The Effects of Tax Shocks: Negative, and Large

Roberto Perotti*

June 2010

PRELIMINARY AND INCOMPLETE

Abstract

In this paper, I argue that on theoretical grounds the discretionary component of taxation should be allowed to have different effects on output than the endogenous component, namely the automatic response of tax revenues to macroeconomic variables. Existing approaches to the study of the effects of the Romer and Romer shocks do not allow for this difference, and I show that as a consequence they exhibit impulse responses that are likely to be biased towards 0. On the other hand the RR specification, as Favero and Giavazzi (2009) correctly argue, is not a specification that can be derived from any representation of the data generating process. I derive a VAR model that can accommodate the different impacts of the discretionary and endogenous component of taxation, and I then show that the impulse responses to a RR shock implied by this specification are about half-way between the large effects of RR and the much smaller effects of Favero and Giavazzi: in general, a one percentage point of GDP increase in taxes leads to a decline in output by about 1.5 percentage points after 12 quarters.

The analysis of shocks to future taxation, instead, is complicated by two factors: the standard errors are extremely large, and the interpretation of the results is complicated once one recognizes that tax shocks can, and in general will, be associated with a negative wealth effect.

Keywords: Tax shocks, Tax multiplier.

JEL Classification Numbers: D91, E21, E62.

*IGIER, Università Bocconi, NBER and CEPR. Email: roberto.perotti@unibocconi.it. I thank Carlo Favero for several discussions that helped clarify the issues, and for pointing out a crucial mistake in an earlier version of the paper. This paper is produced as part of the project Growth and Sustainability Policies for Europe (GRASP), a Collaborative Project funded by the European Commission's Seventh Research Framework Programme, contract number 244725.

1 Introduction

In a seminal paper, Romer and Romer (2009) (henceforth, RR) construct measures of tax shocks from the original documents accompanying tax bills, and show that these shocks have large negative effects on output. Depending on the specification, an increase in taxes by 1 percentage point of GDP can lead to a decline in GDP by between 2 and 3 percentage points after 3 years.

These magnitudes appear implausibly large. Mertens and Ravn (2009) extend the analysis of RR by distinguishing between anticipated and unanticipated changes to taxation, and show that the two have different effects at short horizons. Favero and Giavazzi (2009) instead challenge the specification used by RR, arguing that it cannot be interpreted as a proper (truncated) moving average representation of the output process. When the system is estimated in its VAR form, or its correct truncated MA representation, a unit realization of the RR shock is shown to have much smaller effects on GDP than in RR.

In this paper, I first extend the RR dataset in several dimensions. Among other things, I track the quarterly changes in receipts emanating from each tax bill, and I distinguish between the different types of taxes. Then I argue that, on theoretical grounds, one should expect the discretionary component of tax changes to have stronger effects on output than endogenous changes due to the automatic, cyclical response of taxes to, say, output fluctuations. If one accepts this premise, then I show that the approaches of Favero and Giavazzi (and indeed even the Blanchard and Perotti (2002) approach) generate impulse responses that are likely to be biased towards zero. By adopting a VAR approach that can accomodate the RR shocks but at the same time allows for a different impact of discretionay and endogenous components of taxation, I show that the estimated effects of a tax shock are larger (in absolute value) than those estimated by Favero and Giavazzi, although smaller than in RR. Now a one percentage point of GDP increase in taxation is typically associated with a decline in GDP by about 1.5 percentage point after 12 quarters.

I then provide a test of the null that the effects of the discretionary and of the endogenous components of tax changes have different effects, as well as a test of the (implicit) RR assumption that individuals are liquidity constrained and therefore respond to tax changes only when implemented.

Finally, following Mertens and Ravn (2009) I study the response of output to announcements of future changes in taxation. I show that, once one removes the maintained hypothesis of Mertens and Ravn that tax changes have no wealth effects, it becomes extremely difficult to interpret the results.

2 Data description

I extend the RR data in several dimensions, and in some cases I use somewhat different rules to record the tax changes. In what follows, I detail the main features of my dataset and the main differences with the RR dataset.[1]

1. I collect data on total tax revenues, and also on their components. Specifically, I collect data on the following categories

Table 1: **Breakdown of taxes**

	Individual	**Corporate**	**Indirect**	**Soc. Sec.**
1.	Tax rates	Tax rates	Indirect taxes	Tax rate
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			

Not all categories have equally reliable data; but the distinction between the four main types of taxes is clear and meaningful. In this paper, I mostly aggregate all types of taxes in one variables, as RR do, although I do present some results with corporate and personal income taxes separately.

2. Whenever the source makes a distinction, RR use data on tax liabilities. For many questions, including questions involving liquidity constrained agents, receipts are more appropriate than

[1] The dataset will be posted on my website shortly.

liabilities. Differences between receipts and liabilities can reflect collection lags, tax elusion, tax evasion, and several other factors. Whenever possible, I collect data on both liabilities and receipts, although especially the earlier sources do not always allow a distinction between the two. In this paper, I use data on receipts. From the methodological comments in the sources, the estimates of receipts do not reflect behavioral responses any more than liabilities do.

3. RR typically report the effect of a tax legislation as the first full year effect of liabilities after enactment, and attribute that number to the quarter of enactment. In some cases, this procedure does not capture correctly the impact of tax legislation on tax receipts and liabilities. For instance, changes to depreciation allowances, to the investment tax credit and to capital gains taxation in particular can present a very irregular behavior over time. In some cases, changes in depreciation legislation consist only of accelerated depreciation, hence they change the time profile of receipts, but very little their present discounted value. In several instances, they display a large initial decline in receipts, then large positive changes in receipts. Using the first full-year effect would therefore provide a distorting picture of the effects of the tax measure.

I try to track quarterly changes in receipts and liabilities as closely as possible. Some sources not used by RR, most notably the *Survey of Current Business,* display the quarterly effects of the different tax measures. I complement these sources with a methodology to track quarterly changes, when possible and when quantitatively relevant.

I first keep track of changes in withholding rates for individuals. If these change at the time of enactment, I assume that receipts start being paid regularly at the time of enactment, unless receipt data indicate otherwise. If withholding rates are not changed immediately, some or all of the liability is paid in quarter 1 and 2 of the first next calendar year, when tax declarations are filed and net settlements are carried out.[2] In the first 10 or 15 years of the sample, typically the sources report only the full-year effect of the tax measures; but starting around 1960, they typically report receipts and liabilities over a long horizon, from 5 years to – in the nineties – up to 10 years ahead.

Often the effect on receipts in the first fiscal or calendar year after enactment is well below the

[2] When the source reports the quarterly pattern of receipts, this is indeed the pattern that one observes.

effect in later years, for reasons that can go from slow take up of a given measure to information lags to adjustment lags to tax collection lags. On the other hand, there is a normal trend increase in the effect due to the assumed exogenous increase in GDP over time; and one would not want to identify too many shocks in what is really a single shock. Hence, I adopt the rule that, if in FY x+1 receipts are different from receipts in fiscal year (FY) x by a factor of more than 30 percent, I display a change in x+1:Q1. By convention, the change is assumed to be in the first quarter, unless there is specific information that indicates a different quarter.

Like RR, I also keep track of all legislated changes after enactment. In some cases e.g. tax rates are legislated to change repeatedly after enactment, but not other items: in this case only part of the initial effect changes.

Thus, I end up with the following classification of tax changes, summarized in Table 2. A tax change is "legislated, unanticipated" if the tax change is legislated within 90 days after enactment, and tax receipts start within 90 days after the legislated tax change (row 1). It is "legislated, anticipated" if: either the tax change is legislated within 90 days from enactment, but receipts start more than 90 days after the legislated change (row 2.); or if the legislated change starts more than 90 days after enactment (rows 3 and 4). A tax change is "receipts, anticipated" if it is not associated with a legislated change, and follows form the application of the 30 percent rule.

