in quarter 1 and 2 of the next calendar year, when tax declarations are filed and net settlements are carried out. For corporations, I convert liabilities into receipts, and yearly receipts into quarterly receipts, using the legislation in place in each year determining the timing of tax payments by corporations. These rules are complicated, also because they depend on the choice of the tax year. I illustrate my methodology in the Appendix.

These rules are important not only to calculate the correct time path of receipts, but also the retroactive components. Several tax changes have retroactive components, i.e., they apply to a period before the time of enactment. RR assume that all retroactive liabilities accrue in one installment in the first quarter after enactment.[14] In reality, individuals and corporations pay retroactive liabilities in a variety of ways, and the time pattern of payments has changed several times since 1945. I take all this into account following the methodology also illustrated in the Appendix.

Thus, I end up with the following classification of tax changes, summarized in Table 2. A tax change is "legislated, unanticipated" ("LU" in the last column of the table) if the change to the tax code is legislated to occur within 90 days after enactment, and receipts start within 90 days after the change in legislation (row 1). It is "legislated, anticipated" ("LA") if either the change to the tax code is legislated to take effect within 90 days from enactment, but receipts start more than 90 days after the legislated change (row 2); or the change to the tax code is legislated to take effect more than 90 days after enactment (rows 3 and 4). A tax change is

[13] In the first 10 to 15 years of the sample, typically the sources report only the full-year effect of the tax measures; but starting around 1960, they often report receipts (and in some cases liabilities) over a longer horizon, from five years to—in the nineties—up to 10 years ahead.

[14] This is the assumption I also make in constructing the dataset on liabilities.

TABLE 2—CLASSIFICATION OF DISCRETIONARY TAX CHANGES

1. LC within 90 days after enactment, R within 90 days from LC	LU
2. LC within 90 days after enactment, R change more than 90 days after LC	LA
3. LC more than 90 days after enactment, R change within 90 days from LC	LA
4. LC more than 90 days after enactment, R change more than 90 days after LC	LA
5. LC within 90 days after enactment, R change (without LC) after first R change	NLA
6. LC more than 90 days after enactment, R change (without LC) after first R change	NLA

Note: LC: "Legislated Change;" R: "Receipts;" LU: "Legislated, Unanticipated;" LA: "Legislated, Anticipated;" NLA: "Nonlegislated, Anticipated."

"non legislated, anticipated" ("NLA") if it is not associated with a legislated change to the tax code, and follows from the application of the 30 percent rule (rows 5 and 6). This classification generates two datasets: in the first I keep track only of changes that are explicitly legislated, i.e., of the first two categories, "LU" and "LA." In the second I also include the third category, "NLA," i.e., changes in receipts that are captured by the 30 percent rule. The results in this paper are based on this second definition; the differences with the narrower definition are minor. RR and MR use only legislated changes; the latter also distinguish between anticipated and unanticipated changes.

IV. The Effects of Discretionary Taxation

The key point of this paper is that one should allow the discretionary component $d_{t/t}$ and the endogenous component $\eta y_t + \mu_t$ to have different effects on output. There are at least two reasons for this. First, changes to discretionary taxation consist of changes in tax rates and tax rules, hence they are more distortionary. Second, they are more persistent; in fact, if deviations of output from its reference level sum to 0 over the cycle (such as when the reference level is trend output or potential output), and if agents are not liquidity constrained, then the nondiscretionary component of taxation should have no effect on the agents' behavior.[15] Note also that from an econometric point of view there is no "tax" shock unless there is a change in $d_{t/t}$ (and possibly μ_t).

To put it all together, I consider a minimalist model that however has all the ingredients one needs. The "true" model includes an equation for s_t, like equation (1) repeated here for convenience

$$s_t = d_{t/t} + \eta y_t + \mu_t \tag{5}$$

[15] An important caveat is that a change in tax rates could also affect the elasticity η and therefore it could contaminate the estimation of the discretionary and the endogenous components. Observe, however, that if taxes are proportional a change in tax rates does not affect the elasticity η. Hence, this effect is likely to be second order. One could also argue that, as output changes over the cycle, so does the elasticity of tax revenues, because individuals are moved into different tax brackets. Hence a purely cyclical source of changes in revenues could impact on the behavior of individuals. This effect too is likely to be second order. In the end, any measure of the elasticity of tax revenues, whether estimated as here or constructed from the tax code as international organizations do, is bound to be affected by measurement error.

and an equation for y_t

$$y_t = \alpha y_{t-1} + \gamma_1 d_{t/t} + \gamma_1'(s_t - d_{t/t}) + \gamma_2 d_{t-1/t-1} + \gamma_2'(s_{t-1} - d_{t-1/t-1}) + \varepsilon_t, \tag{6}$$

where ε_t is an i.i.d. error term with 0 mean. ε_t can be correlated with μ_t but, because of decision and implementation lags (see Section II), it is not correlated with $d_{t/t}$. For illustrative purposes, the model includes only two endogenous variables, s_t and y_t, and one lag; in the more general case which I estimate below, the model includes also the interest rate, inflation, and government spending, and a longer lag structure. As in RR, initially I assume that there is no anticipation effect from expectations of changes in future taxation.

If $\gamma_1 = \gamma_1'$ and $\gamma_2 = \gamma_2'$, equation (6) reduces to

$$y_t = \alpha y_{t-1} + \gamma_1 s_t + \gamma_2 s_{t-1} + \varepsilon_t \tag{7}$$

and output depends on total revenues; this is the assumption e.g., of Blanchard and Perotti (2002). If at the other extreme $\gamma_1' = \gamma_2' = 0$, equation (6) becomes

$$y_t = \alpha y_{t-1} + \gamma_1 d_{t/t} + \gamma_2 d_{t-1/t-1} + \varepsilon_t \tag{8}$$

and output depends only on discretionary taxation (this is the assumption of RR). Orthogonality of $d_{t/t}$ to ε_t is the first identifying assumption of RR. A second assumption is that $d_{t/t}$ is unpredictable using lagged variables in the information set of the econometrician. When $\gamma_1' = \gamma_2' = 0$, these two assumptions are sufficient to identify the impulse response of y_t to $d_{t/t}$ from an OLS estimate of (8).[16]

However, in the more general case they are no longer sufficient. When $\gamma_1 \neq \gamma_1'$, $\gamma_2 \neq \gamma_2'$, and $\gamma_1' \neq 0$, $\gamma_2' \neq 0$, using (5) equation (6) becomes

$$y_t = \pi_y^{MR} y_{t-1} + \pi_{d_0}^{MR} d_{t/t} + \pi_{d_1}^{MR} d_{t-1/t-1} + \phi_t^{MR} \tag{9}$$

$$\pi_y^{MR} \equiv \theta(\alpha + \gamma_2'\eta); \qquad \pi_{d_0}^{MR} \equiv \theta\gamma_1; \qquad \pi_{d_1}^{MR} \equiv \theta\gamma_2 \tag{10}$$

$$\phi_t^{MR} \equiv \theta\gamma_1'\mu_t + \theta\gamma_2'\mu_{t-1} + \theta\varepsilon_t, \tag{11}$$

where $\theta \equiv 1/(1 - \gamma_1'\eta)$. I call this the "**MR specification**," where "MR" comes from Mertens and Ravn (2012), who estimate (11) by OLS.[17] Under the two RR assumptions, if $\gamma_i \neq \gamma_1'$ the resulting estimates of the π_i^{MR}s are biased and inconsistent,

[16] RR estimate two versions of (8). The benchmark specification includes lags 0 to 12 of $d_{t/t}$ and no lagged endogenous variable. The second specification includes also lags 1 to 4 of y_t. They interpret both specifications as truncated versions of the MA representation; see below for a discussion.

