An empirical characterization of the dynamic effects of changes in government spending and taxes on output

Olivier Blanchard and Roberto Perotti[*]

First version: March 1998
This version: April 2001

An empirical characterization of the dynamic effects of changes in government spending and taxes on output

1   Introduction

This paper characterizes the dynamic effects of shocks in government spending and taxes on economic activity in the United States during the postwar period. It does so by using a structural VAR approach that relies on institutional information about the tax and transfer systems and the timing of tax collections to construct the automatic response of fiscal policy to economic activity, and, by implication, to identify changes in fiscal policy.

The paper can be seen as doing for fiscal policy what a number of studies (in particular Bernanke and Mihov [1998]) have recently done for monetary policy. Indeed, the structural VAR approach would seem better suited to the study of fiscal policy than of monetary policy, for two reasons. First, fiscal variables move for many reasons, of which output stabilization is rarely predominant; in other words, there are exogenous (with respect to output) fiscal shocks. Second, in contrast to monetary policy, decision and implementation lags in fiscal policy imply that, at high enough frequency—say, within a quarter—there is little or no discretionary response of fiscal policy to unexpected contemporaneous movements in activity. Thus, with enough institutional information about the tax and transfer systems and the timing of tax collections, one can construct estimates of the automatic effects of unexpected movements in activity on fiscal variables, and, by implication, obtain estimates of fiscal policy shocks. Having identified these shocks, one can then trace their dynamic effects on GDP. This is what we do in this paper.

A methodological twist —one that was imposed on us by the data but is likely to be useful in other contexts—is that we combine this structural VAR approach with one akin to an event-study approach. A few episodes of large discretionary changes in taxation or in expenditures are simply too large to be treated as realizations from the same underlying stochastic process

1

and must be treated separately. We trace the effects of these large, one-time, changes by studying the dynamic response of output to an associated dummy variable that we include in the VAR specification. As we show when we look at the 1950s, not all such large fiscal events can be used so cleanly; when they can, as in the case of the 1975 temporary tax cut, we find a high degree of similarity between the impulse responses obtained by tracing the effects of estimated VAR shocks, and by tracing the effects of these special events.

Our results consistently show positive government spending shocks as having a positive effect on output, and positive tax shocks as having a negative effect. The size and persistence of these effects vary across specifications (for instance, whether we treat time trends as deterministic or stochastic) and subperiods; yet, the degree of variation is not such as to cloud the basic conclusion. Turning to the effects of taxes and spending on the components of GDP, one of our results has a distinctly non-Keynesian flavor: We find that both increases in taxes and increases in government spending have a strong negative effect on investment spending. (Keynesian theory, while agnostic about the sign, predicts opposite effects of tax and spending increases on investment.)

Our paper is organized as follows. Section 2 presents the main specification, and discusses identification. Section 3 presents the data and discusses their main properties, both at high and low frequencies. Section 4 discusses the contemporaneous relations between shocks to government spending, net taxes and output. Section 5 presents the dynamic effects of tax shocks. Section 6 does the same for spending shocks. Section 7 discusses robustness, including subsample stability, quarter dependence, cointegration, and the use of alternative net tax elasticities. Section 8 takes up the important issue of anticipated fiscal shocks, i.e. the possibility that, precisely because of implementation lags in policy, what the econometrician treats as unexpected shocks are in fact anticipated by private agents in the economy. Section 9 extends the sample to include the 1950a, and shows what can be learned from the Korean war buildup. Section 10 relates our results to those in the recent papers on the effects of fiscal policy, both those in the VAR and those in the event study tradition. Section 11 takes a first pass at a more disaggregated analysis, and presents the response of the individual output components. Section 12 concludes.

2   Methodological issues

Both government expenditure and taxation affect GDP: Since the two are presumably not independent, to estimate the effects of one it is also necessary to include the other. Hence, we focus on two-variable breakdowns of the budget, consisting of an expenditure and a revenue variable.

We define the expenditure variable as total purchases of goods and services, i.e. government consumption plus government investment. We call it "government spending", or simply "spending" for short. We define the revenue variable as total tax revenues minus transfers (including interest payments). We call it "net taxes", or "taxes" for short. We discuss the data in more detail in section 3.[1]

2.1   The VAR

Our basic VAR specification is

(1) $Y_t = A(L,q)Y_{t-1} + U_t,$

where $Y_t \equiv [T_t, G_t, X_t]'$ is a three-dimensional vector in the logarithms of quarterly taxes, spending, and GDP, all in real, per capita, terms.[2] We use quarterly data because, as we discuss below, this is essential for identification of the fiscal shocks. We allow either for deterministic (quadratic trends in logs), or stochastic (unit root with slowly changing drift) trends, but defer a discussion of this choice to the next section. We also allow for a number of dummy variables; again we defer a discussion of this issue to later.

$U_t \equiv [t_t, g_t, x_t]'$ is the corresponding vector of reduced form residuals, which in general will have non-zero cross correlations.

$A(L,q)$ is a four-quarter distributed lag polynomial that allows for the coefficients at each lag to depend on the particular quarter q that indexes the

---

[1] Any decomposition and choice of two fiscal variables reflects one's theoretical priors. Ours is no exception, and reflects our belief that, in the short run, fiscal policy works mainly through the effect of spending and taxes on aggregate demand and the effect of aggregate demand on output. A researcher who viewed fluctuations instead as real business cycles and believed in Ricardian equivalence, would likely choose a different two-variable decomposition, such as government consumption and government investment, or government spending and the marginal tax rate.

[2] We use the GDP deflator to express the variables in real terms. This allows us to express the impulse responses as shares of GDP. Results using the own deflator to express spending in real terms are very similar.

dependent variable. The reason for allowing for quarter-dependence of the coefficients is the presence of seasonal patterns in the response of taxes to economic activity. Some taxes—such as indirect taxes, or income taxes when withheld at the source—are paid with minimal delays relative to the time of transaction. Other taxes—such as corporate income taxes—are often paid with substantial delays relative to the time of the transaction; in addition, if the bulk of the payment is made in one or two specific quarters of the year, the delay varies depending on the quarter. Suppose, for illustrative purposes, that a tax is paid in the last quarter of each year, for activity over the year: Then, in the last quarter, the tax revenue will depend on GDP in the current and past three quarters; in the other three quarters, it will be equal to zero and thus will not depend on GDP.[3] We have collected evidence on quarter dependence in tax collection over the sample period from various institutional sources. The appendix lists the main relevant features of the tax code.[4]

2.2   Identification

We now discuss our identification methodology. Without loss of generality, we can write

$$
(2) \qquad t_t \;=\; a_1\,x_t + a_2\,e_t^g + e_t^t
$$

$$
(3) \qquad g_t \;=\; b_1\,x_t + b_2\,e_t^t + e_t^g
$$

$$
(4) \qquad x_t \;=\; c_1\,t_t + c_2\,g_t + e_t^x,
$$

where e_t^t, e_t^g, and e_t^x are the mutually uncorrelated structural shocks that we want to recover.

The first equation states that unexpected movements in taxes within a quarter, t_t, can be due to one of three factors: the response to unexpected movements in GDP, captured by $a_1\,x_t$, the response to structural shocks to spending, captured by $a_2\,e_t^g$, and to structural shocks to taxes, captured by e_t^t.

---

[3]Note that, because equation (1) is a reduced form, quarter dependence in the relation of taxes to GDP can show up in all three equations; we thus have to allow for quarter dependence in all three equations.

[4]A warning is in order here. In some cases the pattern of collection lags has changed—slightly—over time. Allowing for changes in quarter dependence in the VAR over time would have quickly exhausted all degrees of freedom. We have not done it; as a result, our adjustment is better than none, but still not quite right.

A similar interpretation applies to unexpected movements in spending in the second equation. The third equation states that unexpected movements in output can be due to unexpected movements in taxes, unexpected movements in spending, or to other unexpected shocks, e_t^x. Our methodology to identify this system can be divided into three steps.

(1) We rely on institutional information about tax, transfer and spending programs to construct the parameters a_1 and b_1. A priori, these coefficients could capture two different effects of activity on taxes and spending: The automatic effects of economic activity on taxes and spending under existing fiscal policy rules, and any discretionary adjustment made to fiscal policy in response to unexpected events within the quarter. The key to our identification procedure is to recognize that the use of quarterly data virtually eliminates the second channel. Direct evidence on the conduct of fiscal policy suggests that it takes policymakers and legislatures more than a quarter to learn about a GDP shock, decide what fiscal measures, if any, to take in response, pass these measures through the legislature, and actually implement them. The same would not be true if we used annual data: to a degree, fiscal policy can be adjusted in response to unexpected changes in GDP within the year.[5]

Thus, to construct a_1 and b_1, we only need to construct the elasticities to output of government purchases and of taxes minus transfers. To obtain these elasticities, we use information on the features of the spending and tax/transfer systems:

We could not identify any automatic feedback from economic activity to government purchases of goods and services; hence, we take $b_1 = 0$.

Turning to net taxes, write the level of net taxes, $\tilde{T}$, as $\tilde{T} = \sum \tilde{T}_i$, where the $\tilde{T}_i$'s are positive if they correspond to taxes, negative if they correspond

---

[5]Although the logic of our approach is similar to that of Bernanke and Mihov [1998] and Gordon and Leeper [1994], our identification methodology is different in one important respect. In those papers (which look at monetary, not fiscal policy), identification is achieved by assuming that private sector variables, like GDP and even interest rates (as in Gordon and Leeper [1994]), do not react to policy variables contemporaneously. By contrast, we assume that economic activity does not affect policy, except for the automatic feedback built in the tax code and the transfer system. Thus, our approach is more akin to that of Sims and Zha [1996], who have argued that cutting off the contemporaneous effects of policy on private sector variables might give a misleading view of the importance of policy innovations.

to transfers.[6] Let B_i be the tax base which corresponds to tax $\tilde{T}_i$ (or, in the case of transfers, the relevant aggregate for the transfer program, i.e. unemployment for unemployment benefits). We can then write the within-quarter elasticity of net taxes with respect to output, a_1, as

$$(5) \qquad a_1 = \sum_i \eta_{T_i,B_i} \; \eta_{B_i,X} \; \frac{\tilde{T}_i}{\tilde{T}},$$

where, η_{T_i,B_i} denotes the elasticity of taxes of type i to their tax base, and $\eta_{B_i,X}$ denotes the elasticity of the tax base to GDP.

To construct these elasticities, we extend earlier work by the OECD (Giorno et al. [1995]), which calculated output elasticities of four separate categories of taxes (direct taxes on individuals, corporate income taxes, social security taxes, and indirect taxes) using the formula above. The OECD elasticities must be modified because they are computed with respect to annual changes, while we need them with respect to quarterly changes. Both the elasticities of taxes to their base, the η_{T_i,B_i}s in expression (5), and the elasticities of the tax bases to GDP, the $\eta_{B_i,X}$s, can be quite different at quarterly and annual frequencies. The quarterly elasticities of tax revenues to their base can differ from the annual elasticities because of tax collection lags. In the United States, collection lags are probably relevant only for corporate income taxes (see Appendix). A more important source of divergence is the difference in the quarterly and annual elasticities of the tax bases with respect to GDP. For instance, the contemporaneous elasticity of profits to GDP, estimated from a regression of quarterly changes in profits on quarterly changes in GDP, is substantially higher than the elasticity obtained from a regression using annual changes —4.50 versus 2.15, respectively. The reverse is true of the elasticity of unemployment to output changes: it is, not surprisingly, lower when using quarterly changes rather than annual changes.

The value of a_1 we obtain following this procedure varies over time, both because the ratios of individual taxes and transfers to net taxes— the terms $\tilde{T}_i/\tilde{T}$ in expression (5) above—and the tax base elasticities of tax revenues— the terms η_{T_i,B_i}s—-have changed over time. The average value of our measure of a_1 over the 1947:1-1997:4 period is 2.08; it increases steadily from 1.58 in

---

[6]The tilde denotes the *level* of net taxes. We have used T earlier to denote the *logarithm* of net taxes per capita.

1947:1 to 1.63 in 1960:1 to 2.92 in 1997:4.[7,8] Using a different method, Cohen and Folette [2000], Cohenetal2000's work implies an estimate of (annual) a_1 ranging from 1.7 in 1951 to 2.2 in 1998.[9] The similarity of results, obtained with two different methods, is reassuring.

(2) With these estimates of a_1 and b_1, we construct the *cyclically adjusted* reduced form tax and spending residuals, $t'_t \equiv t_t - a_1 x_t$ and $g'_t \equiv g_t - b_1 x_t = g_t$ (as $b_1 = 0$). Obviously t'_t and g'_t may still be correlated with each other, but they are no longer correlated with e^x_t. Thus, we can use them as instruments to estimate c_1 and c_2 in a regression of x_t on t_t and g_t.