RR use only legislated changes; Mertens and Ravn (2009) also use only legislated changes, but distinguish between anticipated and unanticipated.

Table 2: **Classification of discretionary tax changes**

1.	LC within 90 days after enactment, R change within 90 days from LC	UL
2.	LC within 90 days after enactment, R change more than 90 days after LC	AL
3.	LC more than 90 days after enactment, R change within 90 days from LC	AL
4.	LC more than 90 days after enactment, R change more than 90 days after LC	AL
5.	LC within 90 days after enactment, R change (without LC) after first R change	AR
6.	LC more than 90 days after enactment, R change (without LC) after first R change	AR

LC: "Legislated change"; *R*: "Receipts"; *UL*: "Unanticipated, legislated"; *AL*: "Anticipated, legislated; *AR*: "Anticipated, receipts.

4. Several tax changes have retroactive components, that is they apply to a period before the

time of enactment. RR assume that all retroactive liabilities are paid in one installments in the first quarter after enactment. In reality, individuals and corporations pay retroactive liabilities in a variety of ways.

Individuals typically pay retroactive liabilities in the first two quarters of the first calendar year after enactment, when filing tax returns, although different laws often specify different timings. As an example, suppose a law is signed on October 1 of year x, and it is retroactive to January 1 of year x; the withholding rates were changed immediately on enactment. The source reports an effect on receipts in FY x+1 of 1400 (FY x+1 starts on October 1 of year x),[3] thus contains 7 quarters worth of receipts: the three retroactive quarters x:Q1 to x:Q3, all paid in x+1:Q1 and x+1:Q2, and the four non-retroactive quarters x:Q4 to x+1:Q3. Hence, the retroactive component is 1400*(3/7), to be spread over two quarters; the annualized retroactive component is 1400*(3/7)*(4/2) = 1200, to be attributed to each of x+1:Q1 and x+1:Q2, while the annualized non-retroactive component is 1400*(4/7)*(4/4) = 800, to be attributed to each quarter starting x:Q3.

Suppose instead the tax measure was enacted on July 1 of year x, retroactive to January 1 of year x. Again withholding rates are adjusted immediately. Receipts in FY x contain one quarter's worth of receipts, the non-retroactive receipts paid in x:Q3. The retroactive part (two quarters' worth of receipts) is paid in x+1:Q1 and x+1:Q2. Hence we have to sum the effects from FY x and x+1: this sum contains again 7 quarters worth of data. Suppose this sum is again 1400. The retroactive part is now two quarters, spread over x+1:Q1 and x+1:Q2; hence it is 1400*(2/7)*(4/2) = 800; the non-retroactive part is 5 quarters, spread over 5 quarters: hence 1400*(5/7)*(4/5) = 800. The point of this second example is that sometimes one needs the sum of the effects in the first two fiscal years to compute the retroactive component.

The same rules are also used to allocate liabilities to the retroactive and non-retroactive components, in case only data on liabilities are available.

The case of corporations is more complicated. It depends first on the choice of the tax year by corporations. Most corporations (at present about 85 percent) choose the calendar tax year, and I will present results for this case. Presently, corporations are required to pay their tax year x

[3] Before 1975, a fiscal year started on July 1 of the previous calendar year.

estimated liabilities in four equal installments in year x. But these rules have changed over time. Until 1949 corporations paid 25 percent of their year x tax liability in each of the quarters of year x+1. In 1950 a new system was introduced, whereby corporations would move gradually to a payment of 50 percent of their year x tax liability in each of the first two quarters of year x+1. The transition lasted until 1954. But in 1954 a new system was again adopted: by 1959, a corporation would pay 25 percent of estimated tax liability for year x in quarters 3 and 4 of year x, and quarters 1 and 2 of year x+1. Any difference between estimated and actual tax liability would be paid or credited in March and June of year x+1. In the new system adopted in 1964, corporations would move slowly towards a system where they would pay 25 percent of their estimated year x liability in each quarter of year x+1. The transition to this new system was accelerated in 1966, so that the new system was fully operational in 1967.

Thus, there are four regimes, but each separated by a transition trajectory from the other. I will assume that regime 1 lasts until 1950 included, Regime 2 from 1951 to 1957, Regime 3 from 1958 to 1965, and Regime 4 from 1966 onwards. Table 3 summarizes the rules in place in different years, determining when a dollar of liabilities on corporate income received in year x would be paid. These rules are important to calculate both the correct time path of receipts and the retroactive components.

Take the case of the 1950 Revenue Act. It was enacted on September 23 1950, retroactive to July 1 1950. In calendar year (CY) 1951, calendar year corporations would pay their CY 1950 liabilities in four equal installments. The effect on full year liabilities was estimated to be 1500. I calculate the CY 1951 effect on receipts as follows. The retroactive component is one quarter, hence one fourth of 1500, to be paid over 4 quarters of CY 1951. Hence the quarterly annualized retroactive effect on receipts in CY 1951 is 375. The non-retroactive effect is also 375. From 1952:Q1, the effect on receipts is the full-year effect on liabilities, 1500.

Now take the case of the corporate tax rate increases in 1993 OBRA. It was enacted on August 10, 1993, and the measures were retroactive to January 1, 1993. The source does not report effects on receipts in FY 1993, but starts in FY 1994, that is 1993:Q4. The effects on receipts in FY 1994 is 4400. This includes 7 quarters worth of receipts: 3 quarters of retroactive effects, and 4 quarters of

Table 3: **Tax payments by corporations**

	Income year				**Following year**			
	April	June	Sept.	Dec.	April	June	Sept.	Dec.
1945					25	25	25	.25
1946					25	25	25	25
1947					25	25	25	25
1948					25	25	25	25
1949					25	25	25	25
1950					25	25	25	25
1951					30	30	20	20
1952					35	35	15	15
1953					40	40	10	10
1954					45	45	5	5
1955					50	50		
1956			5	5	50	50		
1957			10	10	45	45		
1958			15	15	40	40		
1059			20	20	35	35		
1960			25	25	30	30		
1961			25	25	25	25		
1962			25	25	25	25		
1963			25	25	25	25		
1964			25	25	25	25		
1965	9	9	25	25	16	16		
1966	12	12	25	25	12	12		
1967	25	25	25	25				

Each cell displays the percentage of the income earned in the "Income year" to be paid in the quarter indicated by the cell.

non-retroactive effects. Under the rules in place in 1993, calendar year corporations would have to pay changes in their 1993 tax liability in equal installments in the remaining quarters of CY 1993. Hence, in this case all the retroactive component would have to be paid in 1993:4. This implies a annualized retroactive effect on receipts in 1993:4 of 4400*(3/7)*4 = 75550. The non-retroactive component, that also starts in 1994:4, is 4400*(1/7)*4 = 2517.

Consider instead another example, from the 2002 Jobs Creation and Workers Assistance Act. This act was signed on March 9, 2002, and its provisions on depreciation applied to all capital put in place after September 10, 2001. Hence, it was retroactive by about two quarters. Although the retroactive component of depreciation provisions is estimated with a large uncertainty, following the pattern above one can proceed as follows. The retroactive component (2 quarters' worth of receipts) was received in three equal installments in 2002:Q2, 2002:Q3, and 2002:Q4. Over the same period a corporation would receive the three quarters of non-retroactive component of CY 2002. Because these quarters span two different fiscal years, one needs FY 2002 and FY 2003 estimates to compute the effects. The sum of FY 2002 and FY 2003 receipts contains 8 quarters' worth of receipts: 2 are retroactive and 6 are non-retroactive. The three retroactive quarters would be received in the last three quarters of CY 2003. Since the sum of FY 2002 and FY 2003 receipts is -35239-32738 = -67976, the annualized retroactive component is (-35239-32738)/8*2*(4/3) = -22659, in 1992:Q2, 1992:Q3 and 1992:Q4. The annualized non-retroactive component is (-35239-32738)/8*4 = -33989, in all quarters from enactment.[4]

3 Estimates of discretionary taxation

Narrative estimates of tax changes refer to changes in "discretionary" or "cyclically adjusted" taxation. Discretionary taxation is defined relative to some reference level of output. As such, it is not observed, but estimated by various agencies, some of which are the sources of the RR observations. In this section, I introduce notation and start from first principles to study the

[4] I have also collected data on the distributional impact of individual income taxes. I am using these data in a companion paper with Tommaso Monacelli.

notion of discretionary taxation. Some of the material covered in this section has been covered in Mertens and Ravn (23009), some is new.