[17] MR estimate a multi-dimensional version of (9), which however is conceptually the same as the specification used here for illustrative purposes. Their specification also includes expected changes to future discretionary taxation (see below).

because μ_{t-1} is correlated with y_{t-1}. One can obtain consistent estimates by taking μ_{t-1} out of the error term and including it in the r.h.s. of (9). An estimated series for μ_t can be obtained by instrumental variable estimation of (5): $d_{t-1/t-1}$ and y_{t-1} are natural instruments, as they are excluded variables from (5) that are correlated with y_t but uncorrelated with μ_t.

To fix ideas, I will call an OLS regression of y_t on y_{t-1}, $d_{t/t}$ and $d_{t-1/t-1}$ in (9) the "OLS estimate of the MR specification;" if the regressors also include the estimate of μ_{t-1}, I will call this, with some impropriety, "IV estimate of the MR specification."

V. Relation with Other Models

A. *Favero and Giavazzi* (*2012*)

FG estimate a VAR in y_t and s_t (plus the other endogenous variables, omitted here and in what follows for expository purposes) with $d_{t/t}$ as an exogenous term, and then trace the response to a shock to $d_{t/t}$:

$$y_t = \pi_{yy}^{FG} y_{t-1} + \pi_{yd}^{FG} d_{t/t} + \pi_{ys}^{FG} s_{t-1} + \phi_{y,t}^{FG} \tag{12}$$

$$s_t = \pi_{sy}^{FG} y_{t-1} + \pi_{sd}^{FG} d_{t/t} + \pi_{ss}^{FG} s_{t-1} + \phi_{s,t}^{FG}. \tag{13}$$

In terms of the coefficients and error terms of the model (5) and (6), the parameters and residuals estimated by FG are:

$$\pi_{yy}^{FG} \equiv \theta(\alpha + \gamma_2'\eta - \gamma_2\eta); \qquad \pi_{yd}^{FG} \equiv \theta\gamma_1; \qquad \pi_{ys}^{FG} \equiv \theta\gamma_2 \tag{14}$$

$$\pi_{sy}^{FG} \equiv \eta\theta(\alpha + \eta\gamma_2' - \eta\gamma_2); \qquad \pi_{sd}^{FG} \equiv 1 + \eta\theta\gamma_1; \qquad \pi_{ss}^{FG} \equiv \eta\theta\gamma_2 \tag{15}$$

$$\phi_{y,t}^{FG} \equiv \theta\gamma_1'\mu_t + \theta(\gamma_2' - \gamma_2)\mu_{t-1} + \theta\varepsilon_t; \tag{16}$$

$$\phi_{s,t}^{FG} \equiv (1 + \eta\theta\gamma_1')\mu_t + \theta\eta(\gamma_2' - \gamma_2)\mu_{t-1} + \eta\theta\varepsilon_t.$$

I will call (12) and (13) the "**FG specification**." Thus, the difference with the MR specification is that FG have s_{t-1} instead of $d_{t-1/t-1}$ in the output equation.

If equations (12) and (13) are estimated by an OLS regression of y_t and s_t on y_{t-1}, s_{t-1} and $d_{t/t}$, as in FG, once again the resulting estimates are inconsistent, because μ_{t-1} is correlated with both y_{t-1} and s_{t-1}. The first source of the correlation is the same as in the MR specification above. There is now also a second source, instead of using $d_{t-1/t-1}$ as in the true model, this specification uses s_{t-1}, which is obviously positively correlated with μ_{t-1}. As I show below, FG OLS impulse responses deliver consistently weaker negative output effects than MR OLS impulse responses.

Once again, consistent estimates can be obtained by taking μ_{t-1} out of the error term, leading to IV estimates of the FG specification. Note that when $\gamma_1' = \gamma_2' = 0$, from (9) MR OLS responses are consistent, while FG OLS responses continue to be inconsistent. When instead $\gamma_1 = \gamma_1'$ and $\gamma_2 = \gamma_2'$, it is FG OLS responses that are consistent: in (12) and (13) the term in μ_{t-1} disappears from the error term. Also, in this case the forecast error variance in the output equation is lower in the FG approach, for obvious reasons: given $\gamma_1 = \gamma_1'$ and $\gamma_2 = \gamma_2'$, there is no need to decompose s_{t-1} into the discretionary and the remaining component.

B. *Predetermined* $d_{t/t}$

Now assume that $d_{t/t}$ is not unpredictable. There are at least two reasons why $d_{t/t}$ can be predetermined. First, as pointed out by FG, by selecting those changes that were motivated by concerns about the level of debt, RR have automatically selected changes that are correlated with variables in the intertemporal government budget constraint. However, FG also show that in practice this does not seem to be an issue in this sample, because controlling for debt does not change the impulse responses appreciably. Second, the criterion adopted by RR to select exogenous changes to discretionary taxation might prove less than air-tight: policymakers might declare that they are solely concerned about the long-run deficit or the level of public debt, while in reality they are responding to a number of cyclical factors.

If $d_{t/t}$ is predetermined, one can fit a reaction function to it by estimating a VAR that includes $d_{t/t}$ as an endogenous variable. The true model would then consist of (5), (6) and

$$d_{t/t} = k_1 y_{t-1} + k_2 d_{t-1/t-1} + k_2'(s_{t-1} - d_{t-1/t-1}) + \zeta_t. \tag{17}$$

Hence the reduced form estimated in this specification, which I call the **"VAR specification,"** is

$$y_t = \pi_{yy}^{VAR} y_{t-1} + \pi_{yd}^{VAR} d_{t-1/t-1} + \phi_{y,t}^{VAR} \tag{18}$$

$$d_{t/t} = \pi_{dy}^{VAR} y_{t-1} + \pi_{dd}^{VAR} d_{t-1/t-1} + \phi_{d,t}^{VAR}. \tag{19}$$

