

DO TAX CUTS STARVE THE BEAST?
THE EFFECT OF TAX CHANGES ON GOVERNMENT SPENDING

Christina D. Romer

David H. Romer

University of California, Berkeley

August 2007

We are grateful to Alan Auerbach, Raj Chetty, and Barry Eichengreen for helpful comments and suggestions, and to the National Science Foundation for financial support.

DO TAX CUTS STARVE THE BEAST?
THE EFFECT OF TAX CHANGES ON GOVERNMENT SPENDING

ABSTRACT

The hypothesis that decreases in taxes reduce future government spending is often cited as a reason for cutting taxes. However, because taxes change for many reasons, examinations of the relationship between overall measures of taxation and subsequent spending are plagued by problems of reverse causation and omitted variable bias. To deal with these problems, this paper focuses on the behavior of spending following legislated changes in taxes not motivated by current or planned changes in spending, a desire to reduce a persistent budget deficit, or short-run macroeconomic considerations. The results provide no support for the hypothesis that tax cuts restrain government spending; indeed, they suggest that tax cuts may actually increase spending. The results also indicate that the main effect of tax cuts on the government budget is to induce subsequent legislated tax increases. Examination of four episodes of major tax cuts reinforces these conclusions.

Christina D. Romer
Department of Economics
University of California, Berkeley
Berkeley, CA 94720-3880
cromer@econ.berkeley.edu

David H. Romer
Department of Economics
University of California, Berkeley
Berkeley, CA 94720-3880
dromer@econ.berkeley.edu

In a speech urging passage of the 1981 tax cuts, Ronald Reagan made the following argument:

Over the past decades we've talked of curtailing government spending so that we can then lower the tax burden. Sometimes we've even taken a run at doing that. But there were always those who told us that taxes couldn't be cut until spending was reduced. Well, you know, we can lecture our children about extravagance until we run out of voice and breath. Or we can cure their extravagance by simply reducing their allowance (Address to the Nation on the Economy, 2/5/81, p. 2).¹

This idea that cutting taxes will lead to a reduction in government spending has become a staple of conservative economic orthodoxy. Distinguished economists from Milton Friedman to Robert Barro have argued that the most effective way to shrink the size of government is to “starve the beast” by reducing tax revenues (see, for example, Friedman, 1967; Barro, 2003; and Becker, Lazear, and Murphy, 2003).

Of course, this is not the only view of the effects of tax cuts. Another possibility is that government spending is determined with little or no regard to taxes, and thus does not respond to tax cuts. A third possibility is that tax cuts actually lead to increases in expenditure. One way this could occur is through the “fiscal illusion” effect proposed by Buchanan and Wagner (1977) and Niskanen (1978): a tax cut without an associated spending cut weakens the link in voters' minds between spending and taxes, and so leads them to demand greater spending. Another possible mechanism is “shared fiscal irresponsibility”: if supporters of tax reduction are acting without concern for the deficit, supporters of higher spending may do the same.

The question of how tax cuts affect government spending would seem to be one that should be investigated empirically, not answered as a tenet of political faith. And, indeed, there have been attempts to look at the aggregate relationship between revenues and spending. However, such examinations of correlations are of limited value in determining the effect of revenues on spending. Revenues change for a variety of reasons. Many changes are legislated, but many others occur automatically in response to changes in the economy. And legislated tax changes themselves are motivated by numerous factors.

¹ Presidential speeches are from the database of presidential documents available from John Woolley and Gerhard Peters, [The American Presidency Project](http://www.presidency.ucsb.edu) (www.presidency.ucsb.edu).

Some, such as many increases in payroll taxes, are driven by increases in current or planned spending. Others, such as tax cuts motivated by a belief in the importance of incentives, are designed to raise long-run growth.

The relationship between revenues and spending is surely not independent of the causes of changes in revenues. For example, if spending-driven tax changes are common, a regression of spending on revenues will almost certainly show a positive correlation. But this relationship does not show that tax changes cause spending changes; causation, in fact, runs in the opposite direction. To give another example, if automatic and legislated countercyclical tax changes are common, one might expect to see government expenditures rising after declines in revenue, because spending on unemployment insurance and other relief measures typically rises in bad economic times. In this case, both revenues and spending are being driven by an omitted variable: the state of the economy. These examples suggest that looking at the aggregate relationship between revenues and spending without accounting for the causes of revenue changes may lead to biased estimates of the effect of revenue changes on spending.

This paper therefore proposes a test of the starve the beast hypothesis that accounts for the motivations for tax changes. In previous work (Romer and Romer, 2007a), we identified all significant legislated tax changes in the United States since 1945. We then used the narrative record – presidential speeches, executive branch documents, Congressional reports, and records of Congressional debates – to identify the key motivation and the expected revenue effects of each action. In this paper, we use our classification of motivation to isolate tax changes that can legitimately be used to examine the effect of revenue changes on spending from those that are likely to give biased estimates. In particular, we focus on the behavior of spending following tax changes taken for long-run purposes. These are changes in taxes that are explicitly not tied to current spending changes or the current state of the economy. They are, instead, taken to promote various long-run objectives, such as spurring productivity growth, improving efficiency, or, as in the case of the Reagan tax cut discussed above, shrinking the size of government. Examining the behavior of government spending following these long-run tax changes should provide the fairest and least biased test of the starve the beast hypothesis.

Our key empirical test involves regressing the percentage change in real federal government expenditures on the contemporaneous and lagged values of our measure of long-run tax changes. The results are striking. We find no support for the hypothesis that a relatively exogenous decline in taxes lowers government spending over the next five years. In our baseline specification, the estimates in fact suggest a substantial and marginally significant positive impact of tax cuts on government spending. Variations on the baseline specification do not change the sign of the estimates, but often make them no longer significant. The finding of a lack of support for the starve the beast hypothesis, however, is highly robust.

We also examine the behavior of different types of spending following long-run tax cuts. We find that defense spending rises significantly. Some other types of spending, such as current transfer payments, fall slightly, but the effects are not precisely estimated. These findings suggest that tax cuts may affect the composition of government spending.

The finding that spending does not fall in response to tax cuts raises an obvious question: How does the government budget eventually balance? One possibility is that what gives in response to a tax cut is not spending but the tax cut itself. To investigate this possibility, we examine the response of both tax revenues and tax legislation to long-run tax cuts. We find that revenues fall in response to a long-run tax cut in the short run, but then recover after about two years. Examination of our measures of legislated tax changes shows that most of a tax cut taken for long-run purposes is typically counteracted by other types of legislated tax increases within the next several years.

While there are a large number of long-run tax changes spread fairly uniformly over the postwar era, four stand out as the largest and most well known: the tax cut passed over Harry Truman's veto in the Revenue Act of 1948; the Kennedy-Johnson tax cut legislated in the Revenue Act of 1964; the Reagan tax cut contained in the Economic Recovery Tax Act of 1981; and the Bush tax cuts passed (along with some countercyclical actions) in 2001 and 2003. As a check on our analysis, we examine these four episodes in detail. We find that the behavior of spending and taxes in these extreme episodes is consistent with the aggregate regressions. Perhaps more importantly, we find that policymakers often did not even

talk as if their spending decisions were influenced by revenue developments. They did, however, often invoke the tax cuts as a motivation for later tax increases. Finally, we find there were concurrent developments, namely wars, that account for some of the rise in defense spending in these episodes. But, there were other concurrent developments that led to unusually low spending changes or that caused measured spending changes to understate the effects of spending decisions that were taken in these episodes. As a result, it is unlikely that failure of total expenditures to fall after these tax cuts was due to chance or unobserved factors.

As discussed above, ours is not the first study to investigate the starve the beast hypothesis. But, despite the importance of the hypothesis in motivating changes in fiscal policy in recent decades, relatively little work has attempted to evaluate it. The most common approach is some variation of a regression of spending on lagged revenues; examples include Anderson, Wallace, and Warner (1986) and Ram (1988). Two more sophisticated versions of this methodology are pursued by Bohn (1991) and Auerbach (2000, 2003). Bohn, focusing on a long sample period that is dominated by wartime budgetary changes, examines the interrelationships between revenues and spending in a vector autoregression framework that allows for cointegration between the two variables (see also von Furstenberg, Green, and Jeong, 1986, and Miller and Russek, 1990). Auerbach, focusing on recent decades, studies the relationship between policy-driven changes in spending (rather than all changes in spending) and past deficits or projections of what future deficits would be if policy did not change (see also, Calomiris and Hassett, 2002).²

The results of these studies are mixed, but for the most part they suggest that tax cuts are followed by reductions in spending. None of these studies, however, consider the different reasons for changes in revenues, and thus none isolate the impact of independent tax changes on future spending. Indeed, our results point to a potentially important source of bias in studies using aggregate data. We find that the only type of legislated tax changes that are systematically followed by spending movements in the

² In contrast to the other studies, which focus on federal budget data, Holtz-Eakin, Newey, and Rosen (1989) estimate the temporal correlation between taxes and spending at the municipal level.

same direction are ones motivated by decisions to change spending. Since causation in these cases clearly does not run from the tax changes to the spending changes, this relationship is not informative about the starve the beast hypothesis. We also find that this type of tax change is sufficiently common to make the overall relationship between legislated tax changes and subsequent spending changes substantially positive.³

The rest of the paper is organized as follows. Section I discusses the different motivations for tax changes, and which ones are best suited for testing the starve the beast hypothesis. Section II analyzes the relationship between tax changes and both the overall level of expenditures and the different categories of spending. Section III examines how changes in taxes affect future tax revenues and tax legislation. Section IV discusses spending and taxes in four key episodes. Section V presents our conclusions.

I. THE MOTIVATIONS FOR LEGISLATED TAX CHANGES AND TESTS OF THE STARVE THE BEAST HYPOTHESIS

Legislated tax changes classified by motivation are a key input into our tests of the starve the beast hypothesis. Therefore, it is important to describe how we classify motivation, and to discuss which types of tax changes are likely to yield relatively unbiased estimates of the effects of tax changes on government spending. It is also useful to provide a brief overview of our identification of the motivations for tax changes, and of our findings about the patterns of legislated tax changes in the postwar era.

³ One can also test the starve the beast hypothesis indirectly. Perhaps the best known study of this type is Becker and Mulligan (2003). They show that under appropriate assumptions, the same forces that would give rise to a starve the beast effect would cause a reduction in the efficiency of the tax system to reduce government spending. They therefore examine the cross-country relationship between the efficiency of the tax system and the share of government spending in GDP. While they find a strong positive relationship, the correlation between efficiency and spending, like that between taxes and spending, may reflect reverse causation or omitted variables. That is, countries may invest in efficient tax systems because they desire high government spending, or a third factor, such as tolerance of intrusive government or less emphasis on individualism, may lead both to a broader, more comprehensive tax system and to higher government spending.

A. Classification of Motivation

Our classification and identification of the motivations for postwar legislated tax changes are described in detail in Romer and Romer (2007a). That paper shows that the motivations for almost all tax changes have fallen into four broad categories. It also shows, perhaps surprisingly, that most postwar tax changes have had a single, fairly clearly identifiable, dominant type of motivation. And it shows that in the small number of cases where more than one type of motivation was important, it is possible to construct reasonable estimates of the portion of the change that was due to each type.

One type of tax changes are those motivated by countercyclical considerations. These are changes made because policymakers believe that growth will be above or below normal, and therefore change taxes to try to keep growth at its normal, sustainable level. A classic example of such a countercyclical action is the 1975 tax cut. Taxes were reduced because the economy was in a severe recession and growth was predicted to remain substantially below normal. Countercyclical actions can be either tax cuts or tax increases, depending on whether they are designed to counteract unusually low or unusually high expected growth.

A second type of tax changes are ones motivated by contemporaneous changes in spending. Often, policymakers will introduce a new program or social benefit, and raise taxes to pay for it. This was true, for example, in the late 1950s when the interstate highway system was started, and in the mid-1960s when Medicare was introduced. The key feature of these changes is that the spending changes are the impetus for the tax changes. Typically such changes are tax increases, but spending-driven tax cuts are not unheard of.

A third type of tax changes are those made to reduce an inherited budget deficit. By definition, these changes are all increases. A classic example of this type of change is the 1993 Clinton tax increase. This increase was undertaken not to finance a contemporaneous rise in spending, but to reduce a persistent deficit caused by past developments. Because deficit-driven tax increases tend to be passed as parts of budget packages, they are often accompanied by spending reductions.