(3) This leaves two coefficients to estimate, a_2 and b_2. There is no convincing way to identify these coefficients from the correlation between t'_t and g'_t: when the government increases taxes and spending at the same time, are taxes responding to the increase in spending (i.e. $a_2 \neq 0, b_2 = 0$) or the reverse? We thus present the results under two alternative assumptions: Under the first, we assume that tax decisions come first, so that $a_2 = 0$ and we can estimate b_2; Under the second, we assume that spending decisions come first, so that $b_2 = 0$ and we can estimate a_2. It turns out that, in nearly all cases, the correlation between t'_t and g'_t is sufficiently small that the ordering makes little difference to the impulse response of output.[10]

---

[7]The fact that a_1 varies over the sample suggests time variation in the dynamic responses of spending and taxes to activity, and thus time variation of the VAR. Except for tests of subsample stability reported later, we have not explored this time variation further.

[8]One implicit assumption in our construction of a_1 is that the relation between the various tax bases and GDP is invariant to the type of shocks affecting output. For broad based taxes, such as income taxes, this is probably fine. It is more questionable, say, for corporate profit taxes: the relation of corporate profits to GDP may well vary depending on the type of shocks affecting GDP.

[9]We derive this number from their results as follows. We first regress the cyclical component of the budget (as a percent of GDP) from (their) Table 18 on the cyclical gap (as a percent of GDP) from (their) Table 13. The estimated coefficient is .34. We then divide this coefficient by the ratio of net taxes to GDP, which is around .2 at the start of the sample, and .15 at the end of the sample, to obtain the relevant elasticity.

[10]Let us discuss one more issue. As in all small-dimensional VARs, one should worry about omitted variables. Controlling for inflation in particular may be potentially important because all our variables are expressed in real terms, but spending on goods and services is typically budgeted in nominal terms, and personal income tax brackets are not indexed contemporaneously. For these reasons, inflation shocks are likely to affect both spending and taxes.

7

2.3 Impulse responses

Having identified the tax and spending shocks, we can study their effects on GDP. We face two issues here. The first is that the constructed elasticity of net taxes to output, a_1, varies over time. We ignore this here and compute the impulse response using the mean value of a_1 over the sample, i.e. 2.08. Second, one of the implications of quarter dependence is that the effects of fiscal policy vary depending on which quarter the shock takes place. One way to deal with this would be to derive four impulse responses, depending on the quarter where the initial shock occurs. This however would be cumbersome. We adopt a simpler procedure. We use a quarter dependent VAR to obtain the estimated covariance matrix, and thus the coefficients in equation (3) and the contemporaneous effects of fiscal shocks. We then use a VAR estimated without quarter dependence (except for additive seasonality) to characterize the dynamic effects of the shocks. Thus, our impulse responses give, admittedly only in a loose sense, the average dynamic response to fiscal shocks.

3 The data

We define net taxes as the sum of *Personal Tax and Nontax Receipts*, *Corporate Profits Tax Receipts*, *Indirect Business Tax and Nontax Accruals*, and *Contributions for Social Insurance*, less *Net Transfer Payments to Persons* and *Net Interest Paid by Government*. Government spending is defined as *Purchases of Goods and Services*, both current and capital. The source is the Quarterly National Income and Product Accounts, with the exception of *Corporate Profits Tax Receipts*, which are obtained from the Quarterly Treasury Bulletin. All these items cover the general government, i.e. the sum of the federal, state and local governments, and social security funds.[11] All the data are seasonally adjusted by the original source, using some variant

---

If activity shocks affect inflation within the quarter, then our constructed values of a_1 and b_1 will be incorrect, as they ignore the effect of activity through inflation on taxes and spending. However, in our data, the correlation between quarterly inflation and real GDP shocks is small (0.01) so this does not appear to be a major issue. We have not however explored this further.

[11]We do not have data on the corporate profit tax on a cash basis for state and local governments. This represents about 5 percent of total corporate profit tax receipts at the beginning of the sample, and about 20 percent at the end.

of the X-11 method.[12]

3.1 High-frequency properties

Figure 1 displays the behavior of the ratio of government spending (i.e. purchases of goods and services) and of net taxes (i.e. taxes minus transfers) to GDP over the longest available sample, 1947:1 to 1997:4. It is obvious that these series display a few extremely large quarterly changes in taxes and spending, well above three times their standard deviations of 4.3 percent and 1.9 percent, respectively.

There are two particularly striking changes in net taxes. First, the increase in 1950:2 by about 26 percent (more than 6 times the standard deviation), followed by a further increase in 1950:3 by about 17 percent (or about 4 times the standard deviation). These episodes represent in part the reversal of the temporary 8 percent fall in net taxes in 1950:1, caused by a large one-time payment of National Service Life Insurance benefits to the war veterans; but mostly they represent a genuine increase in tax revenues. The second episode is the large temporary tax rebate of 1975:2, which results in a net tax drop by about 33 percent for one quarter.[13]

On the expenditure side, the Korean war stands out: in 1951:1, at the onset of the Korean war buildup, government spending increases by 10 percent (more than 5 times the standard deviation), and continues to grow at about the same quarterly rate during the next two quarters, 1951:2 and 1951:3; after this, spending continues to increase at more than twice the standard deviation in 1951:4 and again in 1952:2. It is difficult to think of the early 1950s as being generated by the same stochastic process as the post-1960 period. Thus, our strategy is to proceed in two steps. For most of the paper, we run regressions starting from 1960:1. In section 8, we extend the sample back to include the 1950s and look at what we can learn from this longer sample.

---

[12]Corporate profit tax receipts are only reported without a seasonal adjustment. We use the RATS EZ-X11 routine to seasonally adjust this series.

[13]See Blinder [1981] for a detailed analysis of this tax cut, and its effects on consumption. The 1975:2 tax rebate (which was combined with a social security bonus for retirees without taxes to rebate) corresponded to an increase in disposable income of about $100 billion at 1987 prices; by comparison, the 1968 surtax decreased disposable income by $16 billion and the 1982 tax cut increased disposable income by $31.6 billion, always at 1987 prices (see Poterba [1988]).

Note that our benchmark sample still includes one large net tax cut episode, the 1975:2 tax cut. This episode is a well-identified, isolated, temporary, tax cut which was reversed after one quarter. Hence, it can be easily and clearly dummied. This allows us to compare two different types of impulse responses to a net tax shock: one tracing the dynamic effects of the estimated net tax shocks (our first approach), the other tracing the effects over time of a unit shock to the 1975:2 dummy variable (our second approach).

3.2   Low-frequency properties

The general visual impression from Figure 1 is one of no clear trends, but clear low–frequency (say decade to decade) movements in both spending and taxes. One may be surprised by the general absence of upward trends in spending and net taxes; but recall that we are looking at government spending *not including transfers*, and that net taxes are taxes *net of transfers*. Thus, the figures hide the trend increases in taxes and transfers, which have indeed taken place during this period.

The main practical issue, for our purposes, is how to treat these low-frequency movements in our two fiscal series in relation to output. We have conducted a battery of integration tests for T, G and X. Formal tests (Augmented Dickey-Fuller and Phillips-Perron, with a deterministic time trend) do not speak strongly on whether we should assume stochastic or deterministic trends for each variable. We have also conducted a battery of cointegration tests. One obvious candidate for a cointegration relation is the difference between taxes and spending, $T - G$:[14] In fact, the stationarity of the deficit is the basic idea underlying the tests of "sustainability" of fiscal policy by Hamilton and Flavin [1986] and Bohn [1991].[15] Figure 2 displays the logarithm of the tax/spending ratio. Again, formal test results do not speak strongly: one can typically reject the null of a unit root at about the 5% level, but no lower.[16]

In light of these results, we estimate our VARs under two alternative

---

[14]Recall that T is net of interest payments.

[15]Note that we test cointegration between the *logarithms* of taxes and spending. This is equivalent to testing cointegration between the logarithms of the net tax/GDP and of the spending/GDP ratios.

[16]This lack of strong evidence for cointegration between T and G is consistent with a number of recent empirical studies: see e.g. Bohn [1998], who tests for cointegration over the 1916-95 period using annual data.

assumptions. In the first, we formalize trends in all three variables as deterministic, and allow for linear and quadratic terms in time in each of the equations of the VAR. In the second, we allow for stochastic trends. We take first-differences of each variable, and, to account for changes in the underlying drift terms, we subtract a changing mean, constructed as the geometric average of past first differences, with decay parameter equal to 2.5 percent per quarter (varying this parameter between 1 and 5 percent makes little difference to the results).[17] For brevity, in what follows we will refer to the two specifications as 'DT' (for 'deterministic trend') and 'ST' (for 'stochastic trend'), respectively. In both specifications, we allow for the current value and four lags of a dummy for 1975:2.

In our benchmark specifications, we do not impose a cointegration restriction between the tax and the spending variable. In Section 7, we discuss results under the restriction that T and G are cointegrated. As it turns out, this makes little difference to the results.

From now on, unless otherwise noted (in particular in section 8), our results are based on the 1960:1-1997:4 sample.

4   Contemporaneous effects

The two panels of Table 2 report the estimated coefficients of the contemporaneous relations between shocks in equation (3), under DT and ST, and, for a_2 and b_2, under the two alternative assumptions that taxes come first, or that spending comes first.[18] For convenience of interpretation, while the original estimated coefficients have the dimension of elasticities, the table reports derivatives, evaluated at the point of means (dollar change in one variable per dollar change in another). Table 2 yields two main conclusions.

---

[17]A popular approach to dealing with changing trends is to apply a two-sided filter such as the Hodrick-Prescott filter, or the filter developed by Baxter and King [1996] and Stock and Watson [1997]. When the filtered data are used for econometric purposes, these two-sided filters have the disadvantage that future events can heavily influence today's trend: for example, a large temporary increase in defense spending in a given quarter will cause the estimated trend to increase long before that quarter, causing a large fall in the detrended series before and a large increase after the event (the graph of filtered defense spending in Stock and Watson [1997] illustrates this point vividly). Our filter is one-sided, and thus not subject to this problem.

[18]Note that, in constructing the cyclically adjusted tax shock t'_t, we use the time-varying elasticity a_1, not its mean.

The first is that the signs of the contemporaneous effects of taxes and of spending on GDP — c_1 and c_2 — are those one would expect — the former negative and the latter positive, and are rather precisely estimated. The two coefficients have very similar absolute values, and are also very similar across the two specifications, DT and ST. Under DT, a unit shock to spending increases GDP within the quarter by 0.96 dollars, while a unit shock to taxes decreases GDP by 0.87 dollars. The estimated negative effect of taxes on output depends very much on the use of instruments: the simple correlation between unexpected movements in cyclically unadjusted taxes, t_t, and unexpected movements in output, x_t, is *positive* and equal to 0.38. This raises the issue of the robustness of the construction of cyclically adjusted taxes to the specific value of a_1; we return to the issue in Section 7, and show that the basic conclusions hold over the plausible range of values for a_1.

The second conclusion is that the correlation between cyclically adjusted tax and spending innovations is low (–0.09 in our sample) yielding relatively low estimated values of a_2 and b_2 under either of the two alternative identification assumptions. These small values imply that the choice between the two ordering of taxes and spending makes little difference to the impulse responses.

5   Dynamic effects of taxes

5.1   Effects of estimated tax shocks

The top panel of Figure 3 shows the effects of a unit tax shock assuming that taxes are ordered first ($a_2 = 0$), under DT; the bottom panel does the same under ST.[19] For visual convenience, the impulse responses in these and the following figures and tables are transformations of the original impulse responses, and give the dollar response of each variable to a dollar shock in one of the fiscal variables. The solid line gives the point estimates. The broken lines give one-standard deviation bands, computed by Monte Carlo simulations (assuming normality), based on 500 replications. Table 3 summarizes the main features of the responses of the three variables. This table is useful in comparing the results from the benchmark specifications to alternative

---

[19]Note that the initial value of the change in taxes is not exactly 1. A unit tax shock ϵ_t^t translates into a less than unit change in taxes t_t, since GDP falls in response to the shock, decreasing tax revenues.

specifications.

Under DT (top panel of Figure 3 and of Table 3), output falls on impact by about 70 cents and reaches a trough 5 quarters out, with a multiplier (defined here and below as the ratio of the trough response of GDP to the initial tax shock) of 0.78. From then on, output increases steadily back to trend. The effect of tax shocks on government spending is small at all horizons, with the largest effect being -0.26 after 12 quarters, but it is precisely estimated.

Under ST (bottom panel of Figure 3 and of Table 3), the response of output is stronger and more persistent. Tax shocks have a very similar effect on output on impact, but the output trough is larger (-1.33 against -0.78 under DT), taking place after 7 quarters instead of 5; after this, the response of output stabilizes at around the peak response. The effect on taxes is slightly more persistent than in the DT case, while the effect on spending is similar.

Thus, under both specifications, tax increases have a negative effect on output. In both cases, the effect on output takes time to build up, with the largest response occurring after 5 or 7 quarters depending on the specification. The negative response of output is more pronounced under the assumption of a stochastic trend.

When taxes are ordered second, the results (not shown) are virtually identical to those in Figure 3. This comes from the low correlation between cyclically adjusted tax and spending innovations, which in turn leads to a small value of a_2.