In terms of the coefficients and error terms of the model (5) and (6), the parameters and residuals estimated in this specification are:

$$\pi_{yy}^{VAR} \equiv \theta(\alpha + \gamma_2'\eta + \gamma_1 k_1 + \gamma_1 k_2'\eta); \qquad \pi_{yd}^{VAR} \equiv \theta(\gamma_2 + \gamma_1 k_2); \tag{20}$$

$$\pi_{dy}^{VAR} \equiv k_1 + k_2'\eta; \qquad \pi_{dd}^{VAR} \equiv k_2; \tag{21}$$

$$\phi_{y,t}^{VAR} \equiv \theta(\gamma_1' + \gamma_1 k_2')\mu_t + \theta\gamma_2'\mu_{t-1} + \theta\gamma_1\zeta_t + \theta\varepsilon_t; \tag{22}$$

$$\phi_{d,t}^{VAR} \equiv k_2'\mu_{t-1} + \zeta_t.$$

Since by assumption $d_{t/t}$ does not respond to contemporaneous innovations, impulse responses are obtained from a Choleski decomposition in which $d_{t/t}$ is placed first, as in Blanchard and Perotti (2002).[18]

Note that in this specification too estimates of the VAR coefficients in general will be inconsistent, unless one takes μ_{t-1} out of the error terms. Thus, once again I will distinguish between OLS and IV estimates of the VAR specification. In the IV estimates, a regression of the residual of the y_t equation on the residual of the $d_{t/t}$ equation gives a coefficient of $\theta\gamma_1$, which is precisely the impact effect of a unit shock to $d_{t/t}$ in the MR specification.

C. *Romer and Romer* (*2010*)

From equation (9), by recursive substitution one can derive the truncated MA representation

$$y_t = \pi_{d_0}^{RR} d_{t/t} + \pi_{d_1}^{RR} d_{t-1/t-1} + \pi_{d_2}^{RR} d_{t-2/t-2} + \pi_y^{RR} y_{t-2} + \phi_t^{RR}, \tag{23}$$

where the coefficients and residual are:

$$\pi_{d_0}^{RR} \equiv \theta\gamma_1; \qquad \pi_{d_1}^{RR} \equiv \theta[\gamma_2 + \theta\gamma_1(\alpha + \gamma_2'\eta)]; \tag{24}$$

$$\pi_{d_2}^{RR} \equiv \theta^2\gamma_2(\alpha + \gamma_2'\eta); \qquad \pi_y^{RR} \equiv \theta^2(\alpha + \gamma_2'\eta)^2$$

$$\phi_t^{RR} \equiv \theta\gamma_1'\mu_t + \theta[\gamma_2' + \theta\gamma_1'(\alpha + \gamma_2'\eta)]\mu_{t-1} \tag{25}$$

$$+ \theta^2\gamma_2'(\alpha + \gamma_2'\eta)\mu_{t-2} + \theta\varepsilon_t + \theta^2(\alpha + \gamma_2'\eta)\varepsilon_{t-1}.$$

I call (23) the "**augmented RR specification**." RR (2010) omit the lagged endogenous variable y_{t-2}; I call (23) without the term y_{t-2} the "**RR specification**."

As FG note, omitting y_{t-2} can lead to inconsistent estimates if y_{t-2} is correlated with other terms in the truncated MA representation. We can see from (23) that this must be the case.[19]

However, the inclusion of y_{t-2} does not solve all problems, because once again μ_{t-2} is correlated with y_{t-2}. The solution here too is to take μ_{t-2} out of the error term, leading as usual to distinguishing between an OLS and an IV estimate of the augmented RR specification. In contrast, an IV estimate of the RR specification is still inconsistent, as the bias emanating from the exclusion of y_{t-2} persists.

Note that the inclusion of y_{t-2} does eliminate the inconsistency under FG's assumptions. When $\gamma_i = \gamma_i'$, the coefficient of $d_{t-1/t-1}$ in (23) is identical to the

[18] As Swanson (2006) points out, however, it is not clear how to interpret a shock to $d_{t/t}$ in this specification. This is the residual of a regression of the private sector's estimate of an innovation in discretionary taxation on lags of itself and other endogenous variables. It is even more difficult to interpret the impulse response to such a shock. Finally, it is inherently difficult to fit a reaction function to what one could interpret as a series of specific policy episodes; indeed, one could argue that the whole purpose of the RR exercise is to capture the policy shocks without having to fit a reaction function.

[19] The presence of this correlation is a small sample result, but one that can potentially be important in practice.

coefficient of μ_{t-1} in (24), and similarly for $d_{t-2/t-2}$ and μ_{t-2}. Hence, a regression of y_t on d_t, s_{t-1}, $s_{t-2,}$ and y_{t-2} (the truncated MA representation estimated by FG) gives consistent estimates.[20]

D. *Blanchard and Perotti* (*2002*)

Blanchard and Perotti (2002) estimate a VAR in y_t and s_t. In terms of the coefficients of the true model, they estimate

$$y_t = \pi_{yy}^{BP} y_{t-1} + \pi_{ys}^{BP} s_{t-1} + \phi_{y,t}^{BP} \tag{26}$$

$$s_t = \pi_{sy}^{BP} y_{t-1} + \pi_{ss}^{BP} s_{t-1} + \phi_{s,t}^{BP}. \tag{27}$$

Essentially, they estimate the FG specification (12) and (13), except that $d_{t/t}$ ends up in the error terms. They then construct a measure of the discretionary shock by computing the "discretionary" or "cyclically adjusted" tax residual, $\phi_{s,t}^{BP,CA} = \phi_{s,t}^{BP} - \eta\phi_{y,t}^{BP}$, which is equal to $d_{t/t} + \mu_t$. If $\gamma_i \neq \gamma_i'$, as usual the coefficients of the system (26) and (27) will be estimated consistently only if μ_{t-1} is taken out of the error term; still, the true impulse response will not be estimated consistently. To see this, note that the impact effect on output of a unit realization of $\phi_{s,t}^{BP,CA}$ is given by the coefficient $\hat{h}$ of the regression of $\phi_{y,t}^{BP}$ on $\phi_{s,t}^{BP,CA}$.[21] This gives

$$\hat{h} = \theta \frac{\gamma_1' Var(\mu_t) + \gamma_1 Var(d_{t/t})}{Var(\mu_t) + Var(d_{t/t})}. \tag{28}$$

If $|\gamma_1'| < |\gamma_1|$ this coefficient implies a smaller (in absolute value) impact multiplier than the correct one, which is $\theta\gamma_1$. This can explain the weaker effects of taxes on GDP estimated by Blanchard and Perotti (2002) relative to narrative estimates of tax shocks.[22]

[20] FG notice a large difference in the OLS impulse responses of (23), depending on whether the lagged endogenous variables y_{t-2} and s_{t-2} are omitted or included. They attribute this difference to $d_{t/t}$ being predetermined, so that there is a correlation between $d_{t/t}$ and $d_{t-1/t-1}$ on one hand and s_{t-2} and y_{t-2} on the other. However, this is not necessarily so: in small samples, the inclusion of s_{t-2} and y_{t-2} could have effects even if they were uncorrelated with $d_{t/t}$ and $d_{t-1/t-1}$.