The fourth type of tax changes are ones intended to raise long-run growth. This is a broad

category designed to capture changes not made to keep or return growth to normal, and that are not explicitly for deficit reduction. It includes tax changes motivated by a wide range of factors. The most common motivation is a belief that lower marginal tax rates will improve incentives, and thereby raise long-run growth. Another common motivation is a belief in small government and a desire to return the people's money to them. Crass political motivations, such as a desire to push growth above normal to improve the president's chances of reelection, would also fall in this category. Many of the most famous tax cuts, such as the 1964 Kennedy-Johnson tax cut and the Reagan tax cuts of the early 1980s, fall under the heading of tax changes to raise long-run growth. Most long-run tax changes are cuts. But, there have been a few tax reforms that increased revenues and that were designed to improve efficiency, and so fall into this category.

B. Which Tax Changes Are Useful for Testing the Starve the Beast Hypothesis?

This description of the different motivations for legislated tax changes makes it clear that some changes are much more appropriate for testing the starve the beast hypothesis than others. Most obviously, spending-driven changes are clearly not valid observations in this context. Causation in these episodes runs from desired changes in spending to changes in taxes. If we have classified these tax changes correctly, there will be a positive correlation between these changes and spending changes by construction. These observations, however, provide no information on whether tax changes cause policymakers to alter their subsequent spending decisions. Including spending-driven tax changes in a regression would bias the results toward finding a positive effect of tax changes on government spending.

Examining countercyclical and deficit-driven tax changes, on the other hand, might tend to bias the results against the starve the beast hypothesis. In both cases, there may be spending changes that are negatively correlated with the tax changes but that are not caused by them. Rather, for both countercyclical and deficit-driven tax changes, both the tax and spending changes may be caused by a third factor.

In the case of countercyclical actions, the third factor is the state of the economy. In bad times,

policymakers may cut taxes and increase spending as a way of raising aggregate demand. Also, some types of spending, such as unemployment compensation and public assistance, increase automatically in recessions. Thus, the relationship between taxes and spending in these episodes may reflect discretionary and automatic responses to the state of the economy, not a behavioral link between tax revenues and spending decisions.

In the case of deficit-driven tax increases, the unobserved third factor is a general switch to fiscal responsibility. Tax increases to reduce inherited budget deficits are often passed as parts of packages that include spending reductions. The spending reductions are not caused by the tax increases; rather, both are driven by a desire to eliminate the deficit. Inclusion of such budget packages in a regression will tend to bias the results away from supporting the starve the beast hypothesis. This concern may be more important in theory than in reality, however. Our narrative analysis of tax changes documents the spending reductions agreed to in conjunction with deficit-driven tax changes. In almost every case, the spending cuts were small relative to the tax increase. Therefore, while one may want to treat the behavior of spending following deficit-driven tax changes with caution, it may in fact yield relatively unbiased estimates.

The tax changes that are surely the most appropriate for testing the starve the beast hypothesis are those taken to spur long-run growth. As described above, these are typically permanent tax cuts not made in response to current macroeconomic conditions or in conjunction with spending changes. As a result, they are exactly the kind of changes that proponents of the starve the beast hypothesis believe are likely to reduce government spending. Indeed, these are the tax changes for which the induced reduction in future spending is sometimes cited as motivation.

This discussion suggests that to the degree that focusing on these observations may lead to bias, it is likely to be in the direction of finding a positive correlation between taxes and spending. The very fact that proponents of these tax cuts sometimes cite reducing the size of government as a motivation implies that there is a potential correlation between spending and tax changes in these episodes driven by a third factor: belief in limited government. Because this possible omitted variable bias works in the direction of

supporting the starve the beast hypothesis, using these observations gives the hypothesis the best reasonable chance of success.

At the same time, our narrative analysis suggests that this potential bias is likely to be small, for two reasons. First, as with deficit-driven changes, we documented the spending changes agreed to at the time of long-run tax changes, and found them to be uniformly small. Second, the desire for smaller government is rarely the primary stated motivation for long-run tax changes; a desire to reap the efficiency gains of lower marginal tax rates is considerably more common, for example.

C. Overview of the Narrative Analysis

The implementation and results of our narrative analysis of postwar tax changes are described in Romer and Romer (2007a). We use a detailed examination of a wide range of policy documents to identify all significant legislated tax changes in the postwar era. We then identify the motivations policymakers gave for each action. We find that policymakers are often both quite explicit and remarkably unanimous in their stated reasons for undertaking tax actions. Only infrequently do they emphasize multiple motivations. In these cases, we divide the tax changes into pieces reflecting the different motivations.

We also use the narrative sources to estimate the revenue impacts of the actions. Specifically, we determine how policymakers expected the actions to affect tax liabilities. Very often, tax bills change taxes in a number of steps. In these cases, our baseline revenue estimates show changes in each of the quarters the various provisions took effect. To make the revenue estimates comparable over time, we express them as a percent of nominal GDP.⁴

Two issues arise in the calculation of the revenue effects of various acts. One fairly minor one concerns the treatment of retroactive changes. Tax actions are often retroactive for a quarter or two. When a change is retroactive in this way, it has a much larger effect on liabilities in the initial quarter than

⁴ The nominal GDP series used to normalize the revenue estimates is from NIPA Table 1.1.5, www.bea.gov, downloaded 8/8/07. Quarterly nominal GDP data are only available after 1947. We therefore normalize the two tax changes in 1946 using the annual NIPA figure.

in subsequent ones. In differences, this results in a large movement in one direction in the initial quarter and a partially offsetting movement in the next quarter. For this study, which examines the longer-run responses of spending and future taxes, the short-run volatility caused by these changes may unnecessarily complicate the analysis. We therefore use the revenue estimates ignoring the retroactive changes as our baseline estimates, and use those including the retroactive changes only in a robustness check.

A more substantive issue involves the dating of the revenue effects. In our baseline series, we date tax changes in the quarter that liabilities actually change. An obvious alternative is to date all tax changes associated with a given bill in the quarter that the law was passed. This alternative involves calculating the present value of tax changes scheduled to occur at specific dates in the future. We consider the effects of this alternative dating in another robustness check.

Figure 1 shows our baseline measure of legislated postwar tax changes classified by motivation. Long-run changes, which are the key actions for our purposes, are shown in blue. The graph makes clear that the vast majority of long-run tax actions are cuts. Only a few, such as those legislated in the Tax Reform Act of 1986, are increases. The graph also makes clear that long-run tax changes have been fairly evenly distributed over the postwar era. The largest long-run tax changes were the tax cuts in the late 1940s, the mid-1960s, the early 1980s, and the early 2000s.

Deficit-driven tax changes are shown in red in Figure 1. While there were a number of small deficit-driven increases in the first half of the postwar era, the vast majority of these changes took place in the 1980s and early 1990s. Spending-driven changes are shown in green. Tax changes explicitly tied to spending changes are typically tax increases, and were both frequent and relatively large in the first half of the postwar era. Finally, countercyclical tax changes, which are shown in yellow, are typically large. Until being resurrected by George W. Bush in 2001 and 2002, such explicitly countercyclical changes were confined to the fairly short period 1965-1975.

II. THE EFFECTS OF TAX CHANGES ON EXPENDITURES

The previous section describes our identification of legislated tax changes motivated by concern about long-run growth. This section investigates the relationship between these relatively exogenous tax changes and subsequent changes in government spending. It examines the effects of tax changes on both total expenditures and the components of spending. We also investigate the behavior of spending following other types of tax changes to see if there is evidence of bias when these tax measures are used.

A. Specification

To estimate the effects of tax changes on government spending, for the most part we employ simple regressions using quarterly data of the form:

$$(1) \quad \Delta \ln E_t = a + \sum_{i=0}^N b_i \Delta T_{t-i} + e_t,$$

where E is real government expenditure of some type and ΔT is our measure of long-run tax changes (expressed as a percent of nominal GDP).

The key feature of our measure of long-run tax changes is that it is based on actions motivated by considerations unrelated to current spending, current macroeconomic conditions, or an inherited budget deficit. Our discussion of why such long-run changes provide the best test of the starve the beast hypothesis suggests that they are unlikely to be systematically correlated with other factors affecting spending. It is for this reason that our baseline specification includes no control variables. However, it is certainly possible that there are correlations in small samples. In our robustness checks, we therefore consider a variety of alternative specifications.

We include a number of lags of the tax variable to allow for the possibility that the response of spending to tax changes may be quite delayed or gradual. In our baseline specification we set N to 20, and so look at the response of spending over a five-year horizon. Because the starve the beast hypothesis

does not make predictions about the exact timing of how spending responds to tax changes, we focus on the cumulative effect of a tax change on expenditures at various horizons. We summarize the regression results by reporting the implied impact of a tax cut of one percent of GDP on the path of expenditures (in logarithms). For our baseline specification, the cumulative impact after n quarters is just minus the sum of the coefficients on the contemporaneous value and first n lags of the tax variable. The starve the beast hypothesis predicts that tax cuts reduce spending. Therefore, the estimated cumulative impact of a tax cut on expenditures should be negative if the hypothesis is correct.

Our data on tax changes begin in 1945Q1, and the data on expenditures begin in 1947Q1. The data for both series extend through 2006Q4. Therefore, in the baseline specification, where we include twenty lags of the tax variable, the longest possible sample is 1950Q1–2006Q4. However, previous work has found some evidence that the behavior and effects of fiscal policy were unusual in the Korean War period (see, for example, Blanchard and Perotti, 2002, and Romer and Romer, 2007b). We therefore also consider the shorter sample 1957Q1–2006Q4.

B. The Effects of Long-Run Tax Changes on Total Expenditures

Measure of Aggregate Expenditures. The most basic relationship we want to consider is that between our measure of relatively exogenous, long-run tax changes and overall government spending. We use quarterly data on government expenditures from the National Income and Product accounts. Our series on long-run tax changes refers only to federal legislation. Therefore, we consider only the behavior of federal expenditures.

What the Bureau of Economic Analysis calls “total expenditures” includes two components that are not appropriate in thinking about the response of spending to tax changes. One is a deduction for the consumption of fixed capital (that is, depreciation). This clearly cannot be affected by current spending decisions, and so could not possibly show a starve the beast type of response. Thus, for a fair test of the hypothesis we do not subtract depreciation. The other component is interest payments on government debt. For a given interest rate, interest payments rise with the amount of debt. As a result, any tax cut

that increases the deficit will almost certainly increase interest payments, even if other types of spending respond strongly. Again, to provide a fair test of the starve the beast hypothesis, we exclude this type of spending. The resulting aggregate that we consider is thus total gross expenditures less interest. For simplicity, we refer to this as total expenditures in what follows.⁵

The NIPA expenditure data are expressed in nominal terms. For some components, such as government consumption expenditures and gross government investment, corresponding deflators exist. However, for some categories, especially those involving transfers, there are no associated deflators. We therefore deflate total gross expenditures less interest by the price index for GDP.⁶

Baseline Results. Table 1 shows the results of estimating equation (1) for total expenditures using twenty lags of the long-run tax variable over the full postwar sample. The coefficient estimates for the individual lags fluctuate between positive and negative. As one would expect, few of the individual coefficients are statistically significant. The overall fit of the regression is modest ($R^2 = 0.20$).

Figure 2 summarizes the results by showing the implied response of total expenditures to a long-run tax cut of one percent of GDP, together with the one-standard-error bands. There is no evidence of a starve the beast effect. The cumulative effect is negative in the quarter of the tax cut and the subsequent three quarters, as the starve the beast hypothesis predicts, but very small, and the t-statistics do not rise above 0.7 in absolute value. After that, the estimated cumulative effect is positive at every horizon except quarters 9 and 10, suggesting fiscal illusion or shared fiscal irresponsibility.

The estimated positive impact of the tax cut on spending is often substantial. Since federal government spending averages roughly twenty percent of GDP in our sample, a tax cut of one percent of GDP is equal to about five percent of government spending. The point estimates suggest that such a tax cut raises spending by four percent or more in quarters 13 through 20. That is, they suggest that spending eventually rises by almost the amount of the tax cut. However, the standard errors are substantial. The t-statistics for the cumulative impact of the tax cut on spending at horizons more than three years are

⁵ Data on total expenditures, consumption of fixed capital, and interest payments are from NIPA Table 3.2, downloaded 8/8/07.

⁶ The price index for GDP is from NIPA Table 1.1.4, downloaded 8/8/07.

generally between 1.5 and 2, and exceed 2 for only one horizon (quarter 14, for which the t-statistic is 2.2).

Robustness. The next step is to examine the robustness of the findings. We consider robustness along several dimensions. Our most important checks are summarized in Figure 3, which shows the implied response of total expenditures to a long-run tax cut of one percent of GDP for a number of variants of the baseline regression.