5.2 Dynamic effects of the 1975:2 net tax cut

We now study the dynamic effects of a unit shock to the 1975:2 dummy variable, under DT and ST respectively.[20] The two panels of Figure 4 display the results: as we are looking at a tax *decrease*, the signs are reversed relative to Figure 3. But the effects are very similar. The impulse response captures well the fact that this tax decrease was temporary: taxes are back to normal after a quarter. The effects on output take some time to build up, reaching a peak after 4 quarters with a multiplier of 0.75 under DT and 1.02 under ST,

---

[20]Note that this exercise can be given a structural interpretation only if the tax cut captured by the dummy variable was unanticipated and was not a response to a contemporaneous output shock. In other words, identification in the "event study" approach requires the same two assumptions that allowed identification in the "structural VAR" approach.

thus close to the multipliers in Table 3. As usual, the response of output is more persistent under ST. The effects on spending are again small.

We find this similarity of results with the impulse responses from identified shocks comforting. The main difference is that the largest output response is reached sooner than in the impulse response to the typical tax shock; this is plausibly explained by the smaller persistence of the tax shock in the 1975:2 episode than of the typical tax shock in Figure 3. (One would also have expected the more temporary tax cut to lead to a smaller change in the present value of taxes, thus a smaller initial effect on consumption, and a smaller multiplier. There is no evidence that this is the case.)

6   Dynamic effects of spending

The two panels of Figure 5 show the effects of a unit spending shock on GDP when spending is ordered first ($b_2 = 0$) under DT and, then, under ST. As in the case of taxes, Table 4 summarizes the main features of the responses to a spending shock under alternative specifications.

Under DT (top panel of Figure 5 and of Table 4) spending shocks are longer lasting than tax shocks: 95 percent of the shock is still there after 2 years. GDP increases on impact by 0.84 dollars, then declines and rises again, to reach a peak effect of 1.29 after almost 4 years. Net taxes also respond positively over the same horizon, probably mostly as a consequence of the response of GDP (notice that the shape of the tax response mimics the shape of the output response).

The peak output response is smaller under ST (bottom panel of Figure 5 and Table 4), 0.90 against 1.29. The peak effect is now reached on impact rather than after 4 years; notice also that the impact response is very similar under DT and ST. The standard error bands are also quite large, so that the response of output becomes insignificant after only 4 quarters. Note the strong response of spending, which stabilizes at about 1.6 after 2 years.

Thus, in all specifications output responds positively to a spending shock. Spending reacts strongly and persistently to its own shock. Depending on the specification, the spending multiplier is larger or smaller than the tax multiplier (Traditional Keynesian theory holds that the spending multiplier should be larger than the tax multiplier; there is no consistent evidence that this is the case.)

As in the case of taxes, the ordering of the two fiscal variables does not

matter for the response to spending shocks: when taxes are ordered first, a shock to spending yields virtually identical effects on output (not shown) to those in Figure 5.

7 Robustness

We take up four issues in this section, sub-sample stability, the importance of allowing for quarter dependence, the effect of imposing cointegration, and robustness to alternative assumptions about the tax elasticity (detailed results are given on the web site).

Subsample stability. Unlike monetary policy, in our sample fiscal policy in the United States does not lend itself easily to a periodization based on alternative policy regimes. Hence, we check for subsample stability by dropping one decade at a time.[21] Table 5 reports the results from this exercise. The left and right parts summarize the impulse responses of output to a net tax shock and a spending shock, respectively, obtained by dropping one decade at a time.

The exclusion of the eighties causes a substantial drop in the magnitude of the tax multiplier, under both ST and DT. The instability is more pronounced under DT: the tax multiplier ranges from -0.49 when the eighties are excluded to -1.45 when the nineties are excluded. There is also evidence of some instability in the spending multiplier, in particular under ST: the multiplier when the eighties are excluded is about three times that when the seventies are excluded — 1.80 against 0.62. We do not have a convicing explanation for the pattern of change of impulse responses over time.[22]

Quarter dependence. How important is quarter dependence for our results? [23] Under DT, an F-test of the hypothesis of no-quarter dependence in the VAR regression gives p-values of 0.24 in the tax equation, 0.67 in the spending equation, and 0.46 in the output equation. Under ST, the p-

---

[21]In performing this exercise, we use the average value of a_1 over the relevant sample.

[22]We cannot conduct formal Chow tests on the model with quarter dependence, because we cannot estimate all the coefficients of the quarter dependent model with just 10 years of data. When we estimate the model without quarter dependence, the Chow test for subsample instability is sometimes significant in the DT specification: specifically, there is some evidence of a structural shift in the tax equation in the eighties, in the government spending equation in the seventies and eighties, and in the output equation in the seventies.

[23]Recall that we use quarter dependence only to identify the shocks, but we compute the impulse responses from a system estimated without quarter dependence.

15

values are 0.12, 0.32, and 0.15. The responses to tax shocks obtained from non-quarter dependent estimation are very similar to those in Table 3. The difference is larger for spending shocks: under DT, the response of output on impact increases by about 40 percent, so that now the peak response occurs immediately rather than after 15 quarters; under ST, the impact response and the spending multipliers increase by about 30 percent. In short, for the United States for the period at hand, we could have ignored quarter dependence and obtained largely similar results. Our work on other countries suggests however that quarter dependence is often important. This is why we maintained it in our benchmark specification.

Cointegration. The imposition of a cointegrating relation between G and T yields very similar results to our benchmark case.[24]

Alternative net tax elasticities. Our identification procedure depends crucially on using ϵ_t^t as an instrument. The construction of this series in turn depend very much on the elasticity of net taxes a_1. One may ask how sensitive the results are to alternative values of a_1. Figure 6 gives the effects of a tax shock on GDP, for DT and ST, when taxes are ordered first ($a_2 = 0$), and for three different values of a_1: the baseline value minus 0.5, the baseline value, and the baseline value plus 0.5. We believe that the range implied by these three values more than covers the relevant range for a_1. A simple way of describing our results is that a change of a_1 from the bottom to the top of the range increases the impact effect on GDP (in absolute value) by 0.5 under both specifications, and increases the maximum multiplier (in absolute value) by 0.5 under DT and by 0.8 under ST. We read this figure as saying that the broad response of output we characterized earlier is robust to the details of construction of a_1.

8   Anticipated fiscal policy

Fiscal policy is subject to two types of lags: *Decision lags*, which imply that it takes some time for policy to be changed in response to shocks; *implementation lags*, which imply that it takes some time for policy changes to be implemented. The first source of lags is what helped us achieve identification. The second source raises however a problem we have ignored until now, namely that what we (the econometrician) identify as fiscal shocks may

---

[24]In keeping with quarter dependence, we allow the coefficient of the cointegrating term $G_{t-1} - T_{t-1}$ to differ depending on the quarter.

be the result of earlier policy changes, and thus be in fact anticipated by the private sector.

To see the problems this raises and how they may be solved, let us go back to our original specification of the joint movement of taxes, spending, and output in equations (1), but now allow the private sector to know fiscal shocks one quarter in advance. Also, for notational simplicity, let us ignore quarter dependence and government spending, so we are looking at a two-variable VAR in T and X:

$$(6) \qquad T_t = a_1 X_t \qquad\qquad + A_{11}(L)T_{t-1} + A_{12}(L)X_{t-1} + e_t^t$$

$$(7) \qquad X_t = c_0 E_t T_{t+1} + c_1 T_t + A_{21}(L)T_{t-1} + A_{22}(L)X_{t-1} + e_t^x.$$

The first equation is as before: Taxes depend on output, current and lagged, and lagged taxes, together with fiscal shocks. The second equation now allows output in quarter t to depend not only on current and lagged taxes, but also on the expectation of taxes in quarter $t + 1$, based on an information set which includes e_{t+1}^t. In other words, this specification assumes that, because of implementation lags, the private sector knows fiscal shocks one quarter ahead, and may therefore react to them one quarter before they are actually observable by the econometrician.

To see the econometric problems this specification raises, rewrite the equation for output as

$$(8) \qquad X_t = c_0 T_{t+1} + c_1 T_t + A_{21}(L)T_{t-1} + A_{22}(L)X_{t-1} + \epsilon_t^{x'},$$

where the error term is defined as $\epsilon_t^{x'} \equiv [e_t^x - c_0(T_{t+1} - E_t T_{t+1})]$.

From the assumption that e_{t+1}^t is known at time t, the term in parentheses in the composite error term is uncorrelated with e_{t+1}^t but will typically be correlated with e_{t+1}^x. Hence, both T_{t+1} and T_t in (8) are correlated with the composite error term. As in our benchmark case, T_t is correlated with e_t^x, through the effect of e_t^x on output and in turn on tax revenues. Also, now T_{t+1} is correlated with both components of the error term: with e_{t+1}^x through the effects of the latter on taxes in quarter $t + 1$, and with e_t^x, through the effect of the latter on T_t and therefore on T_{t+1}.

Can equation (8) be estimated by instrumental variables? The answer is yes, if we are willing to make stronger identification assumptions. The reasoning follows the logic of our previous identification strategy. If we can construct series for e_t^t, then e_t^t and its value led once, e_{t+1}^t, can be used as instruments for T_t and T_{t+1}. Both are correlated with T_t or/and T_{t+1} and

uncorrelated with the two components of the composite error term. Put another way, if we can identify e_t^t in the tax equation, we can then use it and its value led once as instruments in the output equation.

With this in mind, let us turn to the tax equation, and for reasons which will be clear below, rewrite it by separating X_{t-1} from further lagged output terms:

$$(9) \qquad T_t = a_1 X_t + a_{121} X_{t-1} + A_{11}(L) T_{t-1} + \tilde{A}_{12}(L) X_{t-2} + e_t^t,$$

where $\tilde{A}_{12} \equiv L^{-1}(A_{12}(L) - A_{121})$. Note that both X_t and X_{t-1} are likely to be correlated with the error term e_t^t: X_t because of the effect of taxes on output, and X_{t-1} because of the assumption that fiscal shocks in quarter t are known to the private sector in quarter $t-1$.

One way then to achieve identification is to assume that there is no discretionary response of fiscal policy to output shocks this quarter (the assumption we made until now) nor to output shocks *last quarter* (a stronger assumption than before). In other words, decision lags prevent a discretionary response of fiscal policy to output for two quarters.[25]

We proceed under this assumption, and following the same methodology as before, we construct a_1 and a_{121} from information about the tax structure and the response of tax bases to GDP, run the regression above imposing the values of a_1 and a_{121} so constructed, and obtain a time series for e_t^t as the residual.

The construction of a_{121} under the assumption that it only reflects the automatic response of taxes to output is presented in the appendix. It suggests an average value for a_{121} over the 1960:1-1997:4 period of 0.16, compared to 2.17 for a_1. We estimate the system with anticipated effects under three alternative assumptions to construct e_t^t, $a_{121} = 0.0$, 0.16, and 0.5, which can roughly be thought as the lower bound, mean, and upper bound estimates for a_{121}. For simplicity of computation, we ignore quarter dependence in estimation, and for simplicity of presentation, we present the results only under stochastic detrending (ST). Also, while we excluded spending for notational simplicity in the argument above, our VAR specification obviously includes both spending and taxes, in addition to output. We treat spending symmetrically with taxes, allowing for an effect of anticipated spending next quarter

---

[25]Clearly if fiscal policy were anticipated, say, *two* quarters ahead, identification would require that policy cannot respond to output shocks this quarter and in the last *two* quarters.

on output this quarter. Extending our earlier assumption about spending, we assume that there is no automatic response of output on spending either contemporaneously (our earlier assumption) or at a one-quarter lag (an extension of our earlier assumption).

For either of the three values of a_{121}, there is not much evidence of an effect of anticipated tax changes on output. The estimated parameter c_0 varies from 0.07 for $a_{121} = 0$ to -0.03 for $a_{121} = 0.5$, and is never significant. There is a bit more evidence of a positive effect of anticipated government spending on output. The coefficient on anticipated spending in the output equation varies from 0.12 when $a_{121} = 0$ to 0.19 when $a_{121} = 0.5$, and, in this last case, is significant at the 10% level.

Figure 7 shows the effects, from quarter 0 on, of a tax shock of 1 in quarter 1 under the assumption that $a_{121} = 0.16$, the estimated sample average. It does so by simulating the system of equations above (but extended to allow for spending), in response to a shock of e_t^t of 1 at $t = 1$. These simulations are done under rational expectations: Although the shock occurs at $t = 1$, output and taxes start moving at $t = 0$ via the effects of anticipated taxes on output and therefore on taxes. For comparison, the figure also reports the response in the benchmark case when we did not allow for an effect from anticipated fiscal policy. Obviously, this last response starts at $t = 1$.

The response of output at $t = 0$ is small. After that, the response of output is qualitatively similar to that without anticipated effects; the largest difference occurs when we assume $a_{121} = 0$, in which case the long run effect with anticipated effects is -0.95 against about -1.30 in the case of no effects from anticipated fiscal policy.

The difference with the case of no anticipated fiscal policy effects is larger in the case of government spending shocks, displayed in Figure 8. Because the coefficient of anticipated spending in the output equation is large, the effect on output at $t = 0$ is substantial, around 0.5 percentage points. And because both the coefficients of lagged output and of contemporaneous spending in the output equation are also quite large, the response of output rises even more at $t = 1$, to about 2.0, and stays there thereafter. This compares with a response of output in the case of no effects from anticipated fiscal policy, of about 1.0 in the long run.