An indication that this is indeed the case is the following. FG estimate a truncated MA representation using 12 lags of $d_{t/t}$, and lags 13 to 16 of s_t and y_t, and a second MA representation omitting lags 13 to 16 of the endogenous variables. They note that the two impulse responses start diverging precisely after about 12 quarters. This is when the effects of the lagged endogenous variables start kicking in, regardless of whether they are correlated with the shocks $d_{t/t}$ and its lags.

[21] Note the symmetry: In the RR, MR, and FG approaches, one starts from external estimates of $d_{t/t}$ and derives η as part of the estimation of the model. Blanchard and Perotti (2002) do not have estimates of $d_{t/t}$, but they use an external estimate of η to estimate the change in discretionary taxation $\phi_{s,t}^{BP,CA}$. The OECD computes the elasticity based on the tax codes and the distribution of taxable income in the population.

[22] See Mertens and Ravn (2011b) for an alternative explanation, based on an underestimate of the tax elasticity η in Blanchard and Perotti (2002).

VI. Specifications

I estimate and compare the specifications described above, using more realistic lag lengths and sets of endogenous variables. In all these specifications, $d_{t/t}$ is constructed as the dollar impact of a tax measure, divided by GDP.

1) The "**RR specification**:"

$$y_t = A(L)\,d_{t/t} + e_t, \tag{29}$$

where y_t is the log difference of real GDP per capita. As in RR, the order of the lag-polynomial $A(L)$ is 13 (that is, $A(L)$ includes powers 0 to 12 of the lag operator L).

2) The "**Augmented RR specification**:"

$$X_t = A(L)\,d_{t/t} + B(L)\,X_{t-12} + e_t \tag{30}$$

where $A(L)$ is of order 13 and $B(L)$ of order 4. Besides y_t, the vector X_t includes also the log change of real primary government spending per capita g_t, the first difference of the inflation rate $\Delta\pi_t$, and the first difference of the interest rate Δi_t.[23]

3) The "**MR specification**:"

$$X_t = A(L)\,d_{t/t} + B(L)\,X_{t-1} + e_t, \tag{31}$$

where $A(L)$ and $B(L)$ are of order 5 and 4, respectively.

4) The "**VAR specification**:"

$$X_t = B(L)\,X_{t-1} + e_t, \tag{32}$$

where X_t now includes also $d_{t/t}$ and $B(L)$ is of order 4.

5) The "**FG specification**:"

$$X_t = \alpha d_{t/t} + B(L)\,X_{t-1} + e_t \tag{33}$$

with $B(L)$ again is of order 4.

All specifications also include a constant, and are estimated by both OLS and IV. In the latter case, the set of regressors includes also the moving average (lags 1 to 4) of the series μ_t obtained by IV estimation of

$$s_t = d_{t/t} + \eta y_t + \delta_1 \Delta\pi_t + \delta_2 \Delta i_t + \delta_3 g_t + \mu_t. \tag{34}$$

The set of instruments in equation (34) includes lags 0 to 4 of $d_{t/t}$ and lags 1 to 4 of y_t, g_t, $\Delta\pi_t$ and Δi_t.

[23] These are the variables used by FG, except that they also include the log change of real government revenues.

Table 3—Elasticities

	RR data	My data
Unrestricted	1.79	1.79
	(7.53)	(7.40)
Restricted	1.80	1.74
	(7.57)	(7.23)

Notes: Estimates of elasticity in equation (34), t-statistics in parentheses. "RR data" from Romer and Romer (2010); "My data" include all changes in taxation ("LU + LA + NLA"). "Restricted:" the coefficient of $d_{t/t}$ is constrained to be 1.

I estimate all these specifications with the original RR data on $d_{t/t}$, and with the data I have assembled as described in Section III; in this latter case, the series on $d_{t/t}$ includes both the legislated changes and the non legislated changes. In the benchmark case, I use the version of these data that excludes the retroactive changes. In all cases I use only the exogenous changes as defined by RR (2010), that include deficit-driven and growth-driven tax changes in the terminology of these authors.

VII. The Effects of Taxes on GDP

The sample of RR's data on $d_{t/t}$ is 1947:I–2006:II; the sample of my data on $d_{t/t}$ is 1945:I–2009:IV, and includes up to the 2009 *Economic Recovery and Reinvestment Act*. The log change of GDP, government spending, and revenues per capita, all from the NIPA tables as in RR, are available from 1948:II;[24] the interest rate starts in 1947:I.[25] With four lags of the endogenous variables as instruments, the estimated series μ_t from (34) starts in 1949:II. There are 44 quarters with nonzero observations on $d_{t/t}$ in the RR data and 83 in my data. In both cases the p-value of the Ljung-Box Q test for serial correlation with 20 lags is always above 0.90; no partial correlation at any of lags 1, 2, 3, 4, 8 and 12 is ever significant.

Table 3 displays the estimates of the elasticity of revenues to GDP from equation (34). I estimate this equation in two versions, an unrestricted one and a restricted one, in which the coefficient of $d_{t/t}$ is forced to be 1. The difference between the two versions, and between the two datasets, is minimal: the estimate of η ranges between 1.74 and 1.80. As a comparison, the elasticity calculated by Blanchard and Perotti (2002) based on the elasticities provided by the OECD is 2.06 (their sample runs from 1946:I to 1997:IV).

Table 4 displays responses of the different models at six quarters and three years, two typical horizons of interest to policymakers. The initial impulse is an increase in discretionary taxes by 1 percentage point of GDP.

As a benchmark, columns 1 and 2 display results with the original RR data on $d_{t/t}$; row 1 is based on the RR specification. As in RR, the response at three years is

[24] The NIA income account data on the levels of these variables start in 1947:I, but in the FRED dataset the data on population starts in 1948:I.

[25] This series is defined as the average cost of servicing the debt, and it is constructed by dividing interest payments at time t by the federal government debt held by the public at time $t-1$. The two series are from FG. Note that at the numerator they use *net* interest payments from 1960:I. Conceptually, it is not clear that net interest payments are better than gross interest payments.

Table 4—Impulse Responses, OLS, and IV Specifications

		RR data, OLS		My data, OLS		My data, IV	
		6 qrts	12 qrts	6 qrts	12 qrts	6 qrts	12 qrts
		1	2	3	4	5	6
1	RR	−1.15*	−2.72**	−0.98*	−1.89**		
2	augm. RR	−1.33**	−1.39*	−1.65**	−1.57**	−1.73**	−1.59**
3	MR	−1.47*	−1.66**	−1.12*	−1.21*	−1.13*	−1.23*
4	VAR	−1.52*	−1.81**	−1.13*	−1.27*	−1.17**	−1.29**
5	FG	−0.59*	−0.72*	−0.21	−0.25	−1.16**	−1.66**

Note: Standard errors calculated by bootstrapping based on 1,000 replications.
**Significant at the 5 percent level.
*Significant at the 32 percent level.

extremely large, a decline of almost three percentage points. At the same horizon, still with the RR data on $d_{t/t}$, the effect is consistently weaker in the other specifications (rows 2 to 5): by about one percentage point in the augmented RR, MR and VAR specifications, and by two percentage points in the FG specification. The latter response is small, just about -0.7 percentage points of GDP. With my data on $d_{t/t}$ (columns 3 and 4) the responses tend to be even smaller: about -1.20 in the MR and VAR specifications, and virtually 0 in the FG specification.