One obvious concern is the possible importance of the sample period and outliers. As described above, fiscal policy was very unusual in the Korean War period. Panel (a) shows that considering the post-Korea sample weakens the evidence for a perverse effect of tax cuts on spending, but still yields no evidence of a starve the beast effect. The change in the sample makes the initial negative impact even smaller and more insignificant. The response in quarters 4 through 20 is always positive, but considerably smaller than for the full sample and not even marginally significant. To check more generally for the possible influence of outliers, we consider the effects of excluding each of the four largest long-run tax cuts.⁷ In all four cases, the estimated effect of a tax cut on spending remains mainly positive and is never close to significantly negative at any horizon. Dropping the 1948 tax cut, however, renders the positive effect of tax cuts on spending small and insignificant.⁸

A second concern is the horizon over which tax changes may affect spending. Perhaps the main effects of tax changes occur outside the five-year horizon considered in our baseline regression. To test for this, we extend the baseline regression to include forty lags of the tax variable and estimate it over the longest feasible sample (1955Q1–2006Q4). Panel (b) of Figure 3 shows that for horizons beyond five years, the estimated cumulative impact of a tax cut of one percent of GDP on total expenditures is always small, fluctuates between positive and negative, and is never remotely close to statistically significant.

⁷ To exclude a tax cut, we set our series for long-run tax changes to zero from the first to the last quarter in which the bill changed taxes. We treat the 2001 and 2003 cuts as a single measure; thus in this case, we set our series to zero from 2002Q1 to 2005Q1.

⁸ A related exercise along these lines is to split the sample in 1980Q4. For the period 1950Q1–1980Q4, the estimates suggest a large and statistically significant positive effect of tax cuts on spending. For the period 1981Q1–2006Q4, the estimated effects are again virtually always positive, but consistently small and far from significant.

Thus, this change provides no evidence that tax cuts reduce government spending, but also fails to support the hypothesis that they increase it.

A third issue concerns our specification for the dynamics of the normal evolution of spending. Our baseline specification includes only a constant term, implicitly assuming that other influences on spending growth are not serially correlated. Panels (c) and (d) show the results of two approaches to providing for richer dynamics. The first approach is simply to add lagged values of the percentage change in total expenditures to the regression. Including lagged expenditure growth allows for more complicated spending behavior and helps to control for any unobserved serially correlated influences on spending. Since the data on expenditures begin in 1947Q1, we can include eleven lags of the percentage change in expenditures and still retain our baseline sample of 1950Q1–2006Q4. The response of total expenditures to a tax cut for this specification, shown in panel (c), now includes both the direct effect of the tax cut and the indirect effect through lagged expenditure growth.⁹ We carry these dynamic multipliers out to 24 quarters to see if the indirect effects are important. We find, however, that the estimated effect is virtually identical to that in the baseline regression: there is no evidence of a starve the beast effect, and some evidence that tax cuts actually increase government spending.

Our second approach to providing for richer spending dynamics is to estimate a vector autoregression (VAR) with real expenditures and our series for long-run tax changes. For consistency with our regressions, we put the tax changes first and expenditures second, so tax changes can affect spending within the quarter. We enter expenditures in log levels; given the availability of the data, this allows us to include twelve lags while still using our baseline sample. The estimated response of spending to an innovation of minus one percentage point to the tax series, shown in panel (d), is similar to the estimated effect of a long-run tax cut of one percent of GDP in the baseline specification.¹⁰ The point

⁹ We compute the standard errors by taking 10,000 draws of the vector of coefficient estimates from a multivariate normal distribution with mean and variance-covariance matrix given by the point estimates and variance-covariance matrix of the coefficient estimates, and then finding the standard deviation of the implied responses at each horizon.

¹⁰ The standard errors are computed in the same way as for the regression that includes lagged expenditure growth (see n. 9). Note that the experiment considered in panel (d) is slightly different than the experiment of a one-time tax cut of one percent of GDP with no further changes considered in the other panels and in the other figures. In

estimates suggest that the tax cut reduces spending in the short run but then raises it, with a fairly large positive long-run effect. None of the estimated effects are statistically significant. Thus again there is no support for the starve the beast hypothesis. Another finding from the VAR is that the estimated cumulative response of the tax series to an innovation to government spending is very small and highly insignificant at all horizons. This supports the view that long-run tax changes are not responses to spending developments.¹¹

A fourth robustness issue concerns the role of political variables. It is certainly possible that the party of the president or the existence of united government has an influence on government spending. If such variables happen to be correlated with our tax measure, the baseline regression could suffer from omitted variable bias. For this reason, we try adding a variety of political variables to our baseline specification. To give one example, panel (e) of Figure 3 shows the effect of a tax cut on spending when the contemporaneous value and twelve lags of a dummy variable for Democratic administrations are included in the regression. This regression asks whether tax cuts lower spending, taking into account that Democratic presidents may consistently spend more or less than their Republican counterparts. Adding this variable weakens the evidence for fiscal illusion or shared fiscal irresponsibility slightly. The estimated impact of a tax cut on spending remains generally positive, but is slightly smaller than before and never significant. We also consider specifications including a dummy variable for whether the presidency and both houses of Congress are controlled by the same party, and including separate dummies for whether it is the first year of a new Republican or a new Democratic administration. Neither specification provides any support for the starve the beast hypothesis.

A fifth robustness issue involves the specification of our tax variable. Our baseline series dates revenue changes in the quarter that liabilities actually change. An alternative measure, which emphasizes

response to the innovation to our tax measure in the VAR, the cumulative change in the tax measure itself does not remain constant at -1 percentage point; instead it rises from -1 percentage point to roughly -1.25 percentage points after four quarters and then fluctuates around that level.

¹¹ We have also estimated a VAR with twenty lags for the period 1952Q1–2006Q4. The estimated effects of a tax cut on spending in this specification are even more consistently positive, and are marginally significant. The maximum effect is an increase of 3.9 percent after 18 quarters ($t = 1.9$).

expectational effects, calculates the present discounted value of all revenue changes called for by a given piece of legislation, and dates the revenue change in the quarter the law was passed.¹² Panel (f) of Figure 3 shows that the starve the beast hypothesis fares even worse when this alternative tax measure is used. The estimated impact of a tax cut on spending is consistently in the opposite direction from the predictions of the starve the beast hypothesis, often quantitatively large, and sometimes marginally significant. The baseline results are also robust to the treatment of retroactive tax changes. As discussed above, many tax bills have quite volatile (and negatively serially correlated) revenue effects in the first two quarters because of retroactive features. Using the version of our series that includes these retroactive features, however, has little impact on the results.

A final robustness issue involves the appropriate specification of the spending variable. In all of the specifications discussed so far, we have looked at the response of the percentage change in real expenditures to long-run tax changes. The cumulative impact therefore shows the effect of a tax change on the level of real expenditures (relative to normal). We feel this is the appropriate measure for testing the starve the beast hypothesis: Does a tax cut change the spending decisions of policymakers? However, an alternative form of the hypothesis could be that a tax cut leads to a reduction in expenditures as a percent of GDP. In this view, a tax cut could reduce the share of spending not by changing policymakers' spending decisions, but by changing output growth.

To test this alternative version of the hypothesis, we re-estimate equation (1) using two different specifications of the dependent variable. The more sensible of the two expresses real total expenditures as a percent of trend real GDP, where trend real GDP is calculated using a conventional Hodrick-Prescott filter, and then uses the change in this variable as the dependent variable in equation (1).¹³ Detrending real GDP is reasonable because, to the extent that a tax cut causes a temporary boom, it will inherently

¹² See Romer and Romer (2007b) for a detailed description of how we calculate the present value of revenue changes.

¹³ We again calculate real expenditures by dividing nominal expenditures by the price index for GDP. Real GDP is constructed by dividing nominal GDP by the same price index. Data on nominal GDP for this exercise are from NIPA Table 1.1.5, downloaded 8/8/07. We fit a Hodrick-Prescott filter ($\lambda = 1600$) to log real GDP for the full postwar sample (1947Q1–2006Q4).

tend to reduce real expenditures as a percent of actual GDP in the short run. We do not believe that this is the mechanism proponents of even the alternative form of the starve the beast hypothesis have in mind. However, as a further robustness check, we also estimate equation (1) using the change in the ratio of total real expenditures to actual real GDP.

The results of these two exercises are shown in panels (g) and (h) of Figure 3.¹⁴ Panel (g) shows that the results using the change in spending as a share of trend GDP are very similar to the results using the percentage change in spending. A tax cut of one percent of GDP for the most part raises the share of spending in GDP. The estimated maximum effect is large, but only marginally significant. Thus, the results again fail to support the starve the beast hypothesis, and provide moderate support for the alternative view of fiscal illusion or shared fiscal irresponsibility.

Panel (h) shows that a tax cut does not even reduce spending as a share of actual GDP. The estimated effects fluctuate irregularly around zero. The estimates suggest a marginally significant starve the beast effect in a single quarter (quarter 9), but they are more often positive than negative, and the estimated long-run effect is positive, small, and very far from significant. That this second specification fails to support the starve the beast hypothesis is quite surprising. As discussed in Romer and Romer (2007b), the short-run stimulatory effects of tax cuts on output are very strong. Yet even this rapid growth of output is not enough to generate a systematic fall in expenditures as a share of GDP.

The robustness checks yield two conclusions. First, and most important, the lack of support for the starve the beast hypothesis is extremely robust: none of the specifications we consider provide evidence that tax cuts reduce government expenditures. Second, the evidence for the alternative view of fiscal illusion or shared fiscal irresponsibility is only modest. The point estimates consistently suggest that tax cuts raise government expenditures, but they are only occasionally significantly different from zero, and then only marginally so.

¹⁴ These two graphs are on a different scale than the others in Figure 3 because the dependent variable is now a percent of GDP, not a percent of expenditures.

C. Effects of Long-Run Tax Changes on the Components of Spending

The previous results suggest that tax changes have little effect on overall government spending, or perhaps increase it slightly. But, it is possible that tax cuts have more important effects on the composition of spending. To test for this possibility, we look at the behavior of the major components of federal government spending following long-run tax changes.

Breakdowns of Total Expenditures. We consider two ways of dividing government spending. First, total gross expenditures less interest payments is approximately equal to the sum of three types of expenditures: consumption expenditures, total government investment, and current transfer payments. Total government investment is itself the sum of three pieces: gross government investment, capital transfer payments (that is, payments to state and local governments to fund capital projects), and net purchases of nonproduced assets. This breakdown by type of expenditure excludes only a few trivial components of total gross expenditures less interest payments, such as subsidies and the correction for wage accruals less disbursements.¹⁵ The second breakdown we consider is the division of federal government purchases (that is, consumption expenditures plus gross government investment) into two main sectors: national defense and nondefense purchases.¹⁶

To convert the figures on nominal expenditures by category to real values, we deflate each component by the price index for GDP. We also experiment with deflating by the deflator for the specific component when it exists.¹⁷ This alternative method of deflating has no effect on the qualitative results, and little effect on the point estimates.

Results. We estimate versions of equation (1) using the percentage change in various components of expenditures as the dependent variable and our measure of long-run tax changes as the independent variable. As in the baseline case, we include the contemporaneous value and twenty lags of the tax variable and estimate the regression over the full postwar sample (1950Q1–2006Q4).

Figure 4 shows the estimated cumulative impact of a tax cut of one percent of GDP on the level

¹⁵ Data on all of these categories of expenditures are from NIPA Table 3.2, downloaded 8/8/07.

¹⁶ Data on national defense and nondefense purchases are from NIPA Table 3.9.5, downloaded 8/8/07.

¹⁷ The price indexes for various types of government expenditures are from NIPA Table 3.9.4, downloaded 8/8/07.

of expenditures by type. The only significant effect is a positive impact on government consumption expenditures at long horizons. That is, there is some evidence of fiscal illusion or shared fiscal irresponsibility for this type of spending. The estimated impact on total government investment is also generally large and positive, but the standard errors are very large. Finally, the estimated impact on current transfer payments is mainly negative, suggesting some starve the beast response, but with a substantial positive spike in quarter 7. The standard errors are again large, however.

The estimated impact of long-run tax cuts on the different types of expenditure is relatively insensitive to the inclusion of the early 1950s. When the sample period is restricted to the post-1957 period, the most noticeable change is that the spike in the response of current transfer payments disappears.¹⁸ At other horizons, the response of transfers remains generally negative, but is now very small. Interestingly, the spending related to the Korean War has little impact on the estimated responses of government consumption and investment spending. Even in the post-1957 sample, there is a significant positive impact of a tax cut on consumption expenditures.