To understand the concept of discretionary taxation, it is useful to start from the following question: what would tax revenues be if they changed only because of the automatic effects of changes in output? Denote the log of this hypothetical level of taxation by $\widetilde{S}_t$, and the logs of actual revenues and output by S_t and Y, respectively. Then

$$\widetilde{S}_t = S_{t-1} + \eta(Y_t - Y_{t-1}) + \mu_t \tag{1}$$

The difference between the actual revenues S_t and $\widetilde{S}_t$ is the change in discretionary taxation, i.e. the change in revenues we would observe if output remained constant at its reference level Y_{t-1}:

$$\begin{aligned} S_t - \widetilde{S}_t &= S_t - S_{t-1} - \eta(Y_t - Y_{t-1}) - \mu_t && (2) \\ &= D_{t/t} - D_{t-1/t-1} && (3) \end{aligned}$$

where $D_{i/j}$ is the log of discretionary taxation at date i as estimated at date j, and relative to the reference output Y_{t-1}.

Hence the actual change in revenues can be decomposed into the change in discretionary taxation plus a second component, comparing a random term and the automatic response of revenues to changes in output

$$S_t - S_{t-1} = D_{t/t} - D_{t-1/t-1} + \eta(Y_t - Y_{t-1}) + \mu_t \tag{4}$$

The standard procedure to estimate the change in discretionary taxation is to subtract form the actual change in revenues the change in output, and possibly other determinants of revenues, like inflation, multiplied by their elasticities. RR turn this procedure around by providing estimates of the change in discretionary taxation as estimated in official documents, and based on the specific provisions of each tax bill enacted by Congress.

Specifically, a law enacted at time t specifies a path of revisions of discretionary taxation, from time t onward: $D_{t/t} - D_{t/t-1}$, $D_{t+1/t} - D_{t+1/t-1}$, $D_{t+2/t} - D_{t+2/t-1}$ (of course, many of these revisions will be 0). Thus, there are many laws that specify revisions at time t: all the laws enacted from time $t-M$ to time t, where M is the maximum time horizon for a law.

Let $u_{t/t-i}$ be the revision of date t's log change in discretionary taxation relative to date $t-1$, caused by the law enacted at date $t-i$, i.e.

$$u_{t/t-i} \equiv (D_{t/t-i} - D_{t-1/t-i}) - (D_{t/t-i-1} - D_{t-1/t-i-1}) \tag{5}$$

(note that $u_{t/t} = D_{t/t} - D_{t/t-1}$ because $D_{t-1/t-1} = D_{t-1/t}$).

The key variable in the analysis of RR is the change in discretionary taxation between two consecutive periods, $D_{t/t}-D_{t-1/t-1}$, which I denote by $d_{t/t}$:

$$d_{t/t} \equiv D_{t/t} - D_{t-1/t-1} \tag{6}$$

This is equal to the sum of all revisions to date t's change in discretionary taxation, relative to date $t-1$, enacted by all laws between t and $t-M$: as such, $d_{t/t}$ captures the effect of the discretionary action of contemporaneous and past policymakers on the change in tax revenues between t and $t-1$, as opposed to the automatic effects of cyclical factors:

$$\begin{aligned}
D_{t/t} - D_{t-1/t-1} \quad &= \quad \underbrace{[D_{t/t} - D_{t/t-1}]}_{\mathbf{u}_{t/t}} + \underbrace{[(D_{t/t-1} - D_{t/t-2}) - (D_{t-1/t-1} - D_{t-1/t-2})]}_{\mathbf{u}_{t/t-1}} \\
&+ \quad \underbrace{[(D_{t/t-2} - D_{t/t-3}) - (D_{t-1/t-2} - D_{t-1/t-3})]}_{\mathbf{u}_{t/t-2}} + ... \\
&+ \quad \underbrace{[(D_{t/t-M} - D_{t/t-M-1}) - (D_{t-1/t-M} - D_{t-1/t-M-1})]}_{\mathbf{u}_{t/t-M}} \\
&= \quad \sum_{i=0}^{M} u_{t/t-i}
\end{aligned} \tag{7}$$

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

More generally, the change in date $t+j'$s discretionary taxation, expected at date $t-s$, is the the sum of all revisions to the changes in date $t+j$'s discretionary taxation, decided up to date

$t-s$

$$d_{t+j/t-s} \equiv D_{t+j/t-s} - D_{t+j-1/t-s} \tag{8}$$
$$= \sum_{i=0}^{M-s-j} u_{t+j/t-i-s} \qquad j+s \leq M \tag{9}$$

Obviously when $s = j = 0$ we have expression (7), given that $D_{t-1/t} = D_{t-1/t-1}$.

Note that in the expression for $D_{t/t} - D_{t-1/t-1}$, the first component $D_{t/t} - D_{t/t-1}$ (equal to $u_{t/t}$) is unanticipated, the rest is known at date t:

$$\begin{aligned} d_{t/t} &= \underbrace{u_{t/t}}_{\textbf{contemp. revision of change in } \mathbf{D}_{t/t}} + \underbrace{\sum_{i=0}^{M-1} u_{t/t-i-1}}_{\textbf{sum of all past revisions of change in } \mathbf{D}_{t/t}} \qquad (10) \\ &= u_{t/t} + d_{t/t-1} \end{aligned}$$

In fact, the second term on the rhs of (10), $d_{t/t-1}$, is the sum of all revisions to the change in discretionay taxation known at date $t-1$ and implemented at date t .

This allows us to understand that the RR observations are not strictly speaking tax "shocks" in the usual sense, because they contain a large anticipated component. In addition, they could easily be serially correlated. In fact,. $d_{t-1/t-1}$ contains terms like $u_{t-1/t-1}$ which is likely to be correlated with the term $u_{t/t-1}$ appearing in the definition of $d_{t/t}$: the same law approved in $t-1$ can set changes in discretionary taxation for t and $t-1$, and these are likely to be correlated. Empirically, $d_{t/t}$ is not serially correlated (as long as the retroactive component is not included), because the tax laws are far and few between.

4 Alternative models of the effects of discretionary taxation

To understand the differences between alternative estimation procedures, it is important to be clear about the underlying assumptions concerning the private sector behavior. A first distinction is between a model where only past and current changes in taxation matter, and a model where the private sector behavior is also influenced by anticipations of future changes in taxation. This distinction, of course, has been discussed on innumerable occasions. It is the same as the distinction between liquidity constrained agents and forward-looking, unconstrained agents. I will start with the RR assumption - only current and past changes in taxation matter - and then I will move on to discuss anticipation affects as in Mertens and Ravn (2009).