Thus, as FG remark, a specification that takes into account the correct truncation of the MA representation, like theirs, would seem to lead to a much smaller tax multiplier than estimated by RR. However, if $\gamma_i \neq \gamma_i'$ an OLS estimation of the FG specification suffers from attenuation bias. In fact, consider now the IV estimates in columns 5 and 6 (because the IV estimate of the RR specification does not make sense, it is now omitted). The augmented RR, MR and VAR responses are virtually identical to the OLS estimates: as we have seen, a small difference between the OLS and IV responses of these specifications is consistent with γ_i' being small in absolute value. But now the FG response increases substantially, and is very close to the MR response: as we have seen, a large difference between the OLS and IV responses of the FG specification suggests that $\gamma_i \neq \gamma_i'$.

Thus, in all four IV specifications, the range of responses at three years is quite tight, between -1.25 and -1.65; these responses are mid-range between those estimated by RR and those estimated by FG.

Figure 1 presents some of the impulse responses also displayed in Table 4: the solid lines represent the MR IV specification with its 95 percent bootstrapped confidence bands computed from 1000 replications;[26] the various broken lines display (without confidence bands) the MR OLS specification, the RR specification, the FG OLS specification, and the FG IV specification. All use my data on $d_{t/t}$, except the RR specification that, as a benchmark, uses the RR data.

[26] Note that the confidence bands appear wider than in RR and in FG, but only because I use 95 percent confidence bands instead of 68 percent confidence bands.

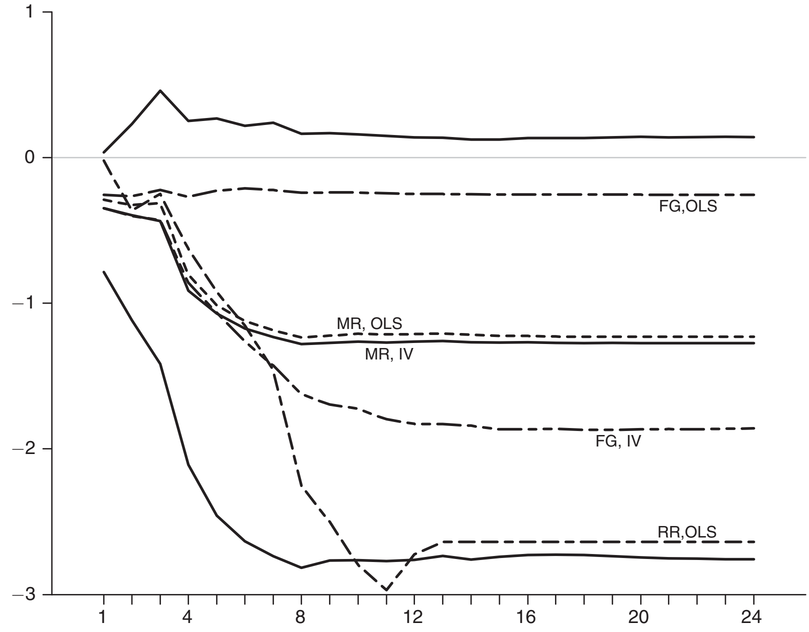

Figure 1. Impulse Responses, Main Specifications

VIII. Anticipated and Surprise Changes to Taxation

The specifications estimated so far do not allow for a distinction between *current* changes in discretionary taxation that come as a surprise and those that were already anticipated. In addition, individuals can react also to expected changes in *future* taxation: the specifications estimated so far do not allow for these.

To address these issues MR estimate by OLS an expanded version of the MR specification:

$$X_t = B(L)X_{t-1} + C(L)u_{t/t} + F(L)d_{t/t-1} + \sum_{i=1}^{K} G_i d_{t+i/t} + e_t, \tag{35}$$

where $d_{t/t}$ has been decomposed into the surprise change in current discretionary taxation, $u_{t/t}$, and the change in current taxation that was anticipated as of time $t-1$, $d_{t/t-1}$; in addition, the specification includes expected future changes to discretionary taxation, $d_{t+i/t}$. I assume $K = 6$ as in MR and set both $C(L)$ and $F(L)$ of order 5, and estimate the IV version of (35) (as usual, the difference with the OLS version is minimal).

The top two panels of Figure 2 show that the surprise and anticipated changes to current taxation, $u_{t/t}$ and $d_{t/t-1}$, have similar long-run negative effects on output, about -1.2 percentage points at three years, as in Table 4 (the figure also displays 68 percent confidence bands with solid lines and 95 percent confidence bands with broken lines). The similarity of these two responses speaks in favor of the presence of liquidity constraints in the economy. Figure 3 presents a more formal test of

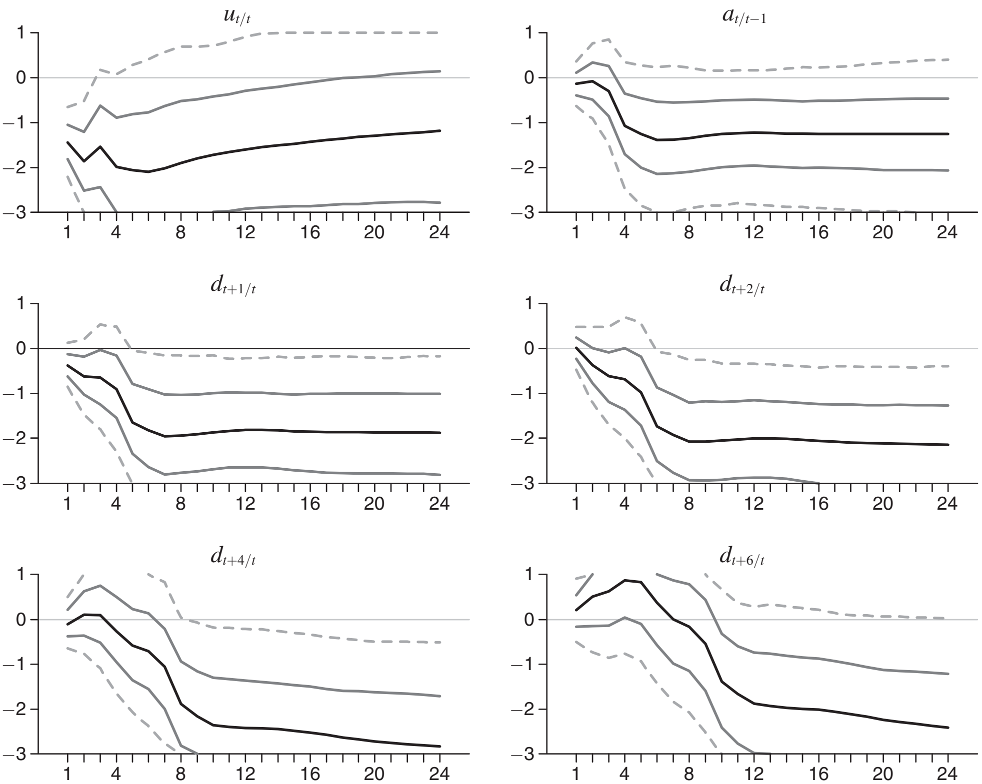

Figure 2. Impulse Responses from (35)