Another approach to the problem of anticipated fiscal policy is again to make use of the tax event of 1975:2. Suppose the large tax cut had been anticipated by, say, one quarter; then if we add one lead value of the 1975:2 dummy in our specification, it will pick up the effect, if any, of the anticipated

tax cut on output. (Note that by including the first lead of the tax dummy variable, we implicitly have to make the same assumption that we make explicitly in our first approach to anticipated fiscal policy: namely, that the tax cut was not implemented in response to either output shocks in either 1975:2 or 1975:1; otherwise, our coefficients on the dummy variable would be biased.) The coefficients on lead values of the tax dummy simply pick up the residuals for GDP (in the benchmark specification) in the quarters before the change. For both 1974:4 and 1975:1, the two residuals are small and negative, leading to estimated negative effects, at one and two leads, of the 1975:2 tax cut on GDP in 1974:4 and 1975:1. Thus, at least under the prior that anticipated tax cuts should lead to either zero or positive effects on output, there is no evidence of anticipation effects of fiscal policy in that particular episode.

We draw two conclusions from this section. First, identifying and tracing the effects of anticipated fiscal shocks can be done within a VAR framework but requires stronger identification restrictions. These restrictions may be too strong. Second, under those identification restrictions, we find that allowing for anticipations of fiscal policy does not alter substantially the results we have obtained so far, more so in the case of the effects of taxes. Although this was not our focus, and we do not see the evidence as very strong one way or the other, we find only weak effects of anticipated tax changes on output, a finding consistent for example with the findings by Poterba [1988] for the Reagan tax cuts.

9   Adding the fifties

So far, we have used only the post-1960 sample. The reason, we argued, was that the 1950s, with their large spending and tax shocks, do not appear to be generated by the same stochastic process as the post-1960 data. Still, there is a number of ways in which the 1950s can be used to provide information on the effects of fiscal policy. We have explored two of them.

9.1   Using dummies for the major tax and spending shocks

The first approach parallels our treatment of the 1975 tax cut. We extend the sample to start in 1949:1, use dummies for the major tax and spending

episodes, then trace both the effects of normal fiscal shocks in the VAR estimated using the extended sample, as well as the effects of the dummies.

The last approach runs however into an identification problem which we did not face in the post 1960 sample. We can still trace the dynamic effects of the identified "normal" fiscal shocks — e_t^t and e_t^g — on output. But if we allow the dummy associated with each major tax and spending shock to have its own distributed lag effect, we cannot identify the effects of each major spending or tax shock: They are too close in time to each other. When looking at output for example in 1950:3, we cannot distinguish between the effect of the 1950:1 tax cut after two quarters, the effect of the 1950:2 tax increase after one quarter, or the contemporaneous effect of high spending in 1950:3. Nevertheless, something — admittedly more informal — can be learned by estimating and examining the deviations of taxes, spending and output from normal in the early 1950s. Let us turn to the results.

We estimate impulse responses using the sample starting in 1949:1, but allowing for 13 dummies from 1950:1 to 1953:1 (equivalently 9 dummies, each with four lags; this is another way of stating the identification problem we just discussed). Under both DT and ST, the tax and spending output multipliers associated with e_t^t and e_t^x are similar to those of Tables 3 and Table 4, respectively. The shapes of the impulse responses (available on the journal's web site) are also very similar to those from the post-1960 sample. The main difference is that the output response to a spending shock under DT is much less persistent, stabilizing after about 8 quarters at around 0.20, instead of 0.65 in the post-1960 sample.

Turn now to the effects of the large fiscal shocks of the early 1950s. Figure 9 gives the deviations of taxes and spending from normal, starting in 1950:1, and the associated deviation of output, as captured by the coefficients on the 13 dummies and their dynamic effects through the VAR dynamics.

The figure shows an initial fall in taxes, followed by a strong increase, which after the first 2 quarters, largely mimics the path of GDP. This captures the large tax cut of 1950:1, and the larger tax increases of 1950:2 and 1950:3.

Spending initially deviates little from normal, then it increases over time. This captures the large buildup in defense spending associated with the Korean war, starting in 1951:1, and continuing until 1952:2.

We can then think of the path of GDP as the response of GDP to this complex path of taxes and spending. The response of GDP shows an initial increase—presumably to the initial tax cut—followed by further increases—presumably in response to the Korean war buildup. The effects of spending

and taxes appear larger than in the post-1960 sample. In the DT case, the maximum increase in output is more than 2.5 times the maximum increase in spending (and this despite the fact that the increase in spending comes with a nearly equal increase in taxes). In the ST case, this ratio is smaller, but still equal to 1.5.

9.2   Defense versus non-defense spending

Another way of approaching the difference between the 1950s and the rest of this sample is to note that the 1950s were dominated by shocks to defense spending, while the rest of the sample is dominated by shocks to non-defense spending. Hence, by dividing government spending between its defense and non-defense components, one can hope to capture explicitly the different persistence of the spending shocks in the early 1950s and in the rest of the sample.

We thus estimate impulse responses from a four-variable VAR, with defense and non-defense government spending on goods and services replacing aggregate government spending on good and services. The system also includes the 1975:2 dummy variable and its first four lags.

The contemporaneous correlation between the two shocks to defense and non-defense spending is small — 0.18 — so the order of the first two variables makes virtually no difference to the results. We discuss the results when the variable whose effect is studied is ordered first. A table with a more complete description of the impulse responses can be found on the journal's web site.

Under DT, we obtain large and similar multipliers for defense and non-defense spending — 2.50 and 2.67, respectively — with a bell-shaped output response in both cases. This is despite the fact that the persistence of the two spending shocks is indeed very different: defense spending keeps increasing considerably for one year after the initial shock, while non-defense spending decreases after the shock.

Under ST, the responses to a defense spending shock display a similar picture—the output response is smaller and the spending response stronger and more persistent. But the output response to a non-defense shock is now small and insignificant, despite the fact that non-defense own response is stronger and more persistent. A potential explanation is in the estimated defense response, which becomes negative very soon and exceeds, in absolute value, the response of non-defense spending. We do not have an explanation for this lack of robustness of the impulse responses to non-defense shocks

across the DT and ST specifications.

10 Relation to other studies

We can now compare our methodology (the combination of a structural VAR and an event study approach) and results to other recent empirical papers on fiscal policy.

The event-study approach has been used by Ramey and Shapiro [1997], Edelberg et al. [1998], and Burnside et al. [2000]. All three studies adopt a 'narrative approach' exploiting the exogeneity of military buildups. They define one dummy variable taking the value of 1 in 1950:3, 1965:1, 1980:1, and trace out its effects on a number of macroeconomic variables, including GDP (hours in the case of Burnside et al. [2000].) Ramey and Shapiro do so in the context of univariate regressions. The other two papers do so in the context of a VAR.

All three papers find a roughly coincidental increase of defense spending and of GDP (business hours in teh case of Edelberg et al.). Based on the graphs in Edelberg, Eichenbaum and Fisher, defense spending peaks after two years at about 2.5 percentage points of GDP, and slowly declines thereafter. GDP peaks after 4 quarters at about 3.5 percent, and then goes back to trend within 2 or 3 years. Because their estimated initial increase in spending is small, the defense spending multiplier, defined as we do as the ratio of the peak response of GDP to the initial spending shock, appears to be about 28; if the multiplier is defined as the maximum increase in GDP to the maximum increase in spending, it is about 1.4.

Assuming the episodes that are captured by the dummy variable approach are truly exogenous, the advantage of this approach is that one does not need to identify the VAR in order to trace out the effects of the dummy variable. But when the dummy variable incorporates episodes with different characteristics, it is not clear exactly what response one is estimating when tracing the effects of a shock to the dummy. And the three episodes captured by the Ramey and Shapiro dummy do differ greatly, in terms both of size and of the accompanying tax policy: the Korean war buildup was accompanied by large tax increases, the (much smaller) Reagan buildup by tax cuts.

A non-exhaustive list of studies following a VAR or a structural VAR approach includes Barro [1981] and Ahmed and Rogers [1995] (who both look at the effect of expenditures, ignoring taxes), Blanchard and Watson

[1986] (who look at the effect of an index of fiscal policy based on a model which does not satisfy Ricardian equivalence), and more recently Rotemberg and Woodford [1992] and Fatas and Mihov [1998]. The last two are closest to ours.[26]

Fatas and Mihov's specification is closest to our benchmark specification, but includes only one fiscal variable in the VAR, the ratio of the primary deficit to GDP. The other variables are output, the GDP deflator, and the interest rate. Their sample starts in 1960:1. Identification is achieved by assuming that the "sluggish" private sector variables, output and prices, do not respond to changes in policy within a quarter; this, in a sense, is the exact opposite of our assumption. An increase in the primary deficit by 1 percentage point of GDP leads to an increase in GDP by about 1 percentage point after about two years, while the primary deficit goes back to trend very quickly. If we interpret the increase in the primary deficit as coming from an increase in spending, then their estimated spending multiplier is similar to what we find over the same sample (see Table 4, DT), although our spending shocks are much more persistent.

Rotemberg and Woodford's specification is closest to our defense/non-defense decomposition. They trace the effects of military spending and military employment shocks on output by Choleski-decomposing a four-variable VAR in personnel military expenditure, military purchases, output, and the real wage. They do not control for other spending or for taxes. Their sample covers 1947 to 1989. The estimated impact elasticity of private GDP to military purchases (when ordered first) is about 0.1, which implies an impact multiplier above 1.0 (the average ratio of military purchases to GDP has been below 10 percent after WWII), and a bit larger if one considers total GDP, as we do. This effect persists for about 4 quarters, and dies out completely after 8 quarters. Thus, they find a smaller defense multiplier than we do.

In short, our basic results that fiscal expansion, be it a decrease in taxes,

---

[26]There is yet a third approach, the study of the effects of fiscal policy in large macroeconometric models. Because of their very structure, (namely the twin assumptions that fiscal policy affects primarily aggregate demand, and that aggregate demand affects output for some time), these models largely postulate rather than document an effect of fiscal policy on activity. But, conditional on these assumptions, they give a sense of magnitudes. Results from 12 models are reported in Bryant [1988]; under the assumption of an unchanged path for money, the average effect of a permanent government spending cut of one percentage point of GDP is a decrease of -1.27% of GDP in first year, declining slowly to -0.65% after five years.

or an increase in spending, increase output for some time, are consistent both in sign and in rough magnitude with other recent findings. We believe that our distinction between the effects of taxes and the effects of spending, and our focus on the effects of both small and large shocks, gives a better sense of the nature and robustness of the findings.

There is however a dimension in which some recent papers have gone further than we have so far, namely in looking at the effects of fiscal policy not only on output but on a large number of macroeconomic variables, from the components of GDP, to real wages, to interest rates. Our methodology makes it difficult to go very far without running out of degrees of freedom. But, in the next section, we take a step in that direction.

11   Effects on output components

We now go back to the post-1960 sample, and decompose the output effects of tax and spending shocks into the effects on each component of GDP. This exercise is interesting in itself, and also because it is a first step in sorting out the relative merits of alternative theories. For instance, both standard neoclassical and Keynesian models imply a positive effect of government spending on output. However, neoclassical models predict a negative effect on private consumption (see e.g. Baxter and King [1993]), while Keynesian models predict the opposite sign. The implications of these models for private investment is more ambiguous. In the neoclassical model, a shock to governmemt spending can raise private investment if the shock is sufficiently persistent and taxes are sufficiently non-distortionary; investment can fall in the opposite case. In a Keynesian model, investment increases if the accelerator effect prevails, it falls if the effect of a higher interest rate prevails.

We estimate a four-variable VAR, with the component of GDP whose response we are studying ordered last. The relation between residuals now becomes

$$
\begin{aligned}
t_t &= a_1\, x_t + a_2\, e_t^g + e_t^t \\
g_t &= b_1\, x_t + b_2\, e_t^t + e_t^g \\
x_t &= c_1\, t_t + c_2\, g_t + e_t^x \\
x_{i,t} &= d_1\, t_t + d_2\, g_t + e_t^{x_i},
\end{aligned}
$$

(10)

where $x_{i,t}$ indicates a component of GDP, so e_t^x and $e_t^{x_i}$ are in general be correlated.

25

11.1   Response to a tax shock

Table 6 displays a summary of the impulse response of the various GDP components to a net tax shock, under DT and ST. The impulse responses of aggregate GDP and government spending change (slightly) with each component added to the VAR, raising the issue of which one to report; in the first two lines in each case, we give the response of GDP and government spending from the three-variable model (see Table 4). The last line in each case displays the unconstrained sum of the responses of the individual components of GDP, which is denoted by 'SUM'.

Under DT, an increase in taxes reduces all the private components of GDP. After a small decline on impact, private consumption falls by a maximum of 0.35 (35 cents) after 5 quarters; investment falls by 0.36 on impact and then increases to become marginally positive after 2 years. The negative effect on imports and exports is very small. The sum of the responses of the components of GDP is close to the response of GDP in the 3-variable system.

A similar picture emerges under ST, with very similar peak negative responses of private consumption and investment, and again small responses of imports and exports. However, the difference between the GDP response in the 3-variable model and the sum of the responses of its components is more pronounced.

11.2   Response to a spending shock

The next two panels in Table 6 display the responses to an increase in government spending. The peak responses of each component are considerably larger than in the case of tax shocks.