A second distinction has not been discussed in the literature, but it turns out to be crucial to an understanding of the estimation procedures: what are the effects of discretionary taxation and of the non-discretionary component? One can think of changes to $D_{t/t}$ in (4) as mostly changes to tax rates, rules about deductions, tax credits, depreciation, etc.; the remaining component of the change in S_t captures instead the automatic effects of deviations of output from its reference level, which occur without any intervention on the part of the policymaker.[5] There are at least three reasons why a change to the discretionay component should have a different effect on output: since it implies a change in tax rates, it is more distortionary; it is more persistent (for instance, if the reference level of output is potential output, deviations of the output gap from 0 should be temporary); and it can affect the reference level of output itself.[6]

The second distinction interacts with the first. Assume that individuals are liquidity constrained: their consumption depends on disposable income, hence on total revenues S_t. Assume instead that individuals are forward-looking with a long horizon, and that taxation is highly distortionary; their behavior will be strongly affected by changes in tax rates, that have wealth and various types of substitution effects; and little affected by the endogenous component, which is likely to be less persistent. Thus, in general the assumption that output depends on total revenues as opposed to discretionary taxation might be more natural when one also assumes that individuals are liquidity constrained.

A third distinction is whether discretionary taxation is exogenous or endogenous. Following RR, I will first assume that it is exogenous.

[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.

[6] Here too there is a slightly moot point. A change in tax rates could also affect the elasticity η, hence (in a non linear way) the automatic component too. This effect does not appear if taxes are proportional, and it is likely to be second order if taxes are only mildly progressive.

4.1 A small model with differentiated effects of discretionary taxation

To put is all together, I consider a minimalist model of output that however has all the ingredients one needs. I will assume that the "true" model includes an equation for the log change in tax revenues

$$s_t = d_{t/t} + \eta y_t + \mu_t \tag{11}$$

and an equation for the log change in output[7]

$$y_t = \alpha y_{t-1} + \gamma_1 d_{t/t} + \gamma_1'(s_t - d_{t/t}) + \gamma_2 d_{t-1/t-1} + \gamma_2'(s_{t-1} - d_{t-1/t-1}) + \varepsilon_t \tag{12}$$

Orthogonality of $d_{t/t}$ to μ_t and ε_t is the identifying assumption of RR. μ_t and ε_t are also structural shocks, as both contemporaneous values of the two endogenous variables s_t and y_t appear in both equations; hence μ_t and ε_t are orthogonal to each other. Obviously, however, μ_t is not orthogonal to $y_{t,}$, as the latter includes μ_t.

Thus, this specification allows the discretionary and the endogenous components of the change in total revenues to have different effects on output. If $\gamma_1 = \gamma_1'$ and $\gamma_2 = \gamma_2'$, equation (12) reduces to

$$y_t = \alpha y_{t-1} + \gamma_1 s_t + \gamma_2 s_{t-1} + \varepsilon_t \tag{13}$$

and output depends on total revenues. If at the other extreme $\gamma_1' = \gamma_2' = 0$, then equation (12) becomes

$$y_t = \alpha y_{t-1} + \gamma_1 d_{t/t} + \gamma_2 d_{t-1/t-1} + \varepsilon_t \tag{14}$$

and output depends only on discretionary taxation.

In the general case that $\gamma_1 \neq \gamma_1'$ and $\gamma_2 \neq \gamma_2'$, using (11), equation (12) becomes

$$y_t = \frac{\alpha + \gamma_2'\eta}{1 - \gamma_1'\eta} y_{t-1} + \frac{\gamma_1}{1 - \gamma_1'\eta} d_{t/t} + \frac{\gamma_2}{1 - \gamma_1'\eta} d_{t-1/t-1} + \frac{\gamma_1'}{1 - \gamma_1'\eta} \mu_t + \frac{\gamma_2'}{1 - \gamma_1'\eta} \mu_{t-1} + \frac{1}{1 - \gamma_1'\eta} \varepsilon_t \tag{15}$$

which can be estimated directly by regressing y_t on y_{t-1}, $d_{t/t}$ and $d_{t/t-1}$. I call this, for lack of a

[7] Obviously this is a simplified model; in the empirical application I allow for 4 lags of all the endogenous variables and of $d_{t/t}$.

better name, the "**P specification**", where "P" stands for "Perotti". Note that this is different from the specification estimated by RR, who omit the term in y_{t-1}; I will return to the RR specification below.

Favero and Giavazzi (2010) (FG henceforth) argue that one should estimate a VAR in y_t and s_t with $d_{t/t}$ as an exogenous term

$$y_t = \theta_1 y_{t-1} + \theta_2 s_{t-1} + \theta_3 d_{t/t} + \varphi_t^y \tag{16}$$

$$s_t = \beta_1 y_{t-1} + \beta_2 s_{t-1} + \beta_3 d_{t/t} + \varphi_t^s \tag{17}$$

and then trace the response to a shock to $d_{t/t}$. Again using (11) and (12), if $\gamma_1 \neq \gamma_1^{'}$ and $\gamma_2 \neq \gamma_2'$, by doing this one would end up estimating the following coefficients:

$$y_t = \frac{\alpha + \gamma_2'\eta - \gamma_2\eta}{1-\gamma_1^{'}\eta} y_{t-1} + \frac{\gamma_1}{1-\gamma_1^{'}\eta} d_{t/t} + \frac{\gamma_2}{1-\gamma_1^{'}\eta} s_{t-1} + \frac{\gamma_1^{'}}{1-\gamma_1^{'}\eta}\mu_t + \frac{\gamma_2'-\gamma_2}{1-\gamma_1^{'}\eta}\mu_{t-1} + \frac{1}{1-\gamma_1^{'}\eta}\varepsilon_t \tag{18}$$

$$s_t = \eta\,\frac{\alpha + \gamma_2'\eta - \gamma_2\eta}{1-\gamma_1^{'}\eta} y_{t-1} + \frac{1+\eta(\gamma_1 - \gamma_1')}{1-\gamma_1^{'}\eta} d_{t/t} + \frac{\eta\gamma_2}{1-\gamma_1^{'}\eta} s_{t-1} + \frac{1}{1-\gamma_1^{'}\eta}\mu_t + \frac{\eta(\gamma_2'-\gamma_2)}{1-\gamma_1^{'}\eta}\mu_{t-1} + \frac{\eta}{1-\gamma_1^{'}\eta}\varepsilon_t \tag{19}$$

where the last three terms in each equation make up the error terms of the equation. I will call the specification (18) and (19) the "**FG specification**".

As we will see, impulse responses from the FG specification deliver consistently smaller (in absolute value) effects than estimates from the P specification. There are two reasons why estimation of (18) and (19) could lead to an estimated impulse response that is biased towards 0.

Consider first the first two periods of the impulse response for y_t to a unit shock to $d_{t/t}$. From the P specification we have:

$$dy_t^P = \frac{\gamma_1}{1-\gamma_1^{'}\eta}; \qquad dy_{t+1}^P = \frac{\alpha + \gamma_2'\eta}{1-\gamma_1^{'}\eta}\frac{\gamma_1}{1-\gamma_1^{'}\eta} + \frac{\gamma_2}{1-\gamma_1^{'}\eta}; \tag{20}$$

From the FG specification instead

$$dy_t^{FG} = \frac{\gamma_1}{1-\gamma_1'\eta}; \qquad ds_t^{FG} = \frac{1+\eta(\gamma_1-\gamma_1')}{1-\gamma_1'\eta}; \tag{21}$$

$$dy_{t+1}^{FG} = \frac{\alpha+(\gamma_2'-\gamma_2)\eta}{1-\gamma_1'\eta}\frac{\gamma_1}{1-\gamma_1'\eta} + \frac{\gamma_2'}{1-\gamma_1'\eta}\frac{1+\eta(\gamma_1-\gamma_1')}{1-\gamma_1'\eta}; \tag{22}$$

It is easy to see that

$$dy_{t+1}^{FG} - dy_{t+1}^{P} = \frac{\eta\gamma_2(\gamma_1-\gamma_1')}{(1-\gamma_1'\eta)^2} \tag{23}$$

which is positive if $\gamma_2 < 0$, $\gamma_1 < 0$, $\gamma_1' < 0$, $\gamma_1 < \gamma_1'$

The second reason why the impulse response may be biased is a standard error - in - variable problem. It is clear that by replacing $d_{t-1/t-1}$ with s_{t-1} one of the regressors, s_{t-1}, becomes correlated with the error term. This biases the estimates of all coefficients, and as we know from standard analyses of the error-in-variable problem, in general it tends to bias the coefficients towards 0 (although this is not a theorem).