Figure 5 shows the results for the sectoral decomposition of government purchases into defense and nondefense purchases. The most striking finding is that the estimated impact of a tax cut on national defense purchases is large and highly statistically significant. A tax cut of one percent of GDP is associated with a rise in defense purchases of 16.5 percent relative to its normal path. The t-statistic for the maximum effect is 5.0. Surprisingly, the estimated positive effect of a tax cut on defense purchases is also very strong in the post-Korea sample. The maximum impact is somewhat smaller for this sample (10.0 percent), but still highly significant ($t = 4.0$).

The fact that tax cuts are correlated with increases in defense spending is open to several interpretations. Most obviously, it could be due to chance: perhaps wars just happened to occur after large tax cuts. The correlation could also reflect a type of omitted variable bias: perhaps policymakers who believe in cutting taxes also believe in a strong military. Finally, a more causal explanation for the

¹⁸ As we discuss further in Section IV, the estimates for the full sample are substantially influenced by a large one-time payment to veterans that occurred seven quarters after the large 1948 tax cut.

correlation could be a somewhat perverse twist on the starve the beast hypothesis. A tax cut and the resulting budget deficit may make policymakers uneasy about spending increases. In this situation, the spending increases that may be easiest to justify are ones for national defense. The case studies discussed in Section IV suggest that there is an element of truth in each of these hypotheses.

The impact of long-run tax changes on nondefense purchases is not statistically different from zero. Figure 5 shows that in the full sample, the estimated impact of a tax cut fluctuates from somewhat positive to somewhat negative, but the effects are far from significant. In the post-Korea sample, the estimated impact is similar, and again not statistically significant. Thus, there is no convincing evidence that tax cuts reduce nondefense spending. That they are associated with increases in one kind of spending, national defense purchases, is consistent with the fact that the point estimates for the effect on total expenditures are positive.

D. The Relationship between Other Types of Tax Changes and Total Expenditures

As discussed above, we focus on the response of government spending to long-run tax changes because this is likely to yield the least biased and fairest test of the starve the beast hypothesis. Nevertheless, it is interesting to look at the behavior of spending following the other types of tax changes we have identified: deficit-driven, countercyclical, and spending-driven actions. This analysis can reveal if the feared biases from using these other types of tax changes to estimate the response of spending appear to be present. It can also provide an indirect check on our classification procedures. For example, if we have classified spending-driven tax changes correctly, they should certainly be positively correlated with spending changes.

For this exercise, we estimate equation (1) using total gross expenditures less interest payments as the dependent variable, and the contemporaneous and twenty lags of a particular type of tax change as the independent variable. We estimate a separate regression for each type of tax change using the full postwar sample period. We again summarize the results by calculating the implied cumulative response of spending to a tax cut (of a given type) of one percent of GDP. These estimated cumulative impacts for

the three additional types of tax actions, along with the one-standard-error bands, are given in Figure 6.¹⁹ To facilitate comparisons, panel (a) of Figure 6 repeats the results for long-run tax actions from Figure 2.

Of the three additional types of tax changes, deficit-driven actions are likely to be the most informative about the starve the beast hypothesis. Like long-run changes, these actions are not taken in response to current or prospective short-run macroeconomic conditions or because spending is moving in the same direction. The reason for excluding these changes from the baseline regression was that deficit-driven tax increases are often part of a budget package that includes spending reductions. In these episodes, both spending and tax decisions are being driven by a third variable, concern about the budget deficit. These observations might therefore bias the results against the starve the beast hypothesis. The cumulative impact of deficit-driven tax changes on total expenditures given in panel (b) of Figure 6 shows this fear is somewhat justified. In the quarter of a deficit driven tax cut and the subsequent two quarters, spending rises substantially. Or, to think about the more realistic case, following a deficit-driven tax increase, spending falls substantially. This is exactly the sort of inverse relationship one would expect if deficit-reduction packages were common. The effect, while large, is not precisely estimated. The t-statistic on the maximum impact is 2.0.

After the first few quarters, the estimated effects of a deficit-driven tax cut turn negative for several years, but return to being positive at distant horizons. None of these estimates are close to statistically significant, however. These results suggest that any spending cuts agreed to at the time of a deficit-driven tax increase disappear within the first year. This reversion of spending to its previous level could reflect a kind of starve (or in this case, feed) the beast response: perhaps spending rises because revenues are rising due to the tax increase. However, the fact that the spending changes do not persist is inconsistent with this interpretation. A more plausible interpretation is that the spending cuts agreed to as part of a budget package are temporary or largely window-dressing, and so get undone quickly. The lack

¹⁹ This way of summarizing the estimates is slightly less intuitive for deficit-driven and spending-driven tax changes than for our baseline case of long-run changes because deficit- and spending-driven tax changes are almost always tax increases. Nevertheless, the interpretation is the same as before: a negative response of spending to a tax cut is supportive of the starve the beast hypothesis; a positive response or no response is not.

of a consistent pattern to the estimates at longer horizons then suggests little ultimate impact of tax changes on expenditure. In this way, the results for deficit-driven tax changes echo those for long-run actions, and provide no support for the starve the beast hypothesis.

Panel (c) of Figure 6 shows the implied impact on spending of a countercyclical tax cut. We excluded such tax changes from our baseline regression because the state of the economy could tend to influence spending and taxes in opposite directions, and so bias the estimates against the starve the beast hypothesis. The results suggest that this is somewhat the case. A countercyclical tax cut is associated with a persistent rise in spending. However, the standard errors are quite large, so it is impossible to reject the hypothesis of no relationship.

Panel (d) shows the behavior of government spending following a spending-driven tax cut. In this case, the effects are negative, large in absolute terms, and highly statistically significant.²⁰ This is exactly the result one would expect: if we classified spending-driven tax changes correctly, there should be a positive correlation between such actions and spending. That the relationship persists is consistent with the spending changes associated with these spending-driven actions being permanent. These findings for spending-driven tax changes both confirm our classification of these tax actions and illustrate the importance of controlling for motivation when testing the starve the beast hypothesis. Including spending-driven actions would clearly bias the results toward finding a positive correlation between spending changes and tax changes.

One way to see how much bias including these spending-driven actions would cause is to define a tax variable that sums all four types of legislated tax changes and then uses this as the explanatory variable in equation (1).²¹ Panel (e) of Figure 6 shows the implied impact of a legislated tax cut of one percent of GDP of any motivation on total expenditures. The estimated response is strongly negative and

²⁰ These findings are somewhat sensitive to the sample period. Some of the largest spending-driven tax changes were during the Korean War. When the post-1957 sample period is used, the maximum impact of a spending-driven tax cut of one percent of GDP is large (−6.8 percent), but only marginally significant ($t = -1.6$).

²¹ As discussed in Romer and Romer (2007a), we identify one legislated tax cut that was designed to accentuate an existing economic boom. For completeness, we include this one procyclical tax action in the sum of all legislated tax changes.

statistically significant for the first three years after a tax cut. The point estimate for the maximum cumulative effect is -3.8 percent ($t = -2.4$). Since none of the other types of tax changes show a consistent negative response, this implied negative effect of the aggregate tax variable reflects the influence of the spending-driven tax changes.

To test this proposition more directly, we define a second composite tax variable that includes all legislated tax changes other than those motivated by spending changes. Panel (f) of Figure 6 shows the implied cumulative response of total expenditures to a non-spending-driven legislated tax change of one percent of GDP. The effects are consistently positive, suggesting that, if anything, tax cuts appear to increase government spending, not decrease it as the starve the beast hypothesis predicts. And, for horizons beyond three years, these positive effects are significantly different from zero.

The results suggest that the inclusion of spending-driven tax changes may explain why much of the previous literature has found evidence for the starve the beast hypothesis. A typical test of the starve the beast hypothesis uses the change in cyclically adjusted revenues, which includes all changes in revenues not related to short-run fluctuations in income, as the measure of tax changes. Data on the change in cyclically adjusted revenues are currently available for the period 1947Q2–2006Q2. We therefore investigate the effects of using the contemporaneous value and eleven lags of this variable as the tax measure for the period 1950Q1–2006Q2.²² When we use this conventional tax variable, the results are indeed supportive of the starve the beast hypothesis. Panel (g) of Figure 6 shows that the cumulative effect of a decline in real cyclically adjusted revenues of one percent of GDP starts out positive, but then turns strongly negative. The maximum impact is a change in government expenditures of -3.1 percent ($t = -2.1$).

Given that we have identified spending-driven tax changes, it is possible to subtract these actions from the change in cyclically adjusted revenues.²³ The results using this series, which are shown in panel

²² For comparability with our tax measure, we use the change in real cyclically adjusted revenues as a percent of real GDP. See Romer and Romer (2007b) for a more detailed discussion of the sources and derivation of this measure.

²³ Since both series are expressed as a percent of GDP, the spending-driven tax changes can be subtracted without further adjustment.

(h), are dramatically different from those using all cyclically adjusted revenues. The impact of a decline in cyclically adjusted revenues less spending driven changes of one percent of GDP is now strongly positive in the short run: the maximum impact is 3.6 percent ($t = 4.7$). It then gradually declines toward zero, but never turns negative over the eleven-quarter horizon we consider. Thus, the results provide no support for the starve the beast hypothesis, and, indeed, are somewhat supportive of the notion of shared fiscal irresponsibility. This suggests that the inclusion of spending-driven changes in conventional revenue measures is an important source of the finding that government spending moves in the same direction as tax revenues.²⁴

III. EFFECTS OF LONG-RUN TAX CHANGES ON FUTURE TAXES

Our analysis finds no evidence that tax cuts lead to reductions in government spending. This finding obviously raises another question: If tax cuts do not reduce government spending, how does the government budget eventually balance? An obvious possibility is that the adjustment occurs on the tax side rather than on the expenditure side. To explore this possibility, we examine the response of both tax revenues and tax legislation to long-run tax changes.²⁵

A. Response of Tax Revenues

To investigate how revenues respond to long-run tax changes, we estimate equation (1) using a measure of real tax revenues as the dependent variable. That is, we regress the percentage change in real

²⁴ The importance of spending-driven tax changes in causing bias is sensitive to the sample period used. Spending-driven changes were largest during the Korean War and tend to cause substantial bias in samples that include this period. In later sample periods, spending-driven changes are smaller and so are a less important source of bias. This may explain why studies such as Ram (1988), Miller and Russek (1990), and Bohn (1991), which use data from the Korean War period and before, find support for the starve the beast hypothesis, while those such as von Furstenberg, Green, and Jeong (1986), which use data starting in 1954, do not.

²⁵ Bohn (1991) also examines the degree to which deficits caused by tax cuts are eliminated by subsequent tax increases.

revenues on a constant and the contemporaneous value and twenty lags of our measure of long-run tax actions. The particular revenue variable we use is the federal total receipts series from the National Income and Product Accounts deflated by the price index for GDP.²⁶ We estimate the revenue response over both the full postwar sample period (1950Q1–2006Q4) and the post-Korean War sample (1957Q1–2006Q4).

Figure 7 shows the implied cumulative response of total receipts to a long-run tax cut of one percent of GDP for both sample periods. Tax receipts decline strongly in the short run in response to a tax cut. The contemporaneous effect is a change in receipts of –1.8 percent in the full sample ($t = -1.9$) and –2.0 percent in the post-Korea sample ($t = -2.2$). Total receipts remain substantially below their pre-tax-cut path for the next year and a half.

In both samples, receipts then recover substantially. For the full sample, the rise in revenues two years after the tax cut is dramatic and marginally significant. This finding is largely driven by the Korean War. As described in Section IV, the large 1948 tax cut was followed roughly two years later by the outbreak of the war. Two major tax increases were passed early in the war, and the war was accompanied by rapid output growth. For this reason, the results for the full sample almost surely overstate the true tendency of revenue to rebound. For the post-Korea sample, receipts rise above their pre-tax-cut path seven quarters after the tax cut, but the effect is modest and the standard errors are large (the t -statistics on the positive effects do not rise above 0.9). Thus, the revenue effects of a tax cut are indistinguishable from zero after roughly two years.

B. Response of Tax Legislation

To understand the behavior of revenues following a long-run tax cut, it is important to investigate the behavior of subsequent tax legislation. Do tax revenues recover because of unusually rapid growth or because policymakers legislate tax increases? Given that we have constructed measures of legislated tax changes classified by motivation, this is an issue we can investigate.