Under DT, private consumption increases by 0.50 (50 cents) on impact, while investment does not move; imports and exports now react strongly, increasing by 0.64 and 0.20, respectively. The positive effect on consumption builds up for 14 quarters, reaching a peak of 1.26; investment declines for the first 5 quarters, with a peak crowding out effect of -1.0. After the initial increase, exports start declining with a maximum negative effect of 0.80 after 9 quarters; similarly, after an initial increase, imports start declining and reach a negative peak of 0.49 after 9 quarters. All these responses are precisely estimated, and the unconstrained sum of the responses is close to the impulse response of GDP in the 3-variable system.

Under ST, the basic picture is qualitatively similar, but the peak re-

sponses of consumption, imports, and exports are smaller; the standard error bands for these components also become wider. Also, note that in this case also the output response is very close to the sum of the responses of the individual components.

Thus, we find that private consumption is consistently crowded out by taxation, and crowded in by government spending. The latter result is difficult to reconcile with a neoclassical model, regardless of the persistence of the spending shock, but it is consistent with a Keynesian model. More surprisingly, we consistently find a considerable crowding out of investment both by government spending and to a lesser degree by taxation; this implies a strong negative effect on investment of a balanced-budget fiscal expansion. This time, these effects are consistent with a neoclassical model, but inconsistent with a standard Keynesian approach. In this approach, an increase in spending may increase or decrease investment depending on the relative strength of the effects of the increase in output and the increase in the interest rate; but, in either case, increases in spending and taxes have opposite effects on investment. This is not the case empirically. Interestingly, using a yearly panel VAR on 18 OECD countries over the same period, Alesina et al. [1999] reach the same conclusion on the effects of fiscal policy on investment.

Note also the decline in imports (after a brief initial surge), which is surprising in light of the considerable increase in GDP. This decline may consist mostly of a decline in imports of capital goods, and therefore might be a reflection of the decline in investment. We have not explored this explanation here, but it clearly deserves further investigation.

12   Conclusions

Our main goal in this paper was to characterize as carefully as possible the response of output to the tax and spending shocks in the post-war period in the United States. From the several specifications we have estimated and the different exercises we have performed, we reach the following conclusions:

- The first is consistent with standard wisdom: When government spending increases, output increases; when taxes increase, output falls.

- In most cases the multipliers are small, often close to one. In the case of spending shocks, the proximate explanation is in the opposite effects they have on the different components of output: While private

consumption increases following spending shocks, private investment is crowded out to a considerable extent. Both exports and imports also fall.

- The responses of investment and imports are difficult to reconcile with most macroeconomic theories; thus, while we do not attempt an explanation here, we believe this certainly deserves further investigation.

Appendix

A.1   The data

All the data, unless otherwise noted, are from the *National Income and Product Accounts*. The Citibase mnemonics are given in parentheses. FG stands for Federal Government; SLG stands for State and Local Governments.

Government spending:
Purchases of goods and services, FG (GGFE) + Purchases of goods and services, SLG (GGSE).

Net taxes:
Receipts, FG (GGFR)[27] + Receipts, SLG (GGSR) [28] – Federal grants-in-aid (GGAID) – Net transfer payments to persons, FG (GGFTP) – Net interest paid, FG (GGFINT) – Transfer payments to persons, SLG (GGST) – Net interest paid, SLG (GGSINT) + Dividend received by government, SLG (GGSDIV).

Defense spending:
Purchases of goods and services, national defense, FG (GGFEN).

A.2   Elasticities

To construct the aggregate net tax elasticity to GDP, a_1, we consider four categories of taxes: indirect taxes (IND), personal income taxes (INC), corporate income taxes (BUS), and social security taxes (SS). For each

---

[27]The Quarterly National Income and Product Accounts report taxes on a cash basis, except for the Corporate Profits Tax and the Indirect Business Tax, which are reported on an accrual basis. We use the Quarterly Treasury Bulletin to obtain data on federal corporate profit taxes on a cash basis. We do not have however this series for the state and local governments, so we use accruals (the proportion of corporate profit tax receipts collected by state and local governments is about 5 percent at the beginning of the sample, and about 20 percent at the end of the sample.) We also could not find data on indirect taxes on a a cash basis; however, the difference between receipts and accruals appears to be very small.

[28]Corporate income taxes collected by state and local governments are not included.

category, we construct the elasticity to GDP (X) as the product of the tax elasticity to its own base, η_{T_i,B_i}, and the elasticity of the tax base to GDP, $\eta_{B_i,X}$ (see expression (5) in the text). For each tax, we must also take into account the possible presence of collection lags and of quarter dependence. On the expenditure side, we also construct an approximate output elasticity of total transfers. The rest of this section describes our assumptions.

Indirect taxes.
We take the tax base to be GDP. (This is an approximation. In many states, food consumption is excluded. In most states, the sale of materials to manufacturers, producers and processors is also excluded (see Advisory Committee of Intergovernmental Relations [1995]). Hence:

$\eta_{B_{IND},X} = 1.0$
$\eta_{IND,B_{IND}} = 1.0$ (from Giorno et al. [1995])

Collection lags: 0.
Quarter dependence: none.

Personal income taxes.
We start from the formula for the elasticity to GDP from Giorno et al. [1995]. Let $T = t(W)W(E)E(X)$, where T is total revenues from the personal income tax, t is the tax rate, W is the wage (or earnings, in the OECD terminology), E is employment and X is GDP. Define the tax base as $B_{INC} = W * E$. Assuming constant elasticities everywhere, define:

elasticity of taxes to earnings: $D = d\log(tW)/d\log W$
elasticity of earnings to employment: $F = d\log W/d\log E$
elasticity of employment to output: $H = d\log E/d\log Y$

Totally differentiating the expression for total tax revenues, after some steps one obtains:

$\eta_{B_{INC},X} = $ H/(F+1)
$\eta_{INC,B_{INC}} = $ (FD+1)/(F+1)

We obtain values of D from Giorno et al. [1995]. We estimate F from a regression of the log change of the wage of production workers on the first lead

and lags 0 to 4 of the log change in employment of production workers. We estimate H from a regression of the log change of employment of production workers on the first lead and lags 0 to 4 of the log change in output. The values of F and H are the estimated coefficients of lag zero of the dependent variable. We find $F = .62$ and $H = .42$.

Collection lags: 0.
Quarter dependence: none.

Note that for personal income taxes we assume the same elasticity for employees and the self-employed: the former have their taxes withheld at the source, while the latter make quarterly payments based on the estimated income of that quarter. As long as there is no systematic pattern in the end-year adjustments (as it should be if the tax system is well designed), our assumption does not introduce any substantial bias in our aggregate elasticity.

Social security taxes.
We follow exactly the same procedure as for personal income taxes. The only difference is in the value of the elasticity of taxes to earnings, D, which we also obtain from Giorno et al. [1995].

Corporate income taxes.
Each corporation can have its own fiscal year different from the tax year. Large corporations are required to make quarterly installment payments, of at least 0.8 of the final tax liability. Until 1980, no penalty was applied if the estimated tax liability was based on the previous year's tax liability; this exception has been gradually phased out since 1980. Hence:

$\eta_{B_{BUS},X}$: we estimate it as the lag 0 estimated coefficient from a regression of quarterly changes of corporate profits on the first lead and lags 0 to 4 of changes in output.
$\eta_{BUS,B_{BUS}}$: Giorno et al. [1995] gives a value of 1 for the annual elasticity. This is the right value of the quarterly elasticity for tax accruals as well. For tax receipts however, things are more complicated. Corporations make partial payments throughout the year, and make up for the difference in quarter 2. Hence, we estimate the elasticity of receipts to accruals from a regression of tax receipts on tax accruals, using quarters 1, 3 and 4, but not quarter 2.

This gives a coefficient of 0.85 for $\eta_{BUS,B_{BUS}}$.

Collection lags: yes. This follows from the fact that we use a value of .85 for $\eta_{BUS,B_{BUS}}$, as explained above. This value appears stable over time.

Quarter-dependence: yes. (for the same reason as above)

Transfers.

Among transfers, unemployment benefits certainly have a large within-quarter elasticity to the cycle. However, unemployment benefits are a small component of transfers: on average, unemployment expenditures (defined as the sum of 'active' and 'passive' measures) represent less than 2 percent of total government expenditures in the United States. Other categories of transfers might be sensitive to the cycle. We do not have reliable quarterly data for our sample on the components of transfers that might be sensitive to the cycle. Based on OECD estimates, we use a value of -0.2 for the elasticities of total tranfers to GDP.

In Section 8, our identification strategy requires us to consider the elasticity of net taxes to lagged output, the term a_{121}. To do so, we assume that there are no non-quarter dependent collection lags, and take into account the lags in the response of the tax bases (the wage, employment, profit) to GDP, using the same regressions as before.

References

Advisory Committee of Intergovernmental Relations, 1995, Significant features of fiscal federalism, 1.

Ahmed, S. and Rogers, J., 1995, Government budget deficits and trade deficits: are present value constraints satisfied in long-term data?, *Journal of Monetary Economics* **36**, 351–374.

Alesina, A., Ardagna, S., Perotti, R., and Schiantarelli, F., 1999, Fiscal policy, profits, and investment, *mimeo*, Boston College.

Barro, R., 1981, Output effects of government purchases, *Journal of Political Economy* **89(6)**, 1086–1121.

Baxter, M. and King, R., 1993, Fiscal policy in general equilibrium, *American Economic Review*.

Baxter, M. and King, R., 1996, Measuring business cycles: approximate band-pass filters for economic time series, *mimeo*, University of Virginia.

Bernanke, B. and Mihov, I., 1998, Measuring monetary policy, *Quarterly Journal of Economics* **113(3)**, 869–902.

Blanchard, O. and Watson, M., 1986, Are all business cycles alike?, pp. 123–156, *The American Business Cycle: Continuity and Change*, Robert Gordon, ed, NBER and University of Chicago Press.

Blinder, A., 1981, Temporary income taxes and consumer spending, *Journal of Political Economy* **89(1)**, 26–53.

Bohn, H., 1991, Budget balance through revenue or spending adjustments? Some historical evidence for the United States, *Journal of Monetary Economics* **27**, 333–359.

Bohn, H., 1998, The behavior of U.S. public debt and deficits, *Quarterly Journal of Economics* **113(3)**, 949–964.

Bryant, R., 1988, *Empirical Macroeconomics for Interdependent Economies*, Brookings Institution, Washington.

33

Burnside, C., Eichenbaum, M., and Fisher, J., 2000, Assessing the effects of fiscal shocks, pp. *mimeo*, Northwestern University.

Cohen, D. and Folette, G., 2000, The automatic fiscal stabilizers: quietly doing their thing, in: Fiscal policy in an era of surpluses: Economic and financial implications, *Federal Reserve Bank of New York Economic Policy Review* **6(1)**, 35–68.

Edelberg, W., Eichenbaum, M., and Fisher, J., 1998, Understanding the effects of a shock to government purchases, *mimeo*, Northwestern University.

Fatas, A. and Mihov, I., 1998, Measuring the effects of fiscal policy, *mimeo*, INSEAD.

Giorno, C., Richardson, P., Roseveare, D., and van der Noord, P., 1995, Estimating potential output, output gaps, and structiral budget deficits, *Economics Department Working Paper* 152, *OECD*, Paris.

Gordon, D. and Leeper, E., 1994, The dynamic impact of monetary policy: an exercise in tentative identification, *Journal of Political Economy* **102(6)**, 1128–1147.

Hamilton, J. and Flavin, M., 1986, On the limitations of government borrowing: a framework for empirical testing, *American Economic Review* **76**, 808–819.

Poterba, J., 1988, Are consumers forward looking? evidence from fiscal experiments, *American Economic Review, Papers and Proceedings* **78**, 413–418.

Ramey, V. and Shapiro, M., 1997, Costly capital reallocation and the effects of government spending, *Carnegie-Rochester Conference Series on Public Policy*.

Rotemberg, J. and Woodford, M., 1992, Oligopolistic pricing and the effects of aggregate demand on economic activity, *Journal of Political Economy* **110(6)**, 1153–1207.

Sims, C. and Zha, T., 1996, Does monetary policy generate recessions?,

mimeo, Yale University.

Stock, J. and Watson, M., 1997, Business cycle fluctuations in u.s. economic time series, *mimeo*, Harvard University.

Table 1: Large changes in net taxes and spending.

sd($\Delta \log G$)		sd($\Delta \log T$)	
0.019		0.049	
$\Delta \log G > 3$ sd		$\Delta \log T > 3$ sd	
1951:1	0.103	1950:2	0.266
1951:2	0.112	1950:3	0.171
1951:3	0.108	1975:2	-0.335
		1975:3	0.240
$2 < \Delta \log G < 3$ sd		$2 < \Delta \log T < 3$ sd	
1948:2	0.039	1947:3	-0.117
1948:4	0.043	1947:4	0.107
1949:1	0.049	1951:1	0.097
1949:2	0.043		
1950:4	0.054		
1951:4	0.051		
1952:2	0.041		
1967:1	0.041		

Table 2: Estimated contemporaneous coefficients.

	c_1	c_2	a_2	b_2
DT				
coeff.	-0.868	0.956	-0.047	-0.187
t-stat.	-3.271	2.392	-1.142	-1.142
p-value	0.001	0.018	0.255	0.255
ST				
coeff.	-0.876	0.985	-0.057	-0.238
t-stat.	-3.255	2.378	-1.410	-1.410
p-value	0.001	0.019	0.161	0.161

DT: Deterministic Trend; ST: Stochastic Trend.
Sample: 1960:1 - 1997:4.
c_1: effect of t on x within quarter;
c_2: effect of g on x within quarter;
a_2: effect of g on t within quarter (assuming b_2 = 0, i.e. when spending is ordered first);
b_2: effect of t on g within quarter (assuming a_2 = 0, i.e. when net taxes are ordered first).
All effects are expressed as dollar for dollar.