If $\gamma_1 = \gamma_1'$ and $\gamma_2 = \gamma_2'$, the P and the FG specifications give the same (unbiased) impulse responses, but the forecast error variance in the output equation is lower in the FG approach. The intuition is obvious: given $\gamma_1 = \gamma_1'$ and $\gamma_2 = \gamma_2'$, there is no need to decompose s_{t-1} into the discretionary and the endogenous components. More generally, if γ_2 is close to γ_2', the FG specification trades off a smaller forecast error variance in the output equation against some bias in the impulse responses.

It should also be clear that in general one would want to estimate a multidimensional system of equation, instead of just the output equations. The reasons are the usual ones, plus one specific to the present context. Suppose that tax revenues respond automatically not only to output, but also to inflation:

$$s_t = d_{t/t} + \eta y_t + \delta\pi_t + \mu_t \tag{24}$$

If there is no equation for inflation, the term $\delta\pi_t$ and its first lag would end up in the error term of the output equation.

4.2 Relation with Blanchard and Perotti

Blanchard and Perotti (2002) estimate the reduced form system

$$y_t = \rho_1 y_{t-1} + \rho_2 s_{t-1} + \chi_t^y \tag{25}$$

$$s_t = \sigma_1 y_{t-1} + \sigma_2 s_{t-1} + \chi_t^s \tag{26}$$

Essentially, BP estimate the FG specification (18) and (19), except that the terms in $d_{t/t}$ end up in the error terms. Hence:

$$\chi_t^y = \frac{\gamma_1}{1-\gamma_1^{'}\eta} d_{t/t} + \frac{\gamma_1^{'}}{1-\gamma_1^{'}\eta}\mu_t + \frac{\gamma_2^{\prime} - \gamma_2}{1-\gamma_1^{'}\eta}\mu_{t-1} + \frac{1}{1-\gamma_1^{'}\eta}\varepsilon_t \tag{27}$$

$$\chi_t^s = \frac{1+\eta(\gamma_1 - \gamma_1^{'})}{1-\gamma_1^{'}\eta} d_{t/t} + \frac{1}{1-\gamma_1^{'}\eta}\mu_t + \frac{\eta(\gamma_2^{\prime} - \gamma_2)}{1-\gamma_1^{'}\eta}\mu_{t-1} + \frac{\eta}{1-\gamma_1^{'}\eta}\varepsilon_t \tag{28}$$

BP then construct a measure of the "discretionary" shock by computing the "cyclically adjusted" tax residual, $\chi_t^{s,CA} = \chi_t^s - \eta\chi_t^y$. Clearly

$$\chi_t^{s,CA} = d_{t/t} + \mu_t \tag{29}$$

The impact effect on output of a unit realization of $\chi_t^{s,CA}$ is given by the coefficient of the regression $\chi_t^y = \zeta\chi_t^{s,CA} + \nu_t$. This gives

$$\widehat{\zeta} = \frac{\frac{\gamma_1^{'}}{1-\gamma_1^{'}\eta} Var(\mu_t) + \frac{\gamma_1}{1-\gamma_1^{'}\eta} Var(d_{t/t})}{Var(\mu_t) + Var(d_{t/t})} \tag{30}$$

If $|\gamma_1^{'}| < |\gamma_1|$ this coefficient implies a smaller (in absolute value) impact multiplier than the correct one. To this, one should add the same error in variable problem that is present in the FG approach. However, the FG approach has a smaller forecast error variance; if one knows the shocks of interest, there is no reason not to use them in estimation.

There is a second advantage of the FG approach over the BP approach: the former does not require knowledge of the elasticity η, while the latter uses η in constructing the cyclically adjusted tax residual. If η is constant, in the FG approach it is estimated together with the other coefficients.

Of course, if η is not constant, both approaches are misspecified, like all the other approaches except the P specification when $\gamma_1^{'} = \gamma_2^{'} = 0$ (from equation (15), η disappears from the estimated equation). The problem however is likely to be more serious in the BP approach, because at least the shocks $d_{t/t}$ utilized in the FG approach are computed by an agency using the changing elasticities over time.

4.3 Relation with Romer and Romer

In the general case when $\gamma_1 \neq \gamma_1^{'}$ and $\gamma_2 \neq \gamma_2^{'}$, from (15) one can estimate the truncated MA representation

$$y_t = \lambda_2 d_{t/t} + (\lambda_1\lambda_2 + \lambda_3) d_{t-1/t-1} + \lambda_1\lambda_3 d_{t-2/t-2} + \lambda_1^2 y_{t-2} + \tag{31}$$

$$\lambda_4 \mu_t + (\lambda_5 + \lambda_1\lambda_4)\mu_{t-1} + \lambda_1\lambda_5\mu_{t-2} + \lambda_6\varepsilon_t + \lambda_1\lambda_6\varepsilon_{t-1} \tag{32}$$

where

$$\lambda_1 \equiv \frac{\alpha + \gamma_2^{'}\eta}{1-\gamma_1^{'}\eta}; \qquad \lambda_2 \equiv \frac{\gamma_1}{1-\gamma_1^{'}\eta} \qquad \lambda_3 \equiv \frac{\gamma_2}{1-\gamma_1^{'}\eta}; \qquad \mu_t \equiv \frac{1}{1-\gamma_1^{'}\eta}\varepsilon_t \tag{33}$$

$$\lambda_4 \equiv \frac{\gamma_1^{'}}{1-\gamma_1^{'}\eta}; \qquad \lambda_5 \equiv \frac{\gamma_2^{'}}{1-\gamma_1^{'}\eta} \qquad \lambda_6 \equiv \frac{1}{1-\gamma_1^{'}\eta}; \tag{34}$$

Clearly the impulse response is the same as that of the P specification. RR do not estimate (31), though, but they leave out the lagged endogenous variable y_{t-2}; I call this the "**RR MA specification**".

As FG note, leaving out the lagged endogenous variable term can generate a bias in the impulse responses estimated by RR if y_{t-2} is correlated with the other terms in the truncated MA representation. We can see from the expression above that this must be the case: since $d_{t-2/t-2}$ enters the expression for y_{t-2}, these two terms are necessarily correlated, although this does not mean that $d_{t-2/t-2}$ is endogenous.

This is different in FG, because they choose a different route to get to the truncated MA representation. They start from the BP specification, and then assume a relation between the BP revenue equation error term and the RR shocks $d_{t/t}$. Assume that we start from the specification

$$y_t = \rho_1 y_{t-1} + \rho_2 s_t + \rho_r s_{t-1} + \omega_t^y \tag{35}$$

$$s_t = \sigma_1 y_{t-1} + \sigma_2 s_{t-1} + \omega_t^s \tag{36}$$

where the two error terms are uncorrelated. The truncated MA representation for y_t is

$$\begin{aligned} y_t \quad = \quad & \omega_t^y + \rho_2 \omega_t^s + (\rho_1 + \rho_2 \sigma_1)\omega_{t-1}^y + [(\rho_1 + \rho_2\sigma_1)\rho_2 + (\rho_3 + \rho_2\sigma_2)]\omega_{t-1}^s + \\ & + [(\rho_1 + \rho_2\sigma_1)^2 + (\rho_3 + \rho_2\sigma_2)]y_{t-2} + [(\rho_3 + \rho_2\sigma_2)(\sigma_2 + \rho_1 + \rho_2\sigma_1)]s_{t-2} \end{aligned} \tag{37}$$

FG the assume a simple relation between $d_{t/t}$ and ω_t^s

$$\omega_t^s = d_{t/t} + \varphi_t \tag{38}$$

where φ_t is white noise. From (37), and unlike in (31), it is no longer the case that y_{t-2} and s_{t-2} are correlated with the error terms.