²⁶ Data on total federal receipts are from NIPA Table 3.2, downloaded 8/8/07.

In all of our previous analyses of spending and revenues, we consider the experiment of a long-run tax cut that is not followed by any additional tax changes based on long-run considerations. Therefore, it does not make sense to ask how long-run tax changes respond to our experiment. But, it is reasonable to ask how other types of legislated tax changes respond to a long-run tax cut. Long-run tax cuts that do not lower spending, and so increase the deficit, might lead to tax increases designed to reduce an inherited budget deficit. Likewise, a long-run tax cut that gives rise to a short-run boom could lead to a countercyclical tax increase. A long-run tax cut could also lead policymakers to switch to a “pay-as-we-go” policy: a budget deficit resulting from a long-run tax cut may make policymakers unwilling to increase spending without increasing taxes. Therefore, one could also see an increase in spending-driven tax increases following long-run tax cuts.

Our empirical framework is again identical to that in equation (1), except that the dependent variable is now a measure of legislated tax changes. That is, we regress legislated tax changes of some motivation on a constant and the contemporaneous value and several lags of our measure of long-run tax changes. In our baseline specification we again use twenty lags. However, we experiment with longer lags. We estimate the responses over both the full postwar sample and the post-Korean War sample.

As before, we summarize the results by examining the cumulative impact of a long-run tax cut of one percent of GDP. A positive impact implies that subsequent tax actions counteract the long-run tax cut. Because the other tax variables are also expressed as a percent of nominal GDP, the dependent and independent variables are on the same scale. Therefore, the impact can be interpreted as the fraction of the long-run tax cut that is undone over the horizon considered.

The estimated impacts of a long-run tax cut of one percent of GDP on tax changes of various types are shown in Figure 8. Panel (a) shows the impact on deficit-driven actions. The effect is positive and highly statistically significant, suggesting that long-run tax cuts tend to be followed by deficit-driven tax increases. The cumulative impact is 0.22 ($t = 3.0$) after eight quarters, and 0.23 ($t = 2.2$) after

sixteen.²⁷ This suggests that about a fifth of a long-run tax cut is undone by deficit-driven tax increases within a few years. These results are highly robust. Starting the sample in 1957 has virtually no impact. And, increasing the number of lags to forty and carrying out the simulations for ten years strengthens the results. Ten years after the long-run tax action, 41 percent of the action has been undone by deficit-driven tax increases ($t = 2.2$).

Panel (b) shows the impact of a long-run tax cut on countercyclical tax actions. The estimated impact is moderate, but not close to significantly different from zero. After twenty quarters, countercyclical tax actions have counteracted 17 percent of a long-run tax cut ($t = 0.5$). Starting the sample in 1957 has no impact because countercyclical actions were unheard of in the early 1950s. Including longer lag lengths suggests that the response diminishes at longer horizons. The estimated effect after ten years is 0.09 ($t = 0.2$).²⁸

Panel (c) shows the impact of a long-run tax cut on spending-driven tax changes. In this case, the effects are virtually zero for the first nine quarters after the long-run tax cut, and then turn strongly positive. The maximum cumulative impact is 0.50 ($t = 2.6$) after fourteen quarters. The impact after twenty quarters is 0.37 ($t = 1.6$). This suggests that spending-driven tax increases occur after a long-run tax cut, and that they counteract close to half of the initial cut. Thus, long-run tax cuts may indeed give rise to more pay-as-we-go policies.

More so that with the other tax changes, there is reason to be concerned that the results for spending-driven actions are being influenced by the observations from the Korean War. Starting the sample in 1957 does indeed weaken the link substantially, but does not eliminate it. The maximum

²⁷ There is a substantial contemporaneous impact (0.10 with a t-statistic of 3.3). The most important observation behind this estimate is 1983Q1. A large part of the tax cuts in the Economic Recovery Tax Act of 1981 were scheduled to go into effect in 1983Q1. Concern about current and prospective deficits, however, led to passage of the Tax Equity and Fiscal Responsibility Act of 1982, which raised revenues mainly by modifying some features of the 1981 act that had already taken effect (Romer and Romer, 2007a). Thus, although the long-run tax cut and the deficit-driven tax increase occurred simultaneously, there is a clear sense in which the deficit-driven increase was a response to the long-run cut.

²⁸ We also experimented with leaving out the 1975 tax cut, which is a huge outlier among countercyclical actions. It is peculiar because it was a rebate that mainly cut taxes dramatically in one quarter and then raised them dramatically in the next. Zeroing out this action reduces the response at medium horizons, but has almost no effect on the longer-run response. The main effect of the change is to cut the standard errors more than in half.

impact of a long-run tax cut is a rise in spending-driven taxes of 0.15 ($t = 2.1$). Likewise, including twenty additional lags reduces the impact substantially for the full sample, but this effect is due entirely to the required shortening of the sample period.

Panel (d) shows the effect of a long-run tax cut on the other types of legislated tax changes combined.²⁹ The effect is positive, large, and significant. The effect is 0.58 ($t = 2.0$) after twelve quarters, 0.81 ($t = 2.3$) after sixteen, and 0.76 ($t = 2.0$) after twenty. This suggests that most of a long-run tax cut is typically undone by legislated tax increases of various sorts within five years.

Figure 9 shows two robustness checks for the effect of a long-run tax cut on this composite of other tax changes. Panel (a) shows the impact of starting the sample in 1957. Both the maximum impact and the statistical significance are somewhat reduced by this change. The impact now peaks at 0.58 ($t = 1.6$) after nineteen quarters. Panel (b) shows the effect of including forty lags of long-run tax changes. The required shortening of the sample reduces the estimated responses over the first twenty quarters slightly. Thereafter, it moves irregularly upward. The response after 40 quarters is large (0.77), but not precisely estimated ($t = 1.3$). Though they weaken the evidence slightly, these two robustness checks confirm that a large fraction of a long-run tax cut is typically reversed by legislated tax increases within the next few years.

C. Discussion

This analysis of the response of revenues and tax legislation to a tax cut yields two main findings. First, although a tax cut leads to a sharp fall in revenues in the short run, it does not have any clear impact on revenues at horizons beyond about two years. Second, most of the tax cut is offset by legislated tax increases over the next several years. Taken together, these findings suggest that a substantial fraction of the rebound in revenues is the result of non-legislated changes.

The key source of the non-legislated changes in revenues is almost certainly the effect of the tax cut on economic activity. In Romer and Romer (2007b), we find that a tax cut of one percent of GDP

²⁹ We again include the one procyclical tax action in the combined measure. See n. 21.

increases real output by approximately three percent over the next three years. Since revenues are a function of income, this growth undoubtedly raises revenues.

There is, however, an important caveat to this finding that tax cuts partially pay for themselves through more rapid growth: some of the output response is almost surely a transitory departure of output from normal, not a permanent change in the economy's normal level of output. The idea that a tax cut has a large, rapid impact on the flexible-price level of output is not particularly plausible. And as we describe in Romer and Romer (2007b), the behavior of inflation and unemployment following tax cuts is consistent with the view that the output effects are largely temporary. To the extent that this is the case, some of the rebound in revenues is also temporary. As a result, in the absence of further legislated changes, there may be some long-run budgetary shortfall in the wake of the tax cut.

Because of these complications, our results do not allow us to describe with complete confidence how the government achieves long-run budget balance following a tax cut. But, we can say that we find no evidence of adjustment on the spending side, and considerable evidence of substantial adjustment on the tax side.

IV. SPENDING AND TAXES IN FOUR KEY EPISODES

In this section, we examine the four episodes in our sample that stand out as having by far the largest long-run tax cuts. This examination serves several purposes. First, it allows us to study the evidence from the narrative record concerning the effect of the cuts on how policymakers made decisions about spending. We can see whether policymakers for the most part simply ignored the tax cuts; were cognizant of the reduced revenues but believed that other considerations were much more important in making spending decisions; or believed that the cuts were in fact causing them to reduce spending. To

keep the narrative analysis manageable, we focus primarily on presidential documents and statements.³⁰ However, in cases where Congressional views appear to be central or at odds with those of the executive branch, we also examine Congressional documents.

Second, we investigate whether the relationships shown by the regressions appear in the key episodes. We examine both the behavior of overall spending and the paths of its major categories. This allows us to check whether the regression results reflect consistent patterns in the data.

Third, we ask whether there were unusual developments in the episodes that could cause the spending regressions to give a misleading picture. For example, if we find that there were often idiosyncratic developments pushing spending up in the episodes, this would suggest that the regressions were overstating the evidence against the starve the beast hypothesis. Another type of development that could cause the regressions to give a misleading picture involves actions that substantially affect spending only with long lags. For example, if we find that there were often important actions in the episodes that raised spending with considerable delays, this would suggest that tax cuts stimulate spending even more than suggested by the regressions.

Finally, we examine the tax side of the episodes. We look at what tax actions were taken following the tax cuts, and thus again check whether the regression results reflect consistent patterns. Perhaps more importantly, we examine the reasons policymakers gave for those actions to see to what extent they appear to have been responses to the cuts. As with spending, we also check whether idiosyncratic factors were an important determinant of tax changes in the episodes.

A. The Revenue Act of 1948

The Revenue Act of 1948 was passed over Harry Truman's veto in April 1948. The bill reduced revenues by 1.9 percent of GDP beginning in 1948Q2. The primary motivation for the cut was a desire to

³⁰ The key presidential documents that we use are the Budget of the United States Government (abbreviated as Budget in citations) and the Economic Report of the President (abbreviated as Economic Report). Presidential speeches are identified by their title and date as given in John Woolley and Gerhard Peters, The American Presidency Project (www.presidency.ucsb.edu). Page numbers are from our printouts of the speeches, and so may be affected by our choices of font size and margins.

improve economic efficiency by reducing marginal tax rates.³¹

The tax cut was followed by a substantial reduction in revenues. It is clear, however, that the cut had little impact on Truman's view of appropriate spending. In his budget message in January 1949, he said of his budget:

Substantial direct assistance is provided for other members of the family of nations, and expenditures in support of our armed forces are materially increased. Funds are included for the necessary strengthening of our economy through the development and conservation of the Nation's productive resources. Increased emphasis is placed on the provision of badly needed measures to promote the education, health, and security of our people (1950 Budget, p. M5).

His main response to the tax cut was to propose a counteracting tax increase. He argued, "In a period of high prosperity it is not sound public policy for the Government to operate at a deficit. ... I am, therefore, recommending new tax legislation to raise revenues by 4 billion dollars" (p. M5). This increase would have offset 80 percent of the 1948 cut.

The 1950 Economic Report provided an even clearer statement of Truman's view that government spending should be determined by considerations other than the level of revenues, and that tax policy should be adjusted accordingly:

In fields such as resource development, education, health, and social security, Government programs are essential elements of our economic strength. If we cut these programs below the requirements of an expanding economy, we should be weakening some of the most important factors which promote that expansion. Furthermore, we must maintain our programs for national security and international peace. ...

Government revenue policy should take into account both the needs of sound Government finance and the needs of an expanding economy (p. 8).

Nonetheless, the fall in revenues appears to have had a marginal effect on Truman's spending policies. In his budget message in January 1949, he stressed "the compelling need for financial prudence by the Government at this time," and added that "it has been necessary to deny many requests for additional funds which would normally be desirable" (1950 Budget, p. M6). In the 1949 Midyear Economic Report of the President, he explained, "When I submitted my budget for the fiscal year 1950 last January, the programs of expenditure that I then recommended were held to a minimum consistent

³¹ Our descriptions in this section of the motivations for tax changes and our figures for their revenue effects are based on Romer and Romer (2007a). The revenue estimates exclude the effects of retroactive features of the bills.

with our basic needs in view of the inflationary strain upon materials and manpower then prevailing” (p. 7). Since Truman viewed the budget deficit as contributing to inflationary pressures (for example, Annual Message to the Congress on the State of the Union, 1/5/49, p. 3), this points to an effect of the tax cut on spending decisions. It is clear, however, that other considerations were more important.

On June 25, 1950, North Korea invaded South Korea. Once that occurred, taxes and the deficit essentially disappeared from Truman’s discussions of spending. Even more so than in peacetime, his view was that spending should be determined by the country’s needs, and taxes adjusted accordingly. For example, in his budget message in January 1951, Truman described the spending side of the budget and then stated, “I shall shortly recommend an increase in tax revenues in the conviction that we must attain a balanced budget to provide a sound financial basis for what may be an extended period of very high defense expenditures” (1952 Budget, p. M6).