Table 3: Responses to tax shocks.

	1qrt	4qrts	8qrts	12qrts	20qrts	peak
			DT			
GDP	-0.69*	-0.74*	-0.72*	-0.42*	-0.22	-0.78*(5)
TAX	0.74*	0.13	-0.21*	-0.20*	-0.11	
GCN	-0.05*	-0.12*	-0.24*	-0.26*	-0.16*	
			ST			
GDP	-0.70*	-1.07*	-1.32*	-1.30*	-1.29*	-1.33*(7)
TAX	0.74*	0.31*	0.17	0.16	0.16	
GCN	-0.06*	-0.10*	-0.17*	-0.20*	-0.20*	

DT: Deterministic Trend; ST: Stochastic Trend. A '*' indicates that 0 is outside the region between the two one-standard error bands. In parentheses besides the peak response is the quarter at which it occurs. All reduced form equations include lags 0 to 4 of the 1975:2 dummy. Sample: 1960:1 - 1997:4.

Table 4: Responses to spending shocks.

	1qrt	4qrts	8qrts	12qrts	20qrts	peak
			DT			
GDP	0.84*	0.45	0.54	1.13*	0.97*	1.29*(15)
GCN	1.00*	1.14*	0.95*	0.70*	0.42*	
TAX	0.13	0.14	0.17	0.43*	0.52*	
			ST			
GDP	0.90*	0.55	0.65	0.66	0.66	0.90*(1)
GCN	1.00*	1.30*	1.56*	1.61*	1.62*	
TAX	0.10	0.18	0.33	0.36	0.37	

See notes to Table 3.

Table 5: Stability of responses to tax and spending shocks.

Net taxes		Spending	
excl.	max. GDP	excl.	max. GDP
period	response	period	response
DT			
60-69	-1.18*(1)	60-69	1.44*(1)
70-79	-0.90*(5)	70-79	1.47*(10)
80-89	-0.49*(2)	80-89	0.96*(3)
90-97	-1.45*(7)	90-97	1.73*(12)
ST			
60-69	-1.45*(11)	60-69	1.25*(1)
70-79	-1.48*(4)	70-79	0.62*(1)
80-89	-0.83*(7)	80-89	1.80*(3)
90-97	-1.52*(7)	90-97	0.85*(12)

See notes to Table 3.

Table 6: Responses of GDP components.

	1qrt	4qrts	8qrts	12qrts	20qrts	peak
			DT, TAX			
GDP	-0.69*	-0.74*	-0.72*	-0.42*	-0.22	-0.78*(5)
GCN	-0.05*	-0.12*	-0.24*	-0.26*	-0.16*	-0.05*(1)
CON	-0.18*	-0.35*	-0.32*	-0.23*	-0.20*	-0.35*(5)
INV	-0.36*	-0.00	-0.00	0.18*	0.16*	-0.36*(1)
EXP	-0.02	0.01	-0.01	0.02	0.05	-0.08 (3)
IMP	-0.01	0.02	-0.14*	-0.06	0.04	-0.14*(7)
SUM	-0.60	-0.48	-0.43	-0.23	-0.18	-0.60 (1)
			ST, TAX			
GDP	-0.70*	-1.07*	-1.32*	-1.30*	-1.29*	-1.33*(7)
GCN	-0.06*	0.04*	-0.01*	-0.00*	-0.00*	0.04*(4)
CON	-0.15	-0.40*	-0.44*	-0.43*	-0.43*	-0.44*(7)
INV	-0.35*	-0.22	-0.30	-0.27	-0.27	-0.35*(1)
EXP	-0.00	-0.01	-0.06	-0.07	-0.07	-0.10 (3)
IMP	-0.01	-0.02	-0.12	-0.12	-0.11	-0.13 (3)
SUM	-0.55	-0.57	-0.68	-0.66	-0.66	-0.73 (6)
			DT, SPE			
GDP	0.84*	0.45	0.54	1.13*	0.97*	1.29*(15)
GCN	1.00*	1.14*	0.95*	0.70*	0.42*	1.14*(4)
CON	0.50*	0.63*	0.91*	1.21*	0.90*	1.26*(14)
INV	-0.03	-0.75*	-0.69*	-0.41*	-0.35*	-1.00*(5)
EXP	0.20*	-0.47*	-0.76*	-0.70*	-0.06	-0.80*(9)
IMP	0.64*	-0.19*	-0.46*	-0.42*	-0.16*	-0.49*(9)
SUM	1.03	0.74	0.86	1.22	1.07	1.39 (15)
			ST, SPE			
GDP	0.90*	0.55	0.65	0.66	0.66	0.90*(1)
GCN	1.00*	1.30*	1.56*	1.61*	1.61*	1.00 (1)
CON	0.33*	0.34	0.42	0.43	0.44	0.46*(2)
INV	0.02	-0.74*	-0.97*	-0.96*	-0.95*	-0.98*(9)
EXP	0.17*	-0.16	-0.30	-0.37*	-0.37	-0.37*(13)
IMP	0.56*	0.03	-0.06	-0.05	-0.04	-0.08 (9)
SUM	0.95	0.72	0.77	0.76	0.78	0.95 (1)

Sample: 1960:1 - 1997:4. See also notes to Table 3.

41
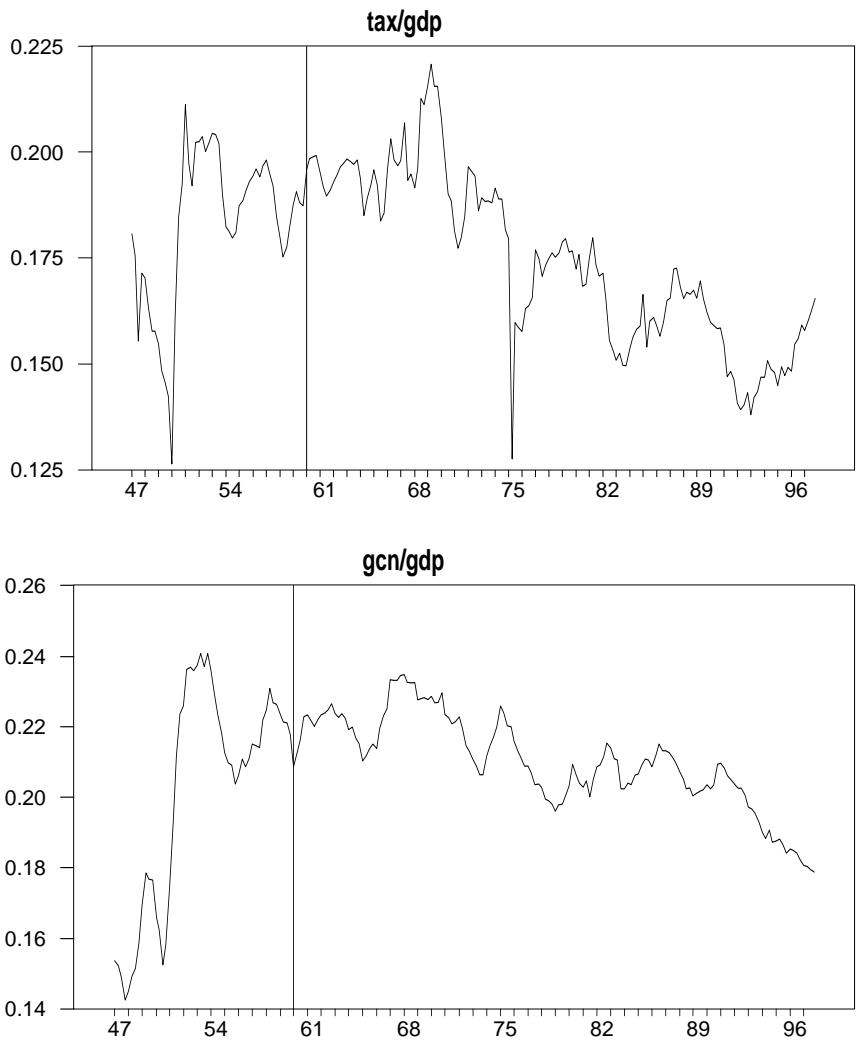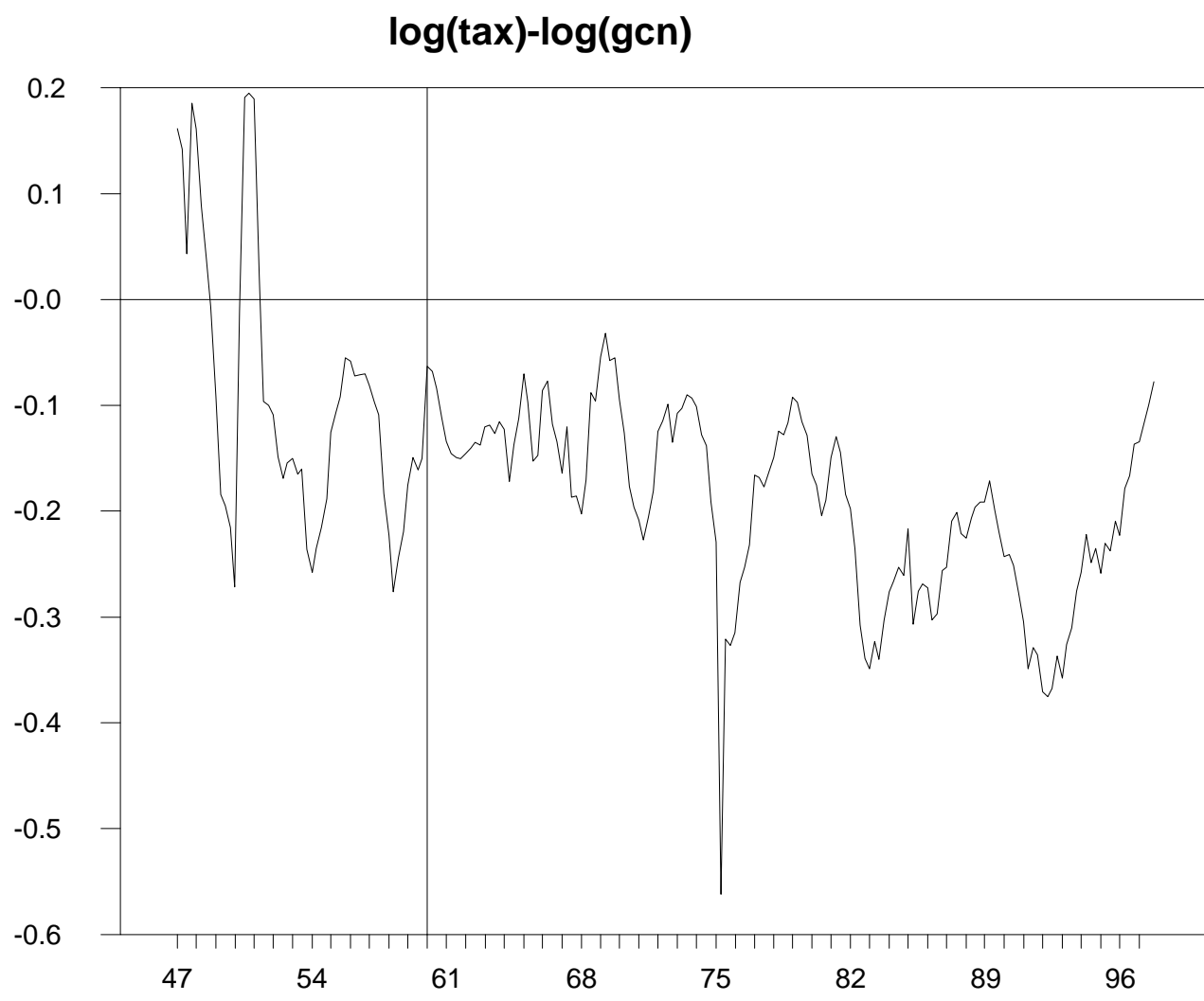
Figure 1: Net taxes and spending, shares of GDP