Now consider a MA representation truncated at lag 12, as in RR. FG note that the truncated MA representation estimated by RR and the correct truncated MA representation (the equivalent of (37), which includes lags from 13 to 16 of the endogenous variables) start to differ approximately after about 10 quarters. They attribute this to correlation between the shock and lags of the endogenous variables higher than 4 quarters. However, this is not necessarily the case. Even if the RR shock were uncorrelated with all lags of the endogenous variable, the latter could well affect the impulse response, and precisely at about lag 13, when their effects start to kick in.

4.4 Endogenous RR shocks

Now assume that the RR innovations $d_{t/t}$ are not exogenous. As we have seen, this is different from the statement that the forgotten terms in the truncated MAR representation is correlated with some lags of $d_{t/t}$.

There are at least two reasons why $d_{t/t}$ can be endogenous. 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 a big concern. Second, and quite simply, the selection criterion of RR might prove less than air-tight. Policymakers might declare that they are solely concerned about the deficit or debt, while in reality they are responding to a number of cyclical factors.

If $d_{t/t}$ is endogenous, one can try to fit a reaction function on it. One obvious way of doing it is to estimate a VAR that includes $d_{t/t}$. Note that if $d_{t/t}$ is instead exogenous, this should do no harm: all coefficients on the rhs in the $d_{t/t}$ equation should be insignificantly different from 0.[8] I call this the "**P, VAR specification**".

As Swanson (20XX) points out, however, it is not clear how to interpret a shock to $d_{t/t}$ now. This is the residual of a regression of the private sector's estimate of an innovation in discretionary taxation. It is even more difficult to interpret the impulse response to such a shock. Third, it is inherently difficult to fit a reaction function to what one could interpret as a series of specific policy episodes; indeed, the whole purpose of the RR exercise is to capture the policy shocks without having to fit a reaction function.

5 Specifications

To put it all together, I estimate and compare several models, estimated with different data.

As a benchmark, I start from the "**RR MA specification**"

$$y_t = A(L)d_{t/t} + \varepsilon_t \tag{39}$$

As in RR, the order of $A(L)$ is 12.

As FG argue, to compute the correct truncated MA representation one should add to the rhs the lagged endogenous variables, starting with lag 13: this is the "**augmented MA specification**":

$$X_t = A(L)d_{t/t} + D(L)X_{t-13} + \varepsilon_t \tag{40}$$

[8] One could also imagine a reaction function on the endogenous component of tax revenues, i.e. on the elasticity. As mentioned above, this would crate a number of complications, including non-linearities.

where X_t is a vector of variables and $D(L)$ is of order 4. In one version, the vector X_t contains only y_t (the "small augmented MA specification"). Alternatively, the vector X_t includes also the log of real primary government spending g_t, the inflation rate π_t, and the interest rate i_t (the "large augmented MA representation"). These are the variables used by FG, except that they also include the log of real government revenues.

I then estimate the specification with the discretionary taxation as one of the regressors, assuming they are exogenous, i.e. the **"P specification"**:

$$X_t = B(L)\ X_{t-1} + C(L)\ d_{t/t} + \varepsilon_t \tag{41}$$

where both $B(L)$ and $C(L)$ are of order 4. Like before, I estimate this specification in two versions, "small" and "large".

I then estimate a specification with $d_{t/t}$ endogenous, i.e. the "**P VAR specification**"

$$X_t = B(L)\ X_{t-1} + \varepsilon_t \tag{42}$$

where X_t includes $d_{t/t}$ and y_t ("small P VAR specification") or $d_{t/t}$, y_t g_t, π_t, i_t ("large P VAR specification"), and $B(L)$ is of order 4.

Finally I estimate the "**FG specification**":

$$X_t = B(L)\ X_{t-1} + \alpha d_t + \varepsilon_t \tag{43}$$

in the two versions "small" and "large", and again with $B(L)$ of order 4.

I estimate all these specifications with three sets of data. The first is the original RR dataset; the second is the dataset that I have described in section 2, using only the legislated changes (as in the RR); the third is the same dataset, using also the large changes that were not specifically legislated. In all cases I use only the exogenous changes as defined by Romer and Romer.

6 Results

Table 4 displays the responses of the different models at 6 quarters and 3 years, two typical horizons of interest to policymakers (standard errors to follow). The sample is 1948:1 to 2009:4. The table displays significance at two levels of confidence: 32 percent (equivalent to 1 standard error bands on each side of the impulse response), denoted with a single star "*"; and 5 percent (two standard error bands), denoted with a double star "**".[9] Like in RR, the change in taxes are scaled by GDP, so that a unit change in $d_{t/t}$ is an increase in discretionary taxes by 1 percentage point of GDP. Like in most of RR's analysis, I exclude the retroactive component of taxation.[10]

Table 4: **Impulse responses, various specifications**

		RR data		P data, leg. changes		P data, all changes	
		6 qrts	12 qrts	6 qrts	12 qrts	6 qrts	12 qrts
		1	2	3	4	5	6
1	RR, MA	-1.17*	-2.74**	-.70*	-2.32**	-.78*	-1.81**
2	augm. MA, small	-1.60**	-1.74*	-1.69**	-2.19**	-1.47**	-1.60**
3	augm. MA, large	-1.29*	-1.35*	-1.52**	-2.06**	-1.31*	-1.47*
4	P, small	-1.16*	-1.12*	-1.32*	-1.18*	-1.10*	-1.00*
5	P, large	-1.40*	-1.65**	-1.99**	-2.43**	-1.51**	-1.58**
6	P, VAR, small	-1.19*	-1.25*	-1.34*	-1.42*	-1.17*	-1.14*
7	P, VAR, large	-1.45*	-1.83*	-1.97**	-2.57**	-1.55**	-1.66**
8	FG, small	-.07	-.07	-.09	-.08	-.18	-.17
9	FG, large	-.45*	-.49*	-.69*	-.77*	-.52*	-.63*

"*": significant at 32 percent level; "**": significant at 5 percent level.

There are five main conclusions to be drawn from Table 4. First, the data matter. In the RR MA specification (row 1)the response at 12 quarters falls by about 1 percentage point as one moves rightward, from the RR data (columns 1 and 2) to my data, only the legislated changes (columns 3 and 4) and then to my data, all changes (columns 5 and 6). In this last case, the response at three years is -1.8 percentage points.

[9] The use of one standard error bands in the literature on fiscal policy may go back to Blanchard and Perotti (2002). Be as it may, I believe there is no reason to use different standards than in the rest of the literature, hence in what follows I will also present results based on two standard error bands.

[10] However, with my methodology of allocating retroactive changes to receipts, the retraoctive component is less negatively serially correlated than in RR.

Second, adding the lagged values of the endogenous variables, as in the augmented MA representations (rows 2 and 3), does make a difference with the RR data, reducing the response at 12 quarters by about half (row 3, column 2), consistent with what FG find; but it makes much less difference with my data (columns 3 to 6). In fact, with my data, all changes (column 6) all three versions of the MA representation give a response at 12 quarters of about -1.5 percentage points.

Third, this is roughly also the response one obtains with the "P, large" specification and with my data, all changes and with the RR data (columns 2 and 6, row 5). The response is smaller (in absolute terms) with the "P, small" specification.