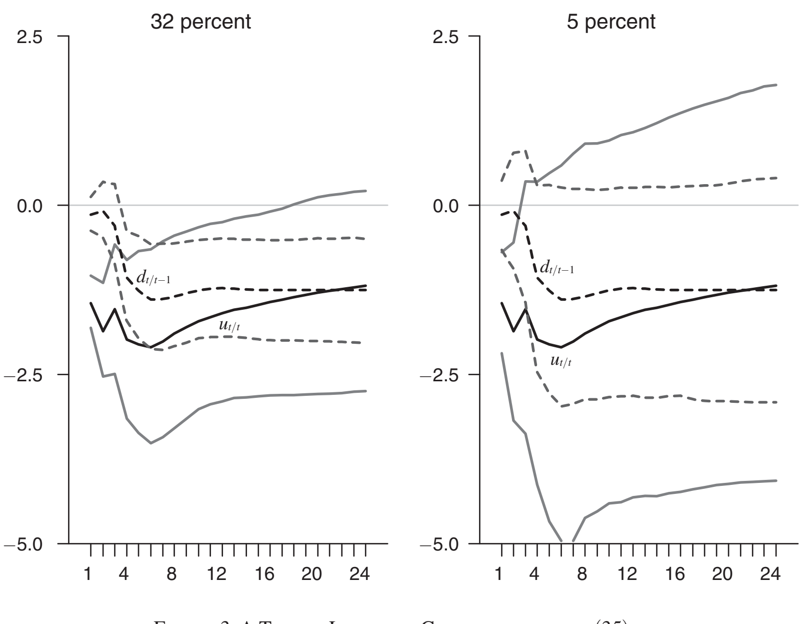

Figure 3. A Test of Liquidity Constraints, from (35)

liquidity constraints: it shows that each of the two responses, to $u_{t/t}$ and $d_{t/t}$, is within the confidence bands of the other.

MR also find large *positive* responses to a positive change in future expected taxation, i.e., to a unit change in $d_{t+i/t}$, for all values of i, with the largest response, 1.5 percentage points of GDP, at $i = 6$. This response speaks in favor of anticipation effects due to intertemporal substitution in a neoclassical model (see Mertens and Ravn 2011a for a formalization in a DSGE model). I find a positive response only at an anticipation horizon of $i = 6$, but it is relatively small and rather imprecisely estimated (see Figure 2). Thus, a dataset that tracks explicitly the starting date and the time path of the different tax items of each tax bill delivers a different response on anticipation effects.

One problem with the specification (35) is that the variables $d_{t+1/t}$ to $d_{t+6/t}$ include both surprise and anticipated changes to future discretionary taxation. Just as $d_{t/t}$ can be decomposed into the surprise component $u_{t/t}$ and the anticipated component $d_{t/t-1,}$ the same reasoning should be applied to changes to future taxation $d_{t+i/t}$, which can be decomposed into $u_{t+i/t}$ and $d_{t+i/t-1}$. This suggests estimating the IV version of:

$$X_t = B(L)X_{t-1} + C(L)u_{t/t} + F(L)d_{t/t-1} + \sum_{i=1}^{K} H_i u_{t+i/t} + \sum_{i=1}^{K} L_i d_{t+i/t-1} + e_t, \quad (36)$$

which derives from (35) after breaking down $d_{t+i/t}$ into $u_{t+i/t}$ and $d_{t+i/t-1}$. Figure 4 displays the responses to the anticipated and unanticipated components, $u_{t+i/t}$ and $d_{t+i/t-1}$, with 68 and 95 percent confidence bands. Here too there is evidence of a small positive response of output to a positive future tax change (both surprise and anticipated) only at an anticipation horizon of 6 quarters (and also 4 quarters for the anticipated component).

This specification still forces the dynamic responses to two different surprise changes, with different anticipation horizons, to share the same coefficients at different lags. This can be seen in Table 5, which lists all coefficients of all variables included in (36), with the components of each variable (for illustrative purposes, the table assumes $K = 2$, $M = 3$, while $C(L)$ and $F(L)$ are of order 3). For instance, the effects over time of a unit realization of $u_{t+2/t}$ are captured by the coefficient H_2, then by L_1, F_0, F_1, and F_2; the effects of a unit realization of $u_{t+1/t}$ are captured by the coefficient H_1, then by F_0, F_1, and F_2. Thus, the response to a given surprise change is still contaminated by the response to other surprise changes, with different anticipation horizons and different lags.

In order to address this issue, I estimate the IV version of:

$$X_t = B(L)X_{t-1} + \sum_{i=0}^{K}\sum_{j=0}^{J} H_i^j u_{t+i-j/t-j} + e_t, \quad (37)$$

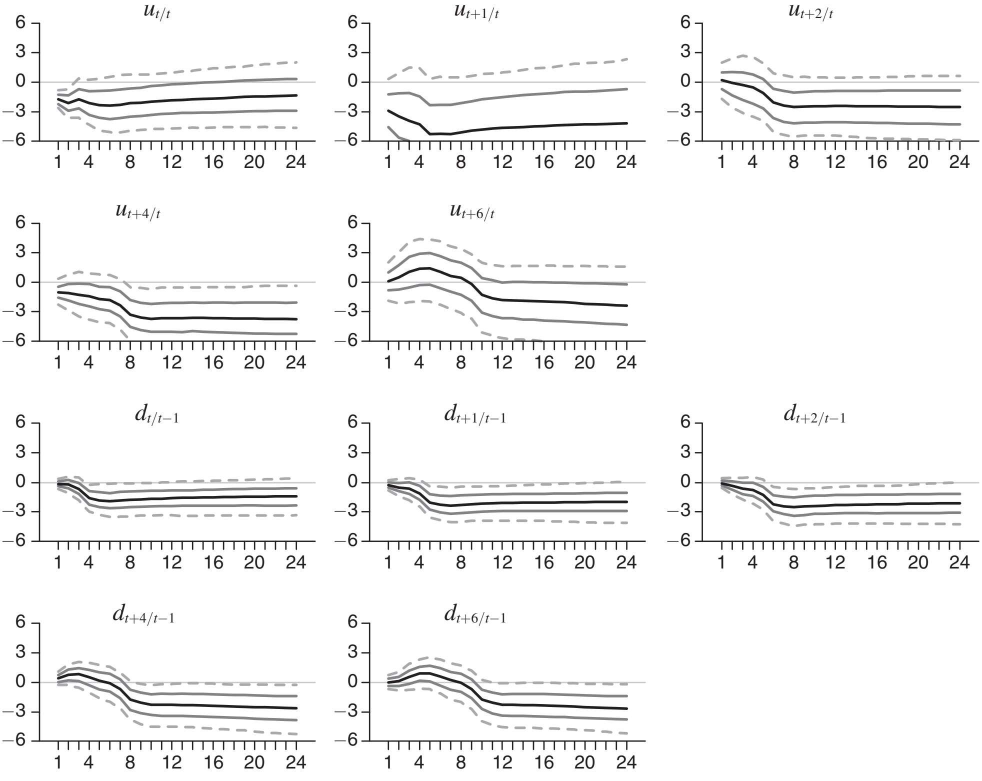

FIGURE 4. IMPULSE RESPONSES FROM (36)