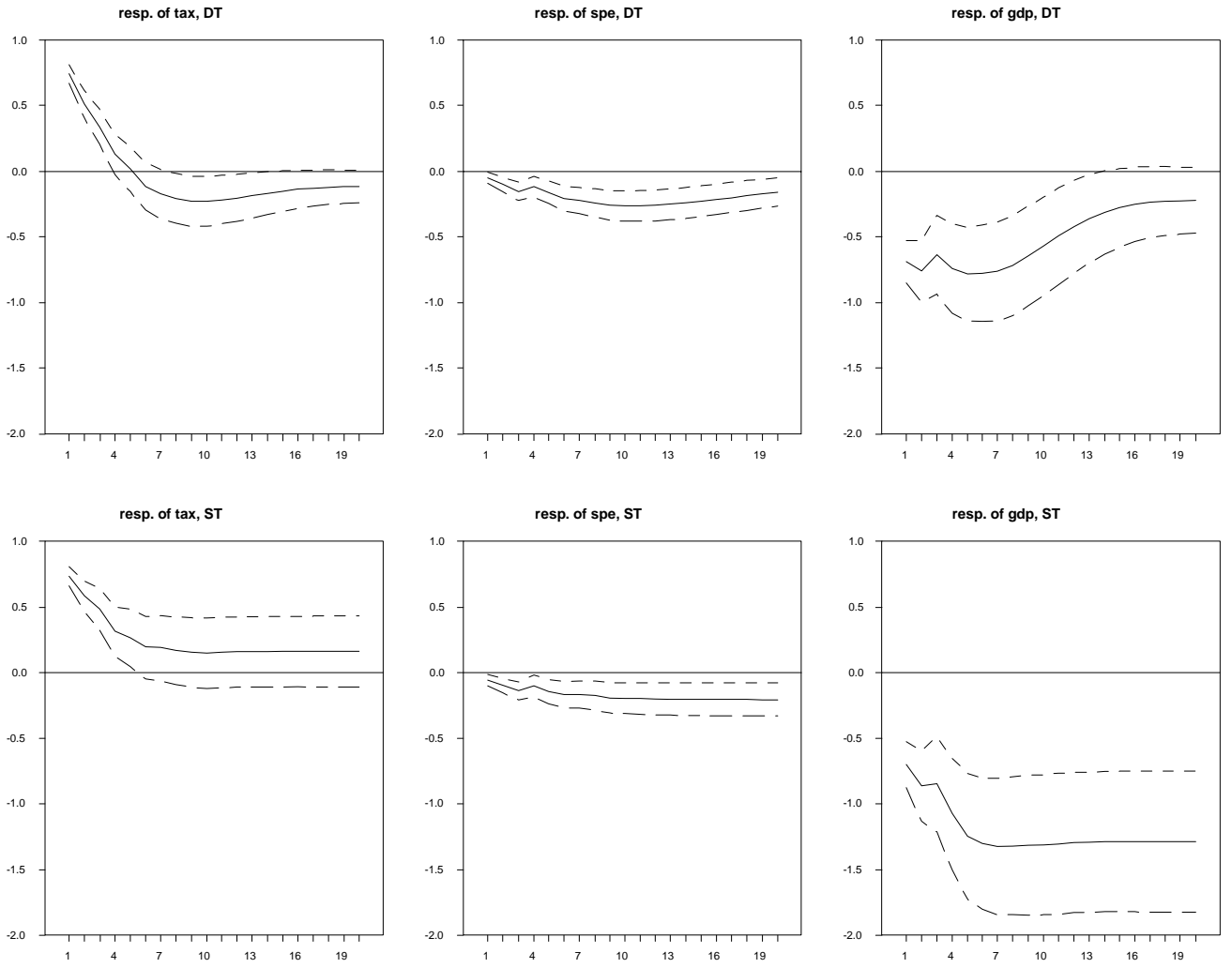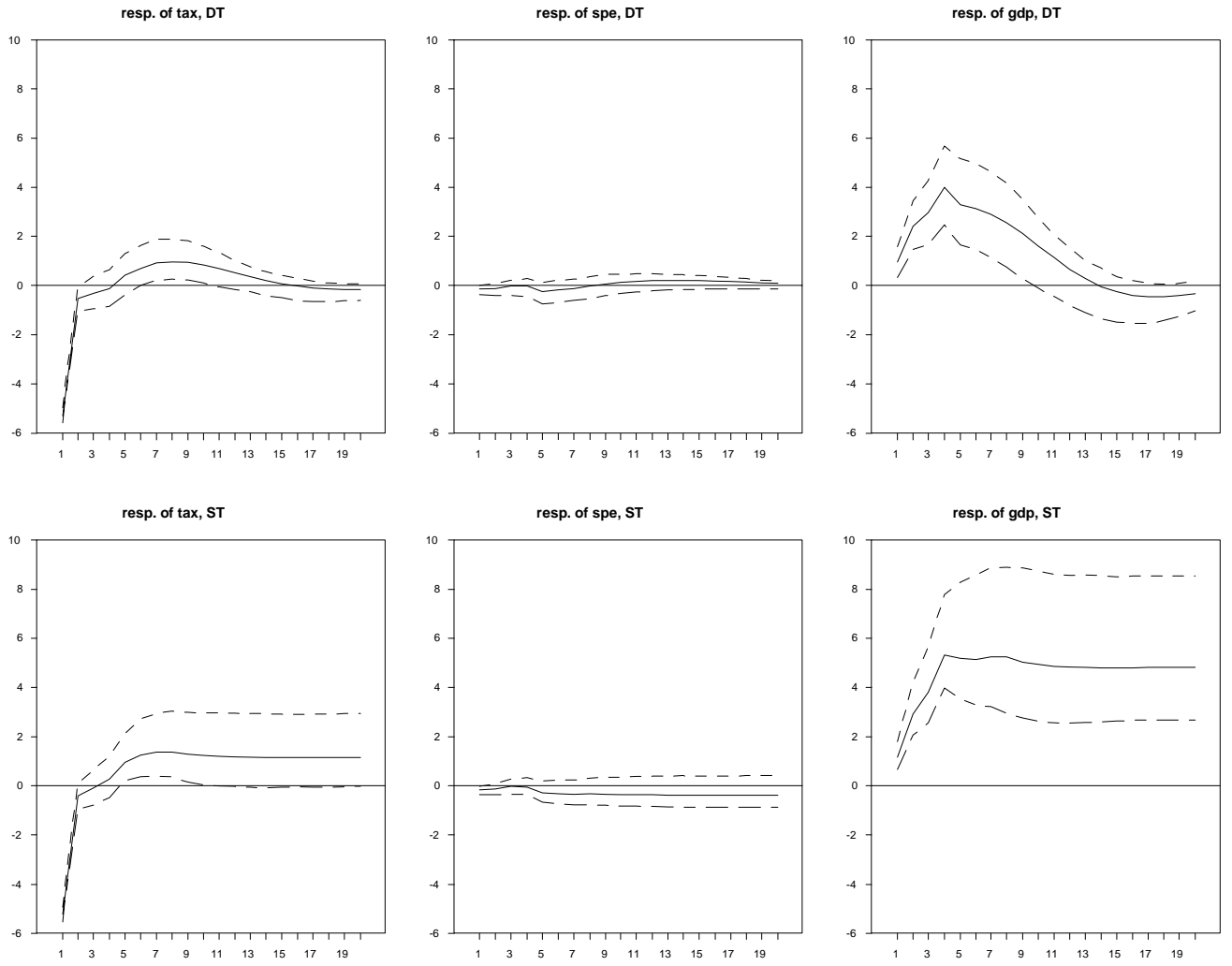Figure 2: log(net taxes) - log(spending)

43

Figure 3: Response to a tax shock
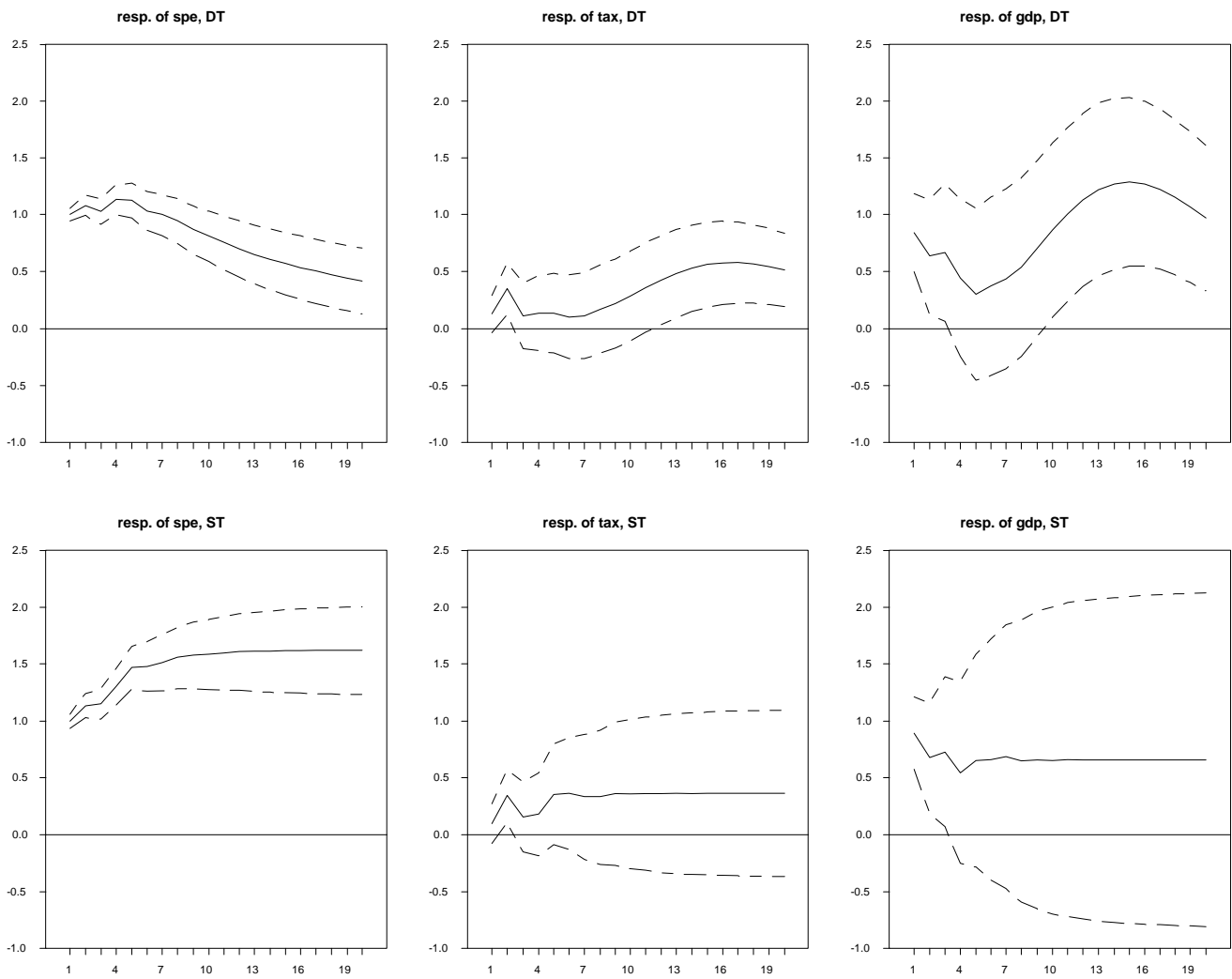
44

Figure 4: Response to a shock to the 1975:2 dummy

45

Figure 5: Response to a spending shock
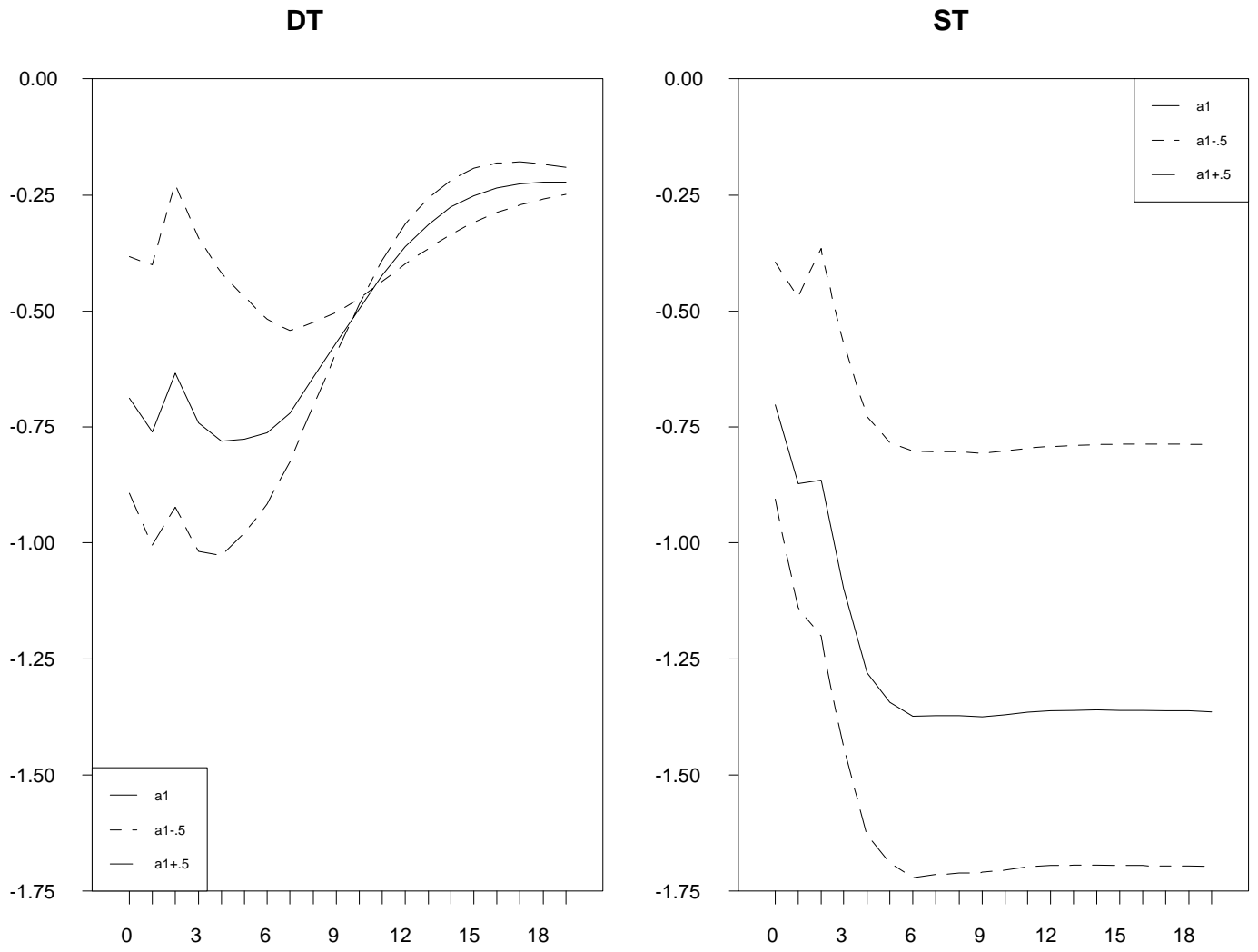
46