Finally, although Congress’s view of the tax cut was obviously very different from Truman’s, Congress does not appear to have desired lower spending than the president. For example, in August 1948, Truman reported that although Congress had not appropriated the full amount he had requested for fiscal 1948 and 1949, this shortfall was offset by two factors: some spending had been authorized but not yet appropriated, and several pieces of legislation had been enacted that would require higher spending, but no spending had yet been authorized. The bottom line was that he expected spending in fiscal 1949 to be significantly higher than what he had requested in January (Statement by the President: The Midyear Review of the Budget, 8/15/48, p. 3). Thus, there is no evidence of a starve the beast effect operating through Congressional actions.

Panel (a) of Figure 10 shows the behavior of government spending in this episode. It plots not only our measure of total expenditures, but also national defense purchases and the two major categories of nondefense spending, nondefense purchases and current transfer payments. The vertical line shows the quarter in which the tax cut took effect. Several things are apparent. First, and most important, there was no discernable slowdown in any major type of spending. Indeed, the growth of overall spending increased after the tax cut. Total expenditures, which had been essentially flat before the tax cut, rose by

16 percent in the two years between the cut and the start of the war. Second, there was a substantial one-time spike in transfer payments in 1950Q1. This rise reflected one-time dividend payments from the National Service Life Insurance trust fund. These payments were the result of a large accumulation of assets in the trust fund, which could not be used for other purposes (Hines, 1943; Survey of Current Business, March 1950, pp. 1-3, and August 1950, p. 7). Third, there was a sharp rise both in defense spending and in overall spending after the outbreak of the war.

Both the National Service Life Insurance dividend payments and the increased military spending after the start of the war clearly reflected unusual developments, not just the normal response of spending to tax cuts. The occurrence of these idiosyncratic upward influences on spending works in the direction of causing the regressions to overstate the impact of tax cuts on subsequent spending increases. The Korean War also acts in the direction of making the regressions overestimate the impact of tax cuts specifically on defense spending.

There was also an important unusual development operating in the opposite direction. The Social Security Amendments of 1950 almost doubled Social Security benefits starting in September 1950 and substantially increased the coverage of the system beginning in January 1951 (Social Security Bulletin, October 1950, pp. 3-14). Because Social Security spending was initially small, these changes had little immediate impact on overall spending. Nonetheless, the expansion of benefits and coverage contributed significantly to the growth of spending over time. The fact that these delayed spending effects are not captured in our regressions operates in the direction of causing the regressions to understate the impact of tax cuts on later spending increases.

On the tax side, the 1948 tax cut was followed by a series of tax increases. The first, and least important, was an increase in Social Security taxes of 0.3 percent of GDP in 1950Q1 that had been legislated before the tax cut was passed. After that, there were larger tax actions. The Social Security Amendments of 1950 increased the base of the payroll tax from \$3000 to \$3600 effective at the beginning of 1951 and called for a gradual increase in the combined Social Security tax rate from 3 percent to 6½ percent over the next two decades (Social Security Bulletin, October 1950, pp. 3-14). And three bills in

1950 and 1951 to finance the Korean War increased taxes by a combined 4.2 percent of GDP.

Both the Social Security and Korean War tax increases were far from inevitable consequences of the spending increases they were associated with. In the case of Social Security, policymakers were grappling with the issue of how to finance the system. A special Congressional commission and the Social Security Administration both recommended that Social Security taxes be limited and that the system move toward increasing reliance on general revenues. Instead, however, the 1950 amendments repealed the provision of the Social Security Act that permitted financing from general revenues and made the system entirely self-financing (Social Security Bulletin, May 1948, pp. 21-28; February 1949, pp. 3-9; October 1950, pp. 3-14).

In the case of the Korean War, the extent of the government's reliance on contemporaneous tax increases is remarkable: total expenditures as a share of GDP rose by 6.0 percentage points from 1950Q2 to its peak in 1952Q3, only moderately more than the expected revenue effects of the tax increases to finance the war. Moreover, Truman explicitly cited the deficit as a reason for this heavy reliance on tax finance. Soon after the start of the war, he wrote to Congressional leaders:

We embark on these enlarged expenditures at a time when the Federal budget is already out of balance. This makes it imperative that we increase tax revenues promptly lest a growing deficit create new inflationary forces detrimental to our defense effort.

We must make every effort to finance the greatest possible amount of needed expenditures by taxation (Letter to the Chairman, Senate Committee on Finance, on the Need for an Increase in Taxes, 7/25/50, p. 1).

Thus, the Korean War tax increases were in part a response to the 1948 tax cut.

The only important unusual tax developments in this episode were ones closely tied to the unusual spending developments. First, although the Korean War tax increases were partly a response to the tax cut, they were also a response to the war. Second, the 1950 Social Security amendments scheduled tax increases long after the three-year and five-year windows considered in our regressions. The first development acts toward making the regressions overstate the relationship between long-run tax cuts and later tax increases, while the second acts in the opposite direction.

B. The Revenue Act of 1964

Lyndon Johnson signed the Revenue Act of 1964 in February 1964. It reduced revenues by 1.3 percent of GDP in 1964Q2 and by another 0.6 percent in 1965Q1. The key motivation for the tax cut was a desire to increase long-run growth.

Because of very rapid growth, revenues recovered quickly after the tax cut. As a result, there were no immediate budget deficits to trigger a starve the beast type of response. Nevertheless, policymakers' statements and behavior can provide some evidence concerning this mechanism.

At almost the same time that he signed the tax bill, Johnson began to propose drastic increases in spending. In February 1964 he gave a speech proposing federal hospital insurance for the elderly and other health initiatives (Special Message to the Congress on the Nation's Health, 2/10/64). In May 1964, he gave his speech on "the Great Society" that called for elimination of poverty, urban renewal, pollution reduction, and education expansion (Remarks at the University of Michigan, 5/22/64). Over the next year, a number of spending increases directed at achieving these goals were passed. The most significant was the dramatic expansion of benefits and the introduction of Medicare contained in the Social Security Amendments of 1965.

The administration believed that spending should be determined by necessity and efficiency. The president's budget message in January 1965 stated: "The expenditures proposed in this budget reflect a careful balancing of national goals against budgetary costs. The budget I now present will, in my judgment, carry out the responsibilities of the Federal Government efficiently and wisely. It was constructed on that basis alone" (1966 Budget, p. 10). The narrative record in this episode is striking in the degree to which revenues are not mentioned as a determinant of expenditures.

Defense spending increased substantially starting in mid-1965 because of escalation of the war in Vietnam. Johnson argued forcefully against allowing budgetary concerns to stop the rise in nondefense spending. He stated:

There are men who cry out: We must sacrifice. Well, let us rather ask them: Who will they sacrifice? Are they going to sacrifice the children who seek the learning, or the sick who need medical care, or the families who dwell in squalor now brightened

by the hope of home? ...

I believe that we can continue the Great Society while we fight in Vietnam
(Annual Message to the Congress on the State of the Union, 1/12/66, p. 2).

Congress went along with his calls for increased spending. For example, the Social Security Amendments of 1967 brought about another substantial increase in benefits and a significant increase in coverage. Thus, the rise in spending following the tax cut was in no way just the consequence of the war. Also, this lack of response to a positive spending shock is indirect evidence that policymakers were unlikely to respond to a decline in revenues by cutting spending.

The leveling off of economic activity in late 1966 slowed revenue growth at the same time that spending was rising rapidly. As a result, the deficit increased substantially. Nevertheless, the administration did not call for substantial spending reductions. Federal expenditures were expected to rise by \$15 billion in 1968 (1968 Economic Report, p. 54). Instead, the administration concluded: “the cost of meeting our most pressing defense and civilian requirements cannot be responsibly financed without a temporary tax increase” (1969 Budget, p. 8).

Over the president’s objection, Congress included a \$6 billion spending reduction (relative to projected) in the bill imposing a ten-percent temporary tax surcharge. Congress pressed for the spending cuts not because revenues had declined, but because they felt it was unfair to take all of the needed macroeconomic restraint in the form of higher taxes. A number of senators expressed sentiments similar to that of Senator Byrd, who stated: “Before any new tax burden ... is placed upon the American taxpayer, the executive branch and the legislative branch should reduce, and eliminate where possible, all nonessential expenditures” (Congressional Record, 90th Congress, 2^d Session, Volume 114—Part 7, 4/2/68, p. 8561). Nevertheless, this spending cut is the one development in this episode that provides some support for the starve the beast hypothesis.

The actual behavior of spending following the 1964 tax cut is completely consistent with policymakers’ stated positions. Panel (b) of Figure 10 shows that total expenditures were basically constant during the first year after the tax cut, but then rose dramatically. Total expenditures increased by 28 percent between 1965Q1 and 1969Q1. The rise in defense purchases was clearly a large source of the

increase. However, current transfer payments increased even more rapidly. Nondefense purchases also increased slightly in this period.

Special factors clearly played a role in the behavior of spending. Much of the rise in defense expenditures was related to the Vietnam War. To the extent that defense spending truly was non-discretionary, some of the rise in spending reflects this exogenous shock rather than a failure of the starve the beast phenomenon. At the same time, the immediate increase in spending called for by the Social Security Amendments of 1965 and 1967 in a fundamental way underestimates the true rise in spending. The creation of the Medicare program and the increases in Social Security benefits and coverage put in place an enormous stream of future spending. So, the present value increase in spending that was passed in the wake of the 1964 tax cut was unquestionably huge.

Policymakers' statements and actions on taxes in this episode are striking. In 1965, the Johnson administration proposed (and succeeded in passing) two significant tax actions. One was the Excise Tax Reduction Act of 1965. The administration viewed this tax cut as a continuation of the 1964 action. In this case, the serial correlation of tax changes reflected serial correlation in views about appropriate policy. The second was the Social Security Amendments of 1965, which included a substantial increase in payroll taxes to pay for an equally substantial increase in benefits, including hospital insurance for the elderly. This tax increase appears to have had little to do with the 1964 tax cut. Policymakers paid for the desired expansion of benefits through expansion of taxes because the decision had been made in the early 1950s that the Social Security system should be self-financing.³²

Beginning in early 1966, policymakers began to worry that the economy was overheating. For example, the 1966 Economic Report stated: "vigorous private demand and required defense spending could upset the balance of supply and demand" (p. 10). While policymakers frequently discussed the need to restrain expenditures, expenditures in fact rose dramatically. Understanding this, policymakers pushed for tax increases instead. The Tax Adjustment Act of 1966 (enacted in March 1966) rescinded the

³² The Social Security Amendments of 1967, enacted in January 1968, also raised taxes substantially to pay for another increase in benefits and coverage.

excise tax reduction that had just occurred the previous January. Public Law 89-800 (enacted in November 1966) suspended the investment tax credit. Together, the two tax increases were expected to raise revenues by 0.3 percent of GDP.³³

This emphasis on tax changes rather than spending changes to deal with overheating was consistent with the administration's fiscal philosophy. The 1967 Economic Report stated: "most economists now agree that the selection of appropriate expenditure levels ... should be made in light of the relative merits of alternative programs, and of the benefits of added public expenditures, compared with private ones, at the margin. ... [I]t is preferable to emphasize changes in tax rates (suitably coordinated with changes in monetary policy) for stabilization purposes" (p. 68).

By far the largest tax increase in the immediate post-1964 period was the 1968 surcharge. The administration first proposed a six-percent surcharge in January 1967 (see, for example, Annual Message to the Congress on the State of the Union, 1/10/67, p. 7). In August 1967, Johnson stated: "If left untended, this deficit could cause ... [a] spiral of ruinous inflation" and "[b]rutally higher interest rates" (Special Message to the Congress: The State of the Budget and the Economy, 8/3/67, p. 1). He requested that the surcharge be increased to ten percent, which was the level ultimately included in the Revenue and Expenditure Control Act of 1968. The act increased taxes by 0.9 percent of GDP in 1968Q3 and by another 0.2 percent in 1969Q1.

Johnson was quite explicit that the surcharge was undoing part of the 1964 tax cut. In the Statement by the President Upon Signing the Tax Bill, he stated: "This temporary surcharge will return to the Treasury about half the tax cuts I signed into law in 1964 and 1965" (6/28/68, p.1). This action, combined with the continued rise in expenditures, is a vivid example that what typically gives in response to a tax cut is not spending but the tax cut itself.

³³ Public Law 90-26 (enacted in June 1967) restored the investment tax credit. As discussed in Romer and Romer (2007a), the motivation for this change involved the conditions in a particular sector (the capital goods market), and concern about longer-run incentives for investment. It does not appear to have been motivated by the 1964 tax cut or short-run macroeconomic conditions.