Fourth, the VAR versions of the P specification also deliver virtually the same multiplier at 12 quarters, - 1.5 percentage points of GDP in the large version and -1.14 percentage points of GDP in the small version. As noted by FG, this is consistent with the fact that $d_{t/t}$ is already virtually unpredictable.

Finally, the responses in the FG approaches are consistently smaller, and in fact practically zero in the small versions; in the larger version they are about -.5 percentage points of GDP (rows 6 and 7).

Thus, all the specifications that are consistent with the more general case $\gamma_1 \neq \gamma_1^{'}$ and $\gamma_2' \neq \gamma_2'$ give the same answer, a tax multiplier at 12 quarters of about -1.5 percentage points of GDP, which is also the same multiplier one obtains with the simple RR approach if one uses my data.

Although the breakdown between the different types of taxes is not the main focus of this paper, it is interesting to note that personal (including indirect and social security) and corporate income taxes seem to have very different effects. Table 5 displays the response to personal income taxes (columns 3 and 4) and to corporate income taxes (columns 5 and 6), as well as, for a comparison, the response to the aggregate of the two (columns 1 and 3), from columns 5 and 6 of the previous table.

The pattern of the responses to personal income tax shocks is nearly identical to the patterns of responses to aggregate taxation. The output response to a corporate income tax shock is uniformly much larger, although the standard errors are also now much larger, and the response is never significant at the 5 percent level.

Table 5: **Impulse responses, corporate and personal income taxes**

		all		personal		corporate	
		6 qrts	12 qrts	6 qrts	12 qrts	6 qrts	12 qrts
		1	2	3	4	5	6
1	RR, MA	-.78*	-1.81**	-.67*	-1.63**	-2.47*	-4.38*
3	augm. MA, large	-1.31*	-1.47*	-1.29*	-1.27*	-2.50	-4.62
5	P, large	-1.51**	-1.58**	-1.37**	-1.46**	-3.98*	-3.23*
9	FG, large	-.52*	-.63*	-.39*	-.56*	-1.81*	-1.58*

"*": significant at 32 percent level; "**": significant at 5 percent level.

6.1 A test of the null $\gamma_1 = \gamma_1'$ and $\gamma_2 = \gamma_2'$

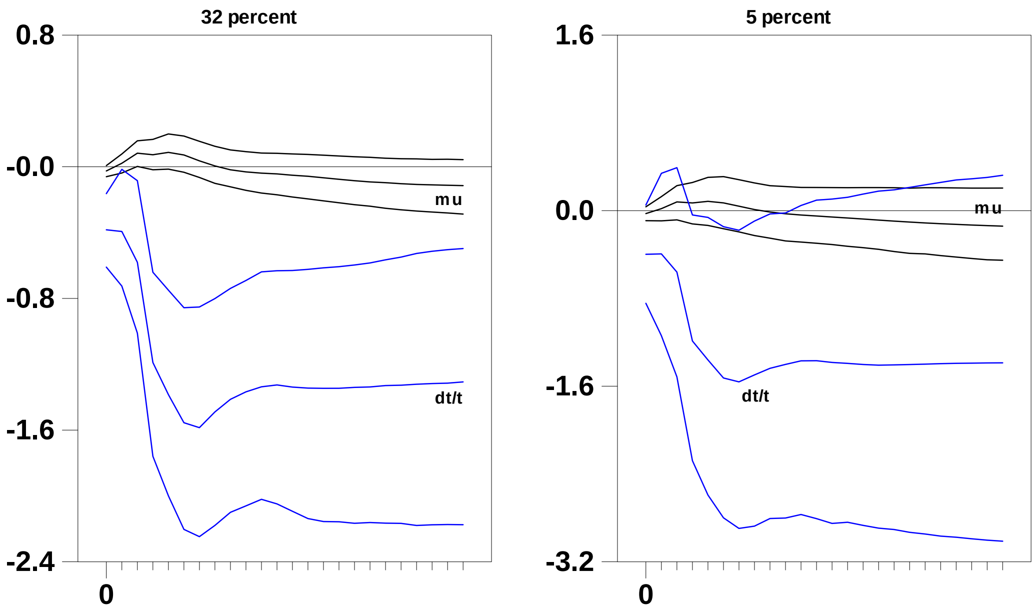
So far, I have merely shown that, if (i) $\gamma_1 \neq \gamma_1'$ and $\gamma_2 \neq \gamma_2'$, then the FG approach provides biased estimated of the impulse responses, and that if (ii) γ_2' is close to γ_2, the FG specification trades off a smaller forecast error variance in the output equation against some bias in the impulse responses. One might be interested in testing the null that $\gamma_1 = \gamma_1'$ and $\gamma_2 = \gamma_2'$.

Consider estimating (11) by instrumental variables. The model provides obvious instruments, as $d_{t.-1/t-1}$ and y_{t-1} are excluded variables from (11) that are correlated with y_t but uncorrelated with μ_t. From this estimation, one can obtain a series for μ_t, and use it in (15) to obtain estimates of $\gamma_1'/(1-\gamma_1'\eta)$ and $\gamma_2'/(1-\gamma_1'\eta)$. These can be compared to the coefficients of $d_{t/t}$ and $d_{t-1/t-1}$.

Performing this test has the usual problem that the coefficients of $d_{t/t}$ and its lags and of μ_t and its lags are estimated very imprecisely, individually. Hence, when testing the coefficients of $d_{t-i/t-i}$ against those of μ_{t-i} pairwise in, say, the "P, large" specification the two coefficients rarely come out significantly different from each other, and never when $i = 0$.

But one implication of the joint hypothesis that $\gamma_i = \gamma_i'$ pairwise is that impulse responses to $d_{t/t}$ and to μ_t are the same. Figure 1 displays such responses from the "P, large" specification, with one standard error bands (left panel) and 2 standard error bands (right panel). In fact, the response to μ_t is essentially 0, while the response to $d_{t/t}$ is negative and large, reaching a peak of -1.6 after about 6 quarters. With one standard error bands, the standard error bands are entirely apart. With two standard error bands, there is some (limited) overlap: at longer horizons, the response to μ_t falls below the upper band of the $d_{t/t}$ response; the $d_{t/t}$ response falls entirely below

P large



the lower band of the μ_t response.

It should be emphasized that this is a rather weak test. We are comparing the response to $d_{t/t}$ with the response to μ_t. But μ_t is itself a residual orthogonal to the automatic component ηy_t, hence one could argue that it has the dimension of a discretionary component.

7 Future expected taxation

In their regressions, RR use the contemporaneous and lagged value of $d_{t/t}$. As we have seen, this is consistent with the assumption that all individuals are liquidity constrained. But at each date t we have more information than this. There are several ways to slice this information, but a good starting point is:

$$D_{t+M/t} - D_{t-1/t-1} = \underbrace{(D_{t+M/t} - D_{t+M-1/t})}_{d_{t+M/t}} + \underbrace{(D_{t+M-1/t} - D_{t+M-2/t})}_{d_{t+M-1/t}} + ... \tag{44}$$

$$+\underbrace{(D_{t+1/t} - D_{t/t})}_{d_{t+1}/t} + \underbrace{(D_{t/t} - D_{t-1/t-1})}_{d_{t/t}} \tag{45}$$

(recall that M is the maximum horizon for a tax law).

Mertens and Ravn (2009) estimate a regression like

$$y_t = C(L)y_{t-1} + D(L)u_{t/t} + F(L)d_{t/t} + \sum_{i=1}^{K} G_i d_{t+i/t} + e_t \tag{46}$$

The rationale is that output and its components depend also on the expected slope of future taxation, because of the various intertemporal substitution effects; importantly, note that Mertens and Ravn *assume* that the RR shocks have no wealth effect, a point to which I will return later.