DT

ST

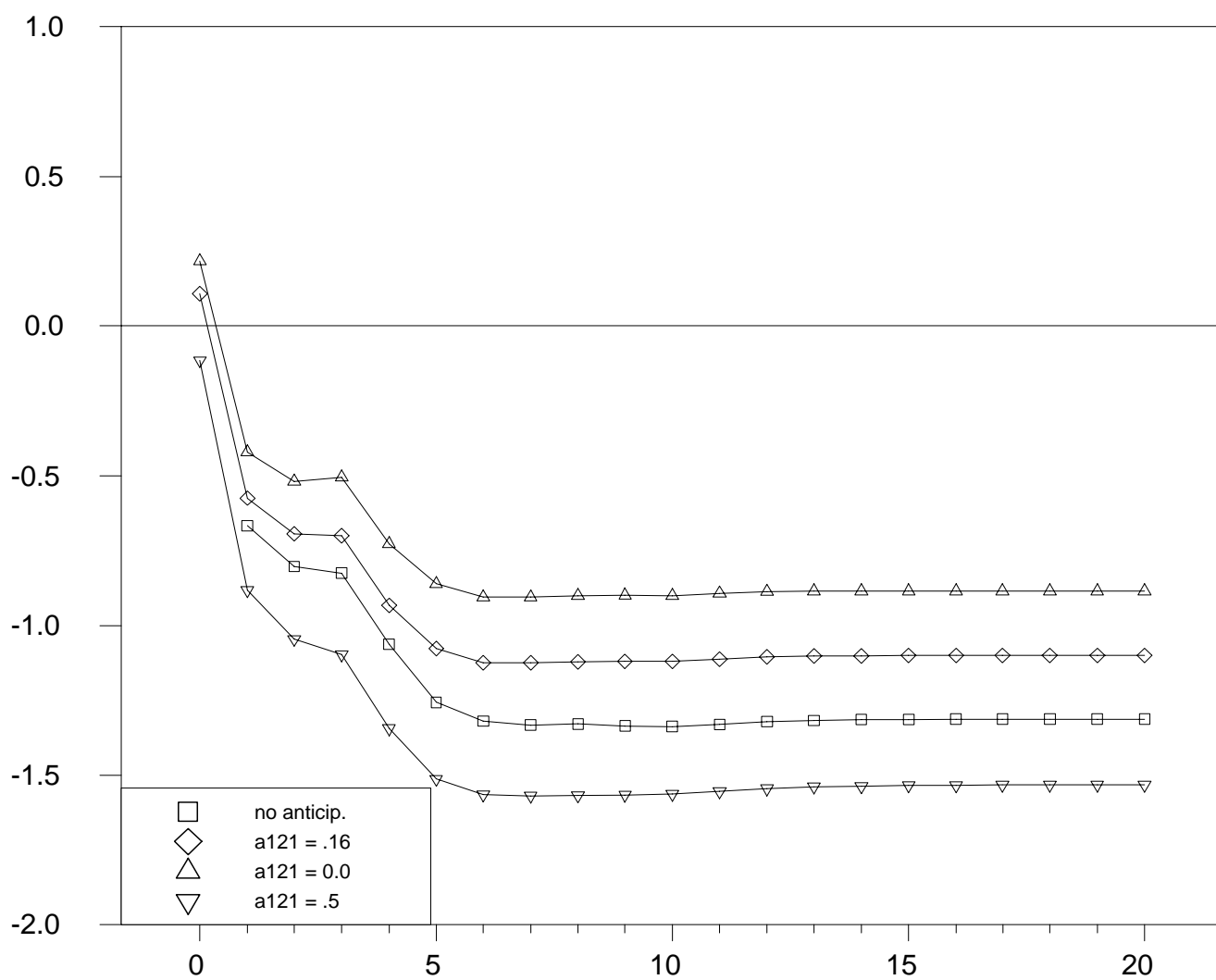

Figure 6: Alternative tax elasticities

47

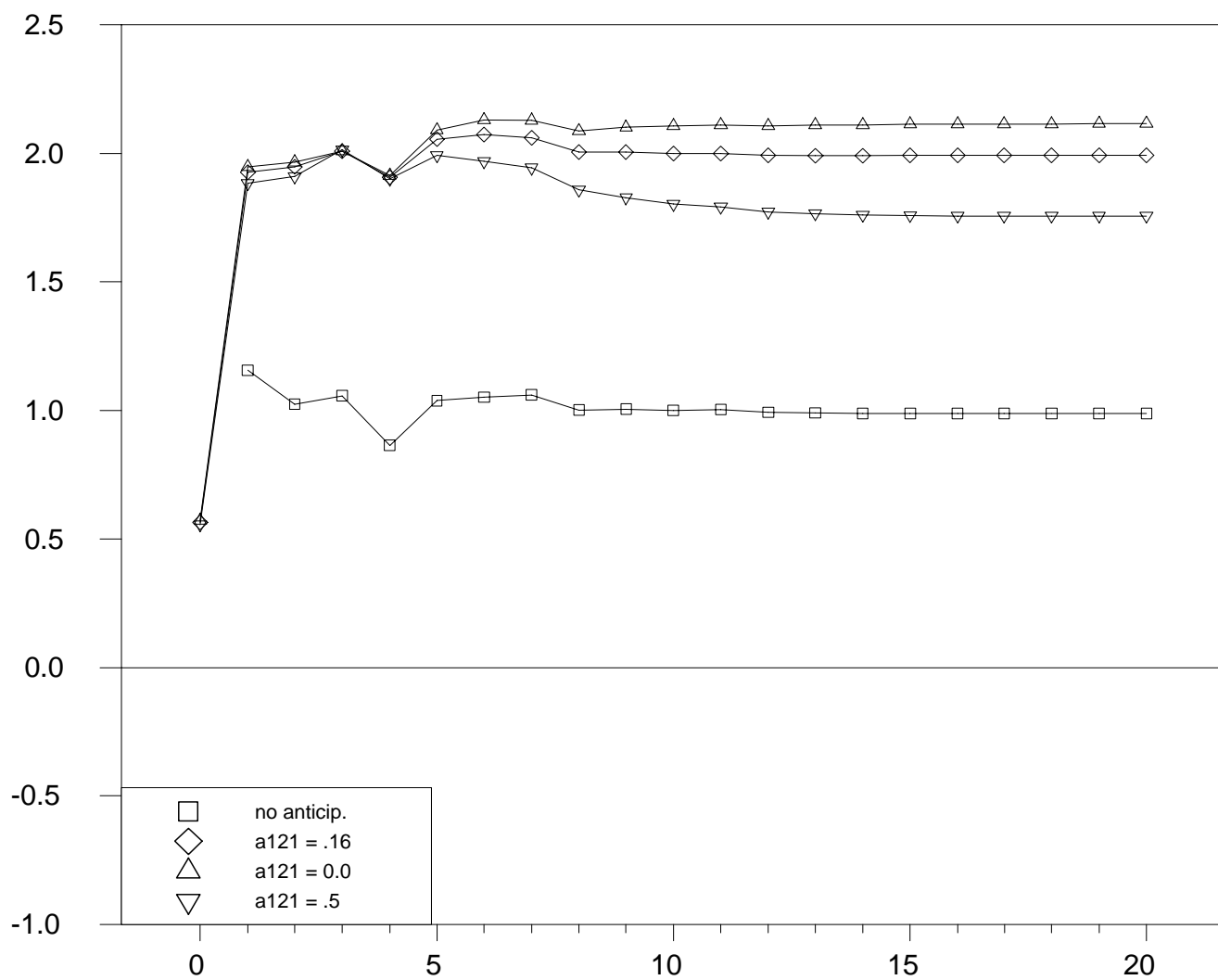Figure 7: Response of output to an anticipated tax shock, ST
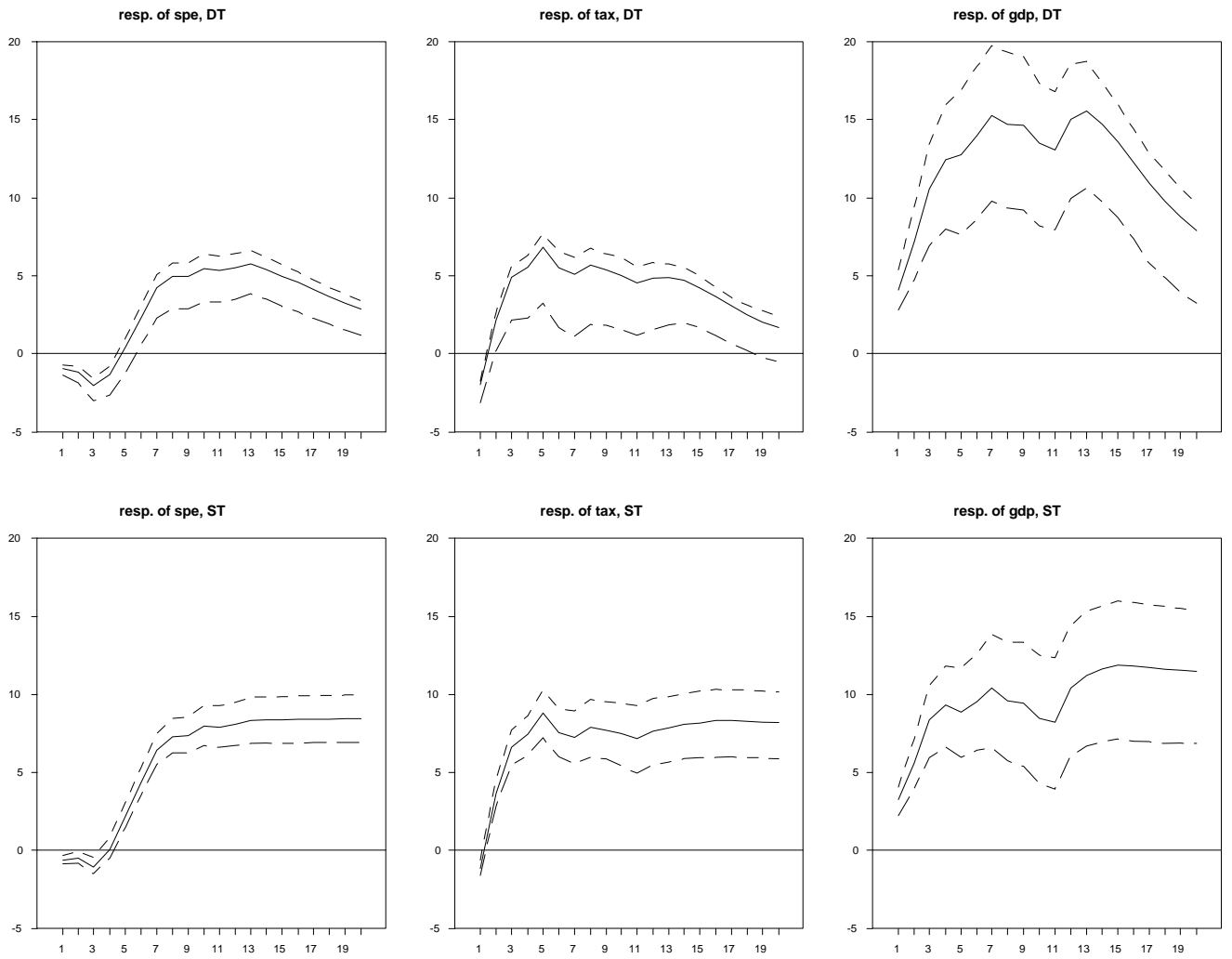
48

Figure 8: Response of output to an anticipated spending shock, ST

49

Figure 9: Response to a shock to the 1950:1 dummy