C. The Economic Recovery Tax Act of 1981

A very large long-run tax cut was enacted in August 1981, shortly after Ronald Reagan took office. The cut lowered taxes by a combined 4.5 percent of GDP in a series of steps.

Reagan was a strong advocate of spending reductions throughout his presidency. For example, in a speech presenting his economic program, he identified “reducing the growth in government spending and taxing” as a central goal, and he argued that “[s]pending by government must be limited to those functions which are the proper province of government” (Address Before a Joint Session of the Congress on the Program for Economic Recovery, 2/18/81, pp. 1, 5). Similarly, in his first budget message, in February 1982, he listed “[r]educing the growth of overall Federal spending by eliminating Federal activities that overstep the proper sphere of Federal Government responsibilities” as one of his fundamental economic goals (1983 Budget, p. M4).

The tax cut was followed by a substantial fall in revenues and a large rise in the budget deficit. As the deficit increased, Reagan often cited it as a further reason for restraining spending. For example, in his February 1986 budget message, he said, “there is a major threat looming on the horizon: the Federal deficit” (1987 Budget, p. M-4). He went on to say: “Spending is the problem – not taxes – and spending must be cut. The program of spending cuts and other reforms contained in my budget will lead to a balanced budget at the end of five years” (p. M-5). Similarly, his February 1988 budget message stated:

Last year, members of my Administration worked with the Leaders of Congress to develop a 2-year plan of deficit reduction – the Bipartisan Budget Agreement. ...

The Bipartisan Budget Agreement reflects give and take on all sides. I agreed to some \$29 billion in additional revenues and \$13 billion less than I had requested in defense funding over 2 years. However, because of a willingness of all sides to compromise, an agreement was reached that pared \$30 billion from the deficit projected for 1988 and \$46 billion from that projected for 1989 (1989 Budget, p. 1-6).

Thus, the narrative record from this episode provides some support for the starve the beast hypothesis.

Panel (c) of Figure 10 plots the behavior of the major categories of spending around the 1981 tax cut. The vertical line is drawn in 1981Q3, the date of the first of the series of cuts. Despite what is suggested by the narrative evidence, the growth of overall spending did not fall. Over the five years

following the tax cut, the growth rate of total expenditures fluctuated around an average of 4.6 percent per year, somewhat above the average growth rate of 3.6 percent in the two years before the cuts. The growth of current transfer payments, however, slowed slightly beginning a few years after the tax cuts were enacted; indeed, transfer payments were essentially unchanged for three years beginning in 1982Q4. This fits with the narrative evidence. Finally, as in the other episodes, defense spending rose sharply.

There were two important unusual spending developments in this episode. First, the tax cuts coincided with a shift in political power toward supporters of lower spending. Reagan's goal of restraining government spending was not shared by his predecessor. For example, in his final budget message, Jimmy Carter, while advocating "budget restraint," stated, "The growth of budget outlays is puzzling to many Americans, but it arises from valid social and national security concerns" (1982 Budget, pp. M4-M5). There was also a large swing in the balance of political power in Congress toward advocates of spending restraint at the time of Reagan's election. Thus, there was clearly an omitted variable acting to reduce spending.

Second, some of the shift in spending toward defense was clearly a result of other developments, not the tax cuts. In particular, Reagan had campaigned on a need to rebuild the military and identified "strengthening the Nation's defenses" as one of his key goals (1983 Budget, p. M4). This ideological shift surely explains some of the rise in defense spending, and so operates in the direction of making the regressions overstate the extent to which tax cuts cause a shift in spending toward defense.

The tax cuts were followed by two types of tax increases. First, the Social Security Amendments of 1983 called for a series of tax increases from 1984 to 1990 to improve the solvency of the Social Security system. Second, there was a series of tax increases that were explicitly motivated by a desire to reduce the budget deficits that developed followed the tax cuts. These included the Tax Equity and Fiscal Responsibility Act of 1982, which undid some of the provisions of the 1981 act; the Deficit Reduction Act of 1984; the Omnibus Budget Reconciliation Act of 1987; and the Omnibus Budget Reconciliation Act of 1990. For example, in a national address on the 1982 act, Reagan stated that it reflected a choice to "reduce deficits and interest rates by raising revenue from those who are not now paying their fair

share,” rather than to “accept bigger budget deficits, higher interest rates, and higher unemployment” (Address to the Nation on Federal Tax and Budget Reconciliation Legislation, 8/16/82, p. 4). Similarly, the 1989 Budget reported that the 1987 act was enacted “[i]n conformance with the Bipartisan Budget Agreement” (p. 4-5), which, as described above, was motivated by concern about the deficit. The 1982 and 1984 actions alone increased taxes by 1.0 percent of GDP.³⁴

D. The Economic Growth and Tax Relief Reconciliation Act of 2001 and the Jobs and Growth Tax Relief Reconciliation Act of 2003

Two long-run tax cuts were passed early in the administration of George W. Bush. The Economic Growth and Tax Relief Reconciliation Act of 2001, enacted in June 2001, included a long-run tax cut of 0.8 percent of GDP in 2002Q1. The Jobs and Growth Tax Relief Reconciliation Act of 2003, enacted in May 2003, included a long-run cut of 1.1 percent of GDP in 2003Q3.

These tax cuts do not appear to have had any substantial impact on the administration’s view of appropriate spending. Throughout the episode, both spending restraint and either preserving the surplus or reducing the deficit received some attention. But, the discussions of spending did not change appreciably in response either to the tax cuts or to the subsequent deterioration of the budget situation.

The administration’s first budget proposals, which predated the tax cuts, put some emphasis on spending restraint and on paying down debt. The president’s first budget document, for example, stated that the budget would “Moderate Growth in Government and Fund National Priorities” and achieve “Debt Reduction” (A Blueprint for New Beginnings: A Responsible Budget for America’s Priorities, 2/28/01, p. 7).³⁵ It also said that “[t]he President’s Budget commits to using today’s surpluses to reduce the Federal Government’s publicly held debt so that future generations are not shackled with the

³⁴ There was one special factor in this episode that works in the direction of making the regressions overstate the extent to which long-run tax cuts are offset by later tax increases: some of the tax increases that followed the tax cuts were the result of the Social Security Amendments of 1977, and thus were not responses to the 1981 act. These particular tax increases, however, were much smaller than those resulting from legislation enacted after 1981.

³⁵ This document was not part of the president’s formal 2002 budget, which was not submitted until April 2001. However, it is included with the other 2002 budget documents on the Government Printing Office website. See <http://www.gpoaccess.gov/usbudget/fy02/index.html>.

responsibility of paying for the current generation's overspending" (p. 22), and that "we must ensure that we rein in excessive Government spending" (p. 23).

In the immediate aftermath of the attacks of September 11, 2001, discussions of budget policy placed less emphasis on spending restraint (see, for example, Address Before a Joint Session of the Congress on the State of the Union, 1/29/02, p. 4). Later presidential statements, however, returned to calls for spending restraint similar to those in 2001. For example, in his 2004 State of the Union Address, Bush stated: "I will send you a budget that funds the war, protects the homeland, and meets important domestic needs, while limiting the growth in discretionary spending By doing so, we can cut the deficit in half over the next five years" (Address Before a Joint Session of the Congress on the State of the Union, 1/20/04, p. 5). Similarly, in the 2007 State of the Union Address, he said, "What we need is spending discipline I will submit a budget that eliminates the Federal deficit within the next 5 years" (Address Before a Joint Session of the Congress on the State of the Union, 1/23/07, p. 1). While these statements were very similar to those before the tax cuts, actual budget conditions had changed substantially: revenues had fallen and the overall budget had shifted from surplus to deficit. The similarity in the rhetoric despite the large changes in the deficit suggests that there was not a link between the level of revenues and the perceived need for spending restraint.

As in the other episodes, the additional defense and security spending in this episode was viewed as being of such importance that the government's revenue situation was essentially irrelevant. For example, in his January 29, 2002 State of the Union Address, Bush stated: "Our first priority must always be the security of our nation, and that will be reflected in the budget I send to Congress. My budget supports three great goals for America: We will win this war; we'll protect our homeland; and we will revive our economy" (Address Before a Joint Session of the Congress on the State of the Union, p. 3). Subsequent Budgets almost all listed homeland security and combating terrorism as the highest priorities.

Panel (d) of Figure 10 plots the behavior of the major categories of spending in this episode. As in the other episodes, overall spending growth did not slow. In the five years following the passage of the first cut in 2001Q2, spending grew by 23 percent. The growth in spending was greatest in defense:

national defense purchases rose by 34 percent over this period, while nondefense purchases rose by 18 percent and current transfer payments by 20 percent.

The events of September 11, 2001 were clearly an important outside influence on spending. Some of the behavior of total expenditures and the shift toward national defense purchases surely reflects the impact of this outside development rather than the effect of the tax cuts. On the other hand, there was one important spending action that is not well reflected in our spending measures. The addition of prescription drug coverage to Medicare, enacted in December 2003, was expected to have only a modest short-run effect on spending but to raise its path substantially over time. Thus, although the change was enacted soon after the tax cuts, most of its impact on spending will almost surely come after the period considered in our regressions.

One notable feature of this episode is that the tax cuts were not soon followed by counteracting tax increases. A modest countercyclical tax cut was enacted in March 2002, in the wake of the September 11 attacks. The only important tax increase was that the bonus depreciation provisions included in the 2002 bill, and then expanded and slightly extended as part of the 2003 tax bill, were allowed to expire as scheduled at the end of 2004. Thus the issue of how the government will eventually deal with the loss of revenues from the 2001 and 2003 tax cuts remains open.

E. Assessment

Examination of the four episodes of major long-run tax cuts reinforces the findings from the statistical work: there is little evidence of a starve the beast effect. The one aspect of the episodes that is somewhat consistent with the hypothesis that tax cuts reduce government spending is the narrative record of the budget process. Although the presidents in two of the episodes (Johnson and Bush) appear to have paid little attention to the impact of the tax cuts on revenues in formulating their budget policies, the presidents in the other two (Truman and Reagan) cited the level of revenues as a consideration in forming budget policy. Even in these cases, however, other factors were clearly much more important, and to a considerable extent the concern over revenues led not to advocacy of spending reductions, but to support

(or acceptance) of tax increases.

The actual behavior of spending in all four episodes provides no support for the starve the beast hypothesis. In no episode was there a discernible slowdown in spending following the tax cut. Indeed, in two of the episodes (1948 and 1964) there was a marked acceleration of spending. This is similar to the overall statistical finding of a positive (though only marginally significant) effect of tax cuts on spending, and suggests that the regression results reflect a consistent pattern in the data rather than the effects of outliers. Likewise, in all four cases we see the shift in expenditures toward defense spending identified by the regressions.

Examination of other influences on spending in the episodes fails to rescue the starve the beast hypothesis. On the one hand, in three of the episodes, there were important external developments that acted to raise defense spending (the outbreak of the Korean War, the expansion of the Vietnam War, and the attacks of September 11, 2001). By itself, this pattern would suggest that the regressions might overestimate the positive effects of tax cuts on spending.

Two considerations, however, point in the opposite direction. First, one tax cut (1981) coincided with the election of a president who had a strong commitment to reducing the size of government. This would suggest that the positive impact of tax cuts on spending might be even larger than that implied by the regressions. Second, there were significant actions taken in the episodes to increase spending that had important effects after the five-year window considered in our baseline regressions. For example, in two of the episodes (1964 and 2001/03), the government enacted major changes in the provision of medical care for the elderly that had very large implications for the long-term path of government spending. Similarly, in all four cases, the government enacted increases in defense spending that lasted for extended periods. Since our regressions miss much of the effects of these developments, this too suggests that the regressions may underestimate the extent to which tax cuts increase spending. Thus, examination of other factors affecting spending in the episodes suggests that on net the regressions do not overstate the evidence against the starve the beast hypothesis.

The discussion of special factors also provides insight into our finding that there is a shift in the

composition of spending toward defense following tax cuts. The fact that there were important outside developments raising defense spending in three of the episodes of major tax cuts suggests that the correlation is partly due to chance. The fourth episode (the Reagan tax cut) is a case where the proponent of a tax cut was also an advocate of a strong military. Thus, it is possible that an omitted variable explains some of the correlation. Finally, the evidence from the 1948 and 2001/2003 episodes is supportive of the view that defense spending is the easiest type to justify in tight budgetary times, and thus that the shift in composition may be somewhat causal.