which derives from (36) after breaking down $d_{t+i/t-1}$ into its components, the lagged revision to future expected taxation $u_{t+i/t-j}$ (see equation 3). Table 6 (which for illustrative purposes assumes $K = 2$ and $J = 2$) makes the point that different shocks to future discretionary taxation, with different anticipation horizons, are now allowed to have their own dynamics. For instance, the response to a unit realization to $u_{t+2/t}$ is captured by the coefficients H_0^2, H_1^2, and H_2^2, each associated only with a specific lag of $u_{t+2/t}$, and no other variable.

In estimation, I assume $K = 6$, $J = 4$, and a polynomial $B(L)$ of order 1 as in MR. Figure 5 displays the responses to surprise shocks at various anticipation horizons. Once again, there is evidence of a positive response of output to a positive shock to future taxation only at one anticipation horizon, $i = 4$, and with rather wide standard error bands.

Thus, using a dataset that tracks the quarterly patterns of shocks to each component of taxation in each tax bill, and separating the anticipated and the unanticipated components, leads to different conclusions about the effects of changes to future expected discretionary taxation: I find little evidence of a positive response of output to changes (whether anticipated or unanticipated) in future expected taxation. In evaluating these responses, though, it is important to bear in mind that the numbers of shocks at long anticipation horizons is small. For instance, there are only

Table 5—Composition of Surprise and Anticipated Changes in Taxation

(1)	$u_{t/t}$ (2)	$d_{t/t-1}$ (3)	$u_{t+1/t}$ (4)	$u_{t+2/t}$ (5)	$d_{t+1/t-1}$ (6)	$d_{t+2/t-1}$ (7)
Lag 0	$u_{t/t}$	$\underbrace{u_{t/t-1} + u_{t/t-2} + u_{t/t-3}}$	$u_{t+1/t}$	$u_{t+2/t}$	$\underbrace{u_{t+1/t-1} + u_{t+1/t-2}}$	$u_{t+2/t-1}$
Coeff. of lag 0	C_0	F_0	H_1	H_2	L_1	L_2
Lag 1	$u_{t-1/t-1}$	$\underbrace{u_{t-1/t-2} + u_{t-1/t-3} + u_{t-1/t-4}}$				
Coeff. of lag 1	C_1	F_1				
Lag 2	$u_{t-2/t-2}$	$\underbrace{u_{t-2/t-3} + u_{t-2/t-4} + u_{t-2/t-5}}$				
Coeff. of lag 2	C_2	F_2				

Note: Coefficients of the lag polynomials $C(L)$, $F(L)$, and the coefficients H_i and L_i in equation (37) are associated with each variable and the lag specified in column 1.

Table 6——Composition of Shocks to Future Taxation

Variable (1)	$u_{t/t}$ (2)	$u_{t+1/t}$ (4)	$u_{t+2/t}$ (5)
Lag 0	$u_{t/t}$	$u_{t+1/t}$	$u_{t+2/t}$
Coeff. of lag 0	H_0^0	H_0^1	H_0^2
Lag 1	$u_{t-1/t-1}$	$u_{t/t-1}$	$u_{t+1/t-1}$
Coeff. of lag 1	H_1^0	H_1^1	H_1^2
Lag 2	$u_{t-2/t-2}$	$u_{t-1/t-2}$	$u_{t/t-2}$
Coeff. of lag 2	H_2^0	H_2^1	H_2^2

Note: Coefficients H_i^j in equation (37) are associated with each variable and the lag specified in column 1.

7 quarters with nonzero values of each of $u_{t+4/t}$ and $u_{t+6/t}$; but in each case, only two shocks are larger than 0.1 percentage point of GDP.

IX. Robustness

Table 7 displays the results of a few robustness checks. For brevity, the table reports only the RR responses (with RR data) and the MR IV responses. In the first panel, the data are in levels, with a linear and a quadratic trend. As expected, these responses display less persistence than those in first differences. The next panel displays responses using data that include also the retroactive changes. The differences with the baseline estimates are small.[27]

The last panel displays responses over the two subsamples, 1950:I–1979:IV and 1980:I–2007:I. Perotti (2002) showed that the spending and tax multipliers seem to have decreased in the second subsamples in the US, UK, Canada, and Australia. FG also shows some evidence that this is the case for the tax multiplier in their specification. I find similar evidence only at 6 quarters, but not at three years.

As mentioned in Section I, it is not obvious how to interpret responses to changes in taxation. For instance, if the present discounted value of government

[27] Note, however, that now the RR data display significant negative serial correlation.

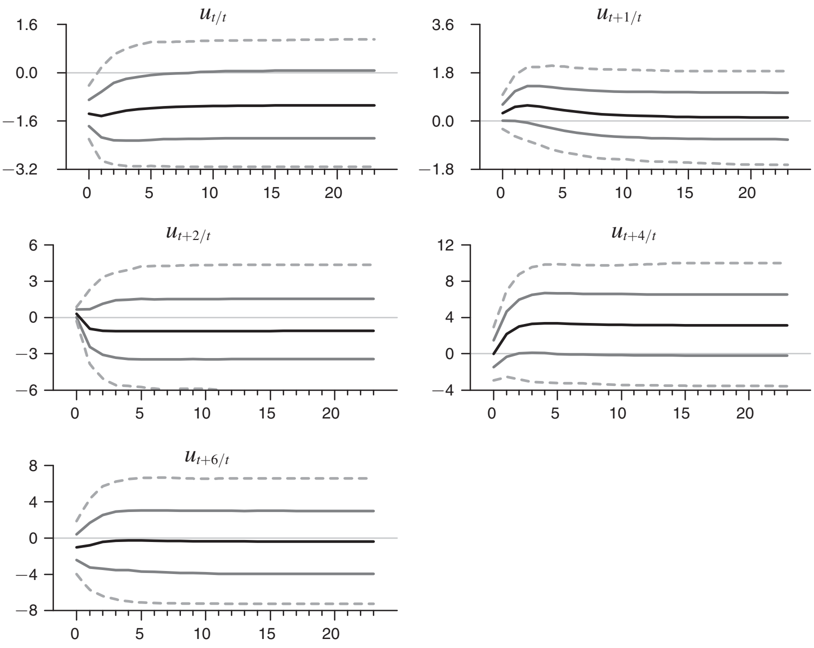

FIGURE 5. IMPULSE RESPONSES FROM (37)