Just as $d_{t/t}$ can be decomposed into $u_{t/t}$ and $d_{t/t-1,}$ the same reasoning applies to changes of future taxation. In fact,

$$d_{t+i/t} = D_{t+i/t} - D_{t+i-1/t} = \sum_{j=0}^{M-i} u_{t+i/t} \tag{47}$$

and $d_{t/t}$ has an unanticipated and an anticipated component. In fact

$$\begin{aligned} d_{t+i/t} &\equiv D_{t+i/t} - D_{t+i-1/t} = \underbrace{(D_{t+i/t} - D_{t+i/t-1})}_{\textbf{surprise change in } \mathbf{D}_{t+i}} - \underbrace{(D_{t+i-1/t} - D_{t+i-1/t-1})}_{\textbf{surprise change in } \mathbf{D}_{t+i-1}} + \underbrace{(D_{t+i/t-1} - D_{t+i-1/t-1})}_{\textbf{anticipated change } \mathbf{D}_{t+i}} \\ &= u_{t+i/t} + d_{t+i/t-1} \end{aligned} \tag{48}$$

Hence the algebraic sum of the first and second term in (48) is $u_{t+i/t}$, the innovation in the expected change in D_{t+i}, the surprise change in the expected slope of taxation. The third term is the anticipated change in D_{t+i}, or $d_{t+i/t-1}$, the anticipated change in the slope of taxation.

Thus, I estimate the regression:

$$y_t = C(L)y_{t-1} + D(L)u_{t/t} + F(L)d_{t/t-1} + \sum_{i=1}^{K} H_i u_{t+i/t} + \sum_{i=1}^{K} L_i d_{t+i/t-1} + e_t \tag{49}$$

where, like in Mertens and Ravn, $K = 6$. The effects of a unit shock in t to taxation in $t + 6$ can be traced by assuming the following sequence of shocks: H_6, L_5, $L_4, L_3, L_2, L_1, D_0, D_1$, D_2, D_3, D_4. Instead, the effects of a contemporaneous, surprise change in taxation at t can be traced by assuming the sequence of shocks D_0, D_1, D_2, D_3, D_4.

When I do this, in general I find that a shock to future taxation first raises output and then depresses it; while a shock to contemporaneous taxation depresses output. This is similar to Mertens and Ravn, except that they estimate (46) instead of (49). However, the standard errors are extremely large; no impulse response (not reported) is significant, and the two responses do not differ significantly from each other.

Exactly how to interpret these results is not obvious. Mertens and Ravn *assume* that there is no wealth effect associated with a change in taxation at any time. However, it is easy to see that if this assumption fails, then the wealth effect confounds the intertemporal substitution effects that are purportedly captured by the coefficients G_i's, H_i's and L$_i$'s. In fact, the wealth effect is given by a term like

$$\sum_{i=0}^{\infty}(1+r)^{-i}\left[(y_{t+i/t}-y_{t+i/t-1})-(D_{t+i/t}-D_{t+i/t-1})\right] \tag{50}$$

i.e. by the revision of the expectation of the present discounted value of disposable income. The terms in $D_{t+i/t}-D_{t+i/t-1}$ are of the type

$$D_{t/t}-D_{t/t-1}=u_{t/t} \tag{51}$$

$$D_{t+1/t}-D_{t+1/t-1}=u_{t/t}+u_{t+1/t} \tag{52}$$

$$\begin{aligned} D_{t+i/t}-D_{t+i/t-1} &= \sum_{j=0}^{i} u_{t+j/t} \qquad i \leq M & (53)\\ &= \sum_{j=0}^{M} u_{t+j/t} \qquad i > M & (54) \end{aligned}$$

Thus, the terms $u_{t+i/t}$ appear repeatedly in these formula; hence, in the regression they pick up both the substitution and the wealth effects. This can easily confound the interpretation of several results. For instance, Mertens and Ravn find that the response of private consumption to a shock to future taxation is not significantly different from the response to a shock to current taxation.

They attribute this result to a prevalence of liquidity constrained individuals. However, if there is a wealth effect, this could be consistent with forward-looking, unconstrained behavior.

8 A test of liquidity constraints

Still, a test of liquidity constraints is available. Recall that $d_{t/t}$ can be written as the sum of the surprise change $u_{t/t}$ and of the anticipated change $d_{t/t-1}$.. The specifications displayed in Table 4 assume that (i) there is no forward-looking behavior and (ii) that the effect of surprise changes are the same as that of anticipated changes. However, one could test the second hypothesis, conditional on the first, by estimating (in the case of the large P specification)

$$X_t = B(L)\ X_{t-1} + C_u(L)\ u_{t/t} + C_a(L)\ d_{t/t-1} + \varepsilon_t \tag{55}$$

and then by testing whether the response to a shock to $u_{t/t}$ in (55) differs significantly from the response to $d_{t/t}$ in (55). The same exercise can be repeated in the other specifications. Essentially, this tests the notion that, under liquidity constraints, tax changes matter only when realized, whether they were anticipated or not.

Table 6 displays responses at 6 and 12 quarters, from the various specifications, to unit realizations of $u_{t/t}$ and $d_{t/t}$. As before, a * (**) indicates significance at 32 (5) percent confidence level. An "A" ("B") indicates that the response, say, to $d_{t/t}$ is outside both 68-percent (95 percent) standard error bands of the response to $u_{t/t}$. The opposite in the case of the response to $u_{t/t}$.

Once again, the response to $d_{t/t}$ at 6 quarters is negative and larger, in absolute value, than the response to $d_{t/t}$. However, except in the P, small specification, there is little evidence that the two responses are significantly different from each other, unless one uses 68 percent standard error bands.

9 Conclusions

In this paper, I argue that on theoretical grounds the discretionary component of taxation should be allowed to have stronger effects on output than the endogenous component, namely the automatic

Table 6: **A test of liquidity constraints**

		6 quarters		12 quarters	
		$d_{t/t}$	$u_{t/t}$	$d_{t/t}$	$u_{t/t}$
1	RR, MA	–.78*A	–.25 A	-1.81**	-2.29*
2	augm. MA, small	-1.47**B	-.31A	-1.60**	-1.00
4	P, small	-1.10*A	-.05	-1.00*A	-.07
5	P, large	-1.51**	-1.01	-1.58**	-1.10*
8	FG, small	–.18	-.44	-.17	-.41
9	FG, large	-.52*	-.72**	-.62*	-.77

"*": significant at 32 percent level; "**": significant at 5 percent level; "A": the response to $d_{t/t}$ is outside both 68-percent standard error bands of the response to $u_{t/t}$ and viceversa in the case of the response to $u_{t/t}$; "B": the response to $d_{t/t}$ is outside both 95-percent standard error bands of the response to $u_{t/t}$ and viceversa in the case of the response to $u_{t/t}$.

response of tax revenues to macroeconomic variables. Existing approaches to the study of the effects of the Romer and Romer shocks do not allow for this difference, and I show that as a consequence they exhibit impulse responses that are biased towards 0. On the other hand the RR specification, as Favero and Giavazzi (2009) correctly argue, is not a specification that can be derived from any representation of the data generating process. I derive a VAR model that can accommodate the different impacts of the discretionary and endogenous component of taxation, and I then show that the impulse responses to a RR shock implied by this specification are about half-way between the large effects of RR and the much smaller effects of Favero and Giavazzi: in general, a one percentage point of GDP increase in taxes leads to a decline in output by about 1.5 percentage points after 12 quarters.

The analysis of shocks to future taxation, instead, is complicated by two factors: the standard errors are extremely large, and the interpretation of the results is extremely complicated once one recognizes that tax shocks can, and in general will, be associated with a negative wealth effect.