The behavior of subsequent tax changes in the episodes is consistent with the pattern shown by the regressions. In three of the cases, there were substantial tax increases within five years that offset a substantial fraction of the initial tax cut. Only in the most recent episode has there not been an offsetting tax increase of some kind. Perhaps more striking than the pattern is what policymakers said about the tax increases. In all three cases they referred directly to the need to raise taxes to counter the macroeconomic and budgetary effects of the original tax cuts. And in two cases (1948 and 1964), the president said explicitly that raising taxes was preferable to cutting spending.

V. CONCLUSIONS

The starve the beast hypothesis – that is, the idea that tax cuts restrain government spending – is a central argument for tax reduction. Despite its importance, however, the hypothesis has been subject to few tests, and those tests are far from definitive.

This paper tests the starve the beast hypothesis by examining the behavior of government spending following tax changes motivated by long-run considerations. Because these tax changes are not motivated by factors that are likely to have an important direct effect on government spending, they are the most appropriate for testing the theory. The results provide no evidence of a starve the beast effect: following long-run tax cuts, government spending does not fall. Indeed, if anything, spending rises,

providing some support for the alternative view of fiscal illusion or shared fiscal irresponsibility. These findings are highly robust. Detailed examination of the four largest postwar episodes of long-run tax cuts reinforces the statistical findings.

We also identify a potentially powerful source of bias in tests of the starve the beast hypothesis that use data on overall revenues and spending. Some tax changes are explicitly motivated by contemporaneous or planned changes in spending. Not surprisingly, these tax changes are followed by large spending changes in the same direction. Causation, however, runs from the decisions to raise spending to the tax changes. For the full postwar sample, this type of tax change is sufficiently common that it causes the overall relationship between tax revenues and spending to be significantly positive. Excluding these spending-driven changes makes the relationship negative and marginally significant.

We reach two other main conclusions concerning the budgetary effects of tax changes. First, long-run tax cuts are offset by legislated tax increases and non-legislated increases in revenues over the next several years. Thus, it appears that in the wake of tax cuts, budget balance is restored mainly on the tax side rather than the spending side. Second, there is suggestive evidence that tax cuts cause a shift in the composition of government expenditures toward defense spending.

There are two main caveats to our conclusions about the starve the beast hypothesis. First, because our estimates are not highly precise, the hypothesis that tax cuts exert some restraining influence on spending cannot be rejected. Second, although we find that the fall in revenues caused by a tax cut disappears after a few years, some of this disappearance is most likely the result of a temporary output boom. Thus, we do not completely resolve the issue of how the government restores long-run budget balance following a tax cut. Since the government's long-run budgetary situation deteriorated substantially over the period we consider, to some extent this limitation is inherent: not all of the offsetting actions have yet occurred. Both caveats suggest that the conclusion that tax cuts do not restrain government spending at all may be too strong. Nonetheless, over the period we consider, there is no evidence of any such effect.

The finding that tax cuts do not appear to restrain government spending could obviously have

implications for policy. At the very least, policymakers should be aware that the historical experience suggests that tax cuts tend to lead to tax increases rather than to spending cuts. The finding also has implications for models that assume the existence of a starve the beast effect. For example, Bohn (1992) argues that one reason for Ricardian equivalence to fail is that a tax cut implies that government spending will be lower; as a result, a tax cut leads households to revise their estimates of the present value of their present and future liabilities down, and so increase their consumption. Similarly, a restraining effect of tax cuts on government spending plays a central role in the theories of strategic debt accumulation of Persson and Svensson (1989), Tabellini and Alesina (1990), and others. If decision-makers understand that tax cuts do not in fact reduce government spending, none of these mechanisms can operate.

REFERENCES

- Anderson, William, Myles S. Wallace, and John T. Warner. 1986. "Government Spending and Taxation: What Causes What?" Southern Economic Journal 52 (January): 630-639.
- Auerbach, Alan J. 2000. "Formation of Fiscal Policy: The Experience of the Past Twenty-Five Years." Federal Reserve Bank of New York Economic Policy Review 6 (April): 9-23.
- Auerbach, Alan J. 2003. "Fiscal Policy, Past and Present." Brookings Papers on Economic Activity, no. 1, pp. 75-138.
- Barro, Robert J. 2003. "There's a Lot to Like about Bush's Tax Plan." Business Week, February 24, p. 28.
- Becker, Gary S., Edward P. Lazear, and Kevin M. Murphy. 2003. "The Double Benefit of Tax Cuts." Wall Street Journal, October 7, p. A20.
- Becker, Gary S., and Casey B. Mulligan. 2003. "Deadweight Costs and the Size of Government." Journal of Law and Economics 46 (October): 293-340.
- Blanchard, Olivier, and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." Quarterly Journal of Economics 117 (November): 1329-1368.
- Bohn, Henning. 1991. "Budget Balance through Revenue or Spending Adjustments? Some Historical Evidence for the United States." Journal of Monetary Economics 27 (June): 333-359.
- Bohn, Henning. 1992. "Endogenous Government Spending and Ricardian Equivalence." Economic Journal 102 (May): 588-597.
- Budget of the United States Government. U.S. Office of Management and Budget. Various years.
- Buchanan, James M., and Richard E. Wagner. 1977. Democracy in Deficit: The Political Legacy of Lord Keynes. New York: Academic Press.
- Calomiris, Charles W., and Kevin A. Hassett. 2002. "Marginal Tax Rate Cuts and the Public Tax Debate." National Tax Journal 55 (March): 119-131.
- Congressional Record. U.S. Congress. Various issues.
- Economic Report of the President. U.S. Office of the President. Various years.
- Friedman, Milton. 1967. "Fiscal Responsibility." Newsweek, August 7, p. 68.
- Hines, Frank T. 1943. "National Service Life Insurance." Annals of the American Academy of Political and Social Science 227 (May): 83-93.
- Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen. 1989. "The Revenue-Expenditures Nexus: Evidence from Local Government Data." International Economic Review 30 (May): 415-429.

- Miller, Stephen M., and Frank S. Russek. 1990. "Co-Integration and Error-Correction Models: The Temporal Causality between Government Taxes and Spending." Southern Economic Journal 57 (July): 221-229.
- Niskanen, William A. 1978. "Deficits, Government Spending, and Inflation: What is the Evidence?" Journal of Monetary Economics 4 (August): 591-602.
- Persson, Torsten, and Lars E. O. Svensson. 1989. "Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preferences." Quarterly Journal of Economics 104 (May): 325-345.
- Ram, Rati. 1988. "Additional Evidence on Causality between Government Revenues and Government Expenditure." Southern Economic Journal 54 (January): 763-769.
- Romer, Christina D., and David H. Romer. 2007a. "A Narrative Analysis of Postwar Tax Changes." Unpublished manuscript, University of California, Berkeley (August).
- Romer, Christina D., and David H. Romer. 2007b. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." Unpublished manuscript, University of California, Berkeley (July).
- Social Security Bulletin. Social Security Administration. Various issues.
- Survey of Current Business. Bureau of Economic Analysis. Various issues.
- Tabellini, Guido, and Alberto Alesina. 1990. "Voting on the Budget Deficit." American Economic Review 80 (March): 37-49.
- von Furstenberg, George M., R. Jeffrey Green, and Jin-Ho Jeong. 1986. "Tax and Spend, or Spend and Tax?" Review of Economics and Statistics 68 (May): 179-188.
- Woolley, John and Gerhard Peters. The American Presidency Project (www.presidency.ucsb.edu).

Table 1
Estimated Impact of Tax Changes on Total Expenditures

Variable	Coefficient
Constant	0.72 (0.25)
Tax Change:	
Lag 0	0.33 (0.85)
Lag 1	0.46 (0.85)
Lag 2	0.05 (0.85)
Lag 3	-0.27 (0.84)
Lag 4	-0.98 (0.88)
Lag 5	-1.59 (0.88)
Lag 6	0.17 (0.88)
Lag 7	-1.46 (0.76)
Lag 8	2.72 (0.77)
Lag 9	2.63 (0.77)
Lag 10	-1.02 (0.77)
Lag 11	-1.52 (0.76)
Lag 12	-2.32 (0.76)
Lag 13	-2.21 (0.76)
Lag 14	-1.00 (0.80)
Lag 15	0.37 (0.80)
Lag 16	0.34 (0.81)
Lag 17	-0.06 (0.78)
Lag 18	0.43 (0.78)
Lag 19	0.91 (0.78)
Lag 20	0.02 (0.80)
 R ²	 0.20
D.W.	1.89
s.e.e.	2.74

Note: Estimates of equation (1) in the text using long-run tax changes and total gross expenditures less interest payments. The sample period is 1950Q1–2006Q4. The numbers in parentheses are standard errors.

Figure 1
Legislated Tax Changes Classified by Motivation

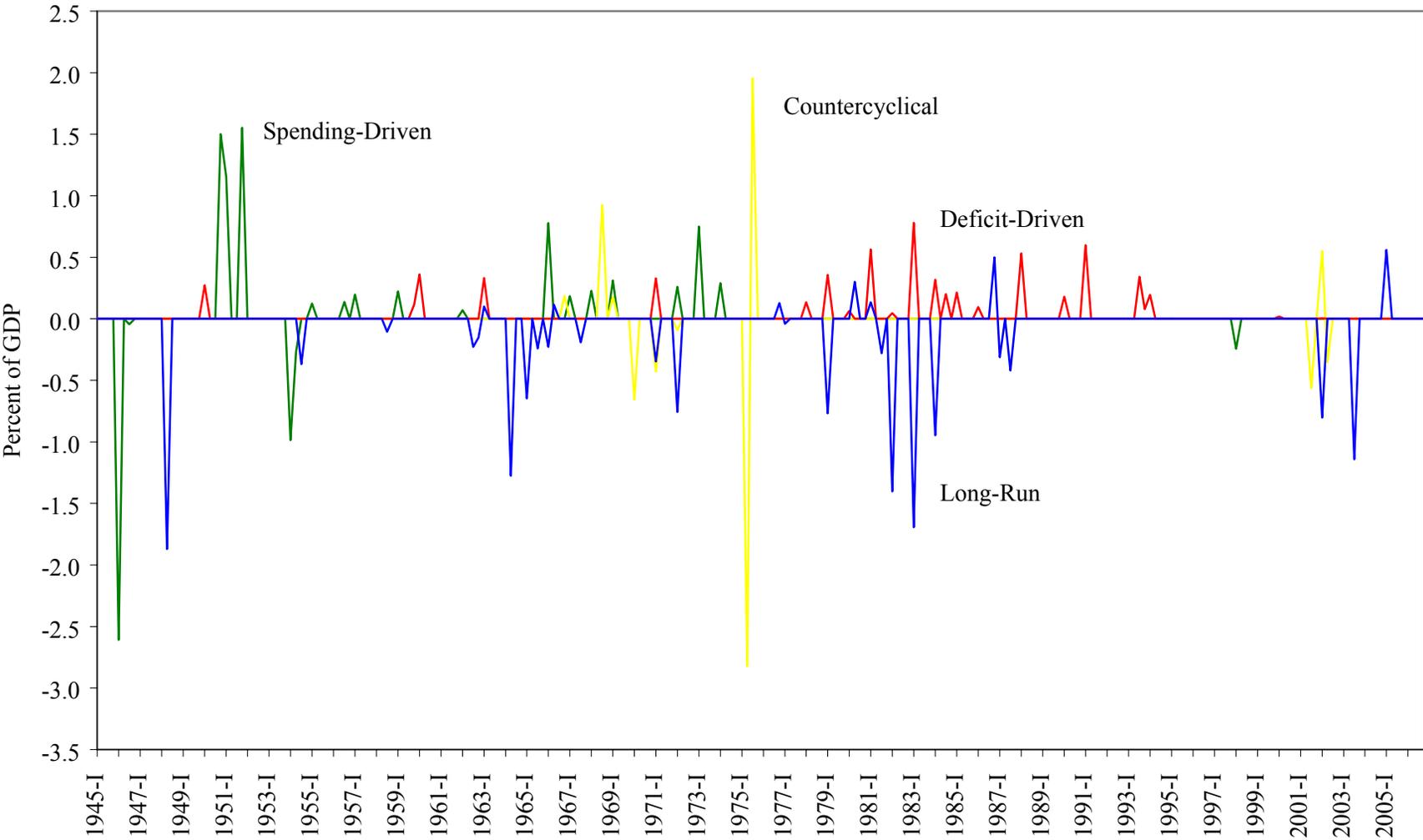


Figure 2
Estimated Impact of a Tax Cut of 1% of GDP on Total Expenditures