spending changes together with taxation, then the coefficients G_is, H_is and L_is in equations (35), (36), and (37) do not just capture the intertemporal substitution effect of tax changes, but also the effects of the shock to wealth caused by the change in government spending. Controlling for the expected change in the present value of government spending on goods and services is difficult; including government spending in the VAR controls at most for the "average" response of government spending to a tax shock (see Cochrane 1998). I attempt to control for the response of the present value of expected government spending by including the present value of revisions to expectations of government spending on goods and services, using the Survey of Professional Forecasters.[28] Because data from this survey are available only from 1981:III, the resulting impulse response should be compared with that of the second subsample in Table 7; as it turns out, controlling for the expectation of government spending makes a minuscule difference (results not shown).

X. Conclusions

RR's seminal contribution has been criticized because it implies implausibly large negative effects of exogenous changes in taxation on GDP—a decline by three percent in response to a one percent of GDP increase in taxation. The contribution of

[28] In each quarter t, this survey presents forecasts up to four quarters ahead. I include forecasts both of federal and of state and local spending. See, however, Perotti (2011b) for a discussion of problems associated with this variable.

Table 7—Robustness

	My data	
	6 qrts	12 qrts
	Trends	
RR[1]	−1.33*	−1.18*
MR[1]	−1.23**	−0.64*
	Retroactive changes	
RR[2]	−1.09*	−2.80**
MR[2]	−1.18*	−1.26**
	Subsamples	
MR, 47:1–79:4	−1.87*	−0.82
MR, 80:1–09:4	−1.03*	−1.03*

Notes: MR specification is estimated in its IV version. 1: in levels, with linear and quadratic trends; 2: series for $d_{t/t}$ includes retroactive changes.
**Significant at the 5 percent level.
*Significant at the 32 percent level.

this paper is threefold. First, I introduce a novel dataset that expands on the RR dataset in several dimensions, including breaking down aggregate taxation into its major items, and tracking the starting date of implementation and the exogenous changes over time of each item.

Second, I argue that on theoretical grounds the discretionary component of taxation should be allowed to have different effects on output than the endogenous component (i.e., the automatic response of tax revenues to macroeconomic variables). Existing approaches to study the effects of the RR shocks do not allow for this difference. In particular, FG correctly argue that the specification estimated by RR is not a truncated MA representation of any process, and it is likely to be biased. They show that the impulse response from the correct MA representation delivers much smaller effects on GDP, in some cases insignificantly different from 0.

However, I show that when the discretionary and the automatic components of taxation have different effects on GDP, the FG impulse responses are biased towards 0. I derive a VAR model that accommodates the different impacts of the discretionary and endogenous components of taxation, and I then show that the impulse responses to a tax shock implied by this specification are about half-way between the large effects of RR and the much smaller effects of FG: typically, a one percentage point of GDP increase in taxes leads to a decline in output by about 1.3 percentage points after 12 quarters.

Third, I show that, using a dataset that explicitly keeps track of the starting dates and time path of changes to each item of each tax bill leads one to find no significant difference in the effects of anticipated and unanticipated changes in taxation, and virtually no statistically significant evidence of large positive effects of anticipated future tax changes. This evidence speaks in favor of the presence of liquidity constraints, or other factors that impede the response to anticipated changes in taxation.

REFERENCES

Baxter, Marianne, and Robert G. King. 1993. "Fiscal Policy in General Equilibrium." *American Economic Review* 83 (3): 315–34.

Blanchard, Olivier, and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *Quarterly Journal of Economics* 117 (4): 1329–68.

Blanchard, Olivier Jean. 1990. "Suggestions for a New Set of Fiscal Indicators." Organisation for Economic Co-operation and Development (OECD) Working Paper.

Cochrane, John H. 1998. "What Do the VARs Mean? Measuring the Output Effects of Monetary Policy." *Journal of Monetary Economics* 41 (2): 277–300.

Favero, Carlo, and Francesco Giavazzi. 2012. "Measuring Tax Multipliers: The Narrative Method in Fiscal VARs." *American Economic Journal: Economic Policy* 4 (2): 69–94.

Galí, Jordi, J. David López-Salido, and Javier Vallés. 2007. "Understanding the Effects of Government Spending on Consumption." *Journal of the European Economic Association* 5 (1): 227–70.

Leeper, Eric M., Alexander W. Richter, and Todd B. Walker. 2012. "Quantitative Effects of Fiscal Foresight." *American Economic Journal: Economic Policy* 4 (2): 115–44.

Mertens, Karel, and Morten Overgaard Ravn. 2012. "Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated US Tax Policy Shocks." *American Economic Journal: Economic Policy* 4 (2): 145–81.

Mertens, Karel, and Morten O. Ravn. 2011a. "Understanding the Aggregate Effects of Anticipated and Unanticipated Tax Policy Shocks." *Review of Economic Dynamics* 14 (1): 27–54.

Mertens, Karel, and Morten O. Ravn. 2011b. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States." Center for Economic Policy Research Discussion Paper 8554.

Perotti, Roberto. 2002. "Estimating the Effects of Fiscal Policy in OECD Countries." European Central Bank Working Paper 168.

Perotti, Roberto. 2011a. "The Effects of Tax Shocks on Output: Not So Large, But Not Small Either." National Bureau of Economic Research Working Paper 16786.

Perotti, Roberto. 2011b. "Expectations and Fiscal Policy: An Empirical Investigation." Growth and Sustainability Policies for Europe Working Paper 19.

Perotti, Roberto. 2012. "The Effects of Tax Shocks on Output: Not So Large, But Not Small Either: Dataset." *American Economic Journal: Economic Policy*. http://dx.doi.org/10.1257/pol.4.2.214.

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

Swanson, Eric. 2006. "Econometric Estimation When the 'True' Model Forecasts or Errors Are Observed." Unpublished. http://www.ericswanson.us/paper/etak.pdf.

This article has been cited by:

1. James Cloyne. 2013. Discretionary Tax Changes and the Macroeconomy: New Narrative Evidence from the United Kingdom. *American Economic Review* **103**:4, 1507-1528. [Abstract] [View PDF article] [PDF with links]
2. Karel Mertens,, Morten O. Ravn. 2013. The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States. *American Economic Review* **103**:4, 1212-1247. [Abstract] [View PDF article] [PDF with links]
3. Carlo Favero,, Francesco Giavazzi. 2012. Measuring Tax Multipliers: The Narrative Method in Fiscal VARs. *American Economic Journal: Economic Policy* **4**:2, 69-94. [Abstract] [View PDF article] [PDF with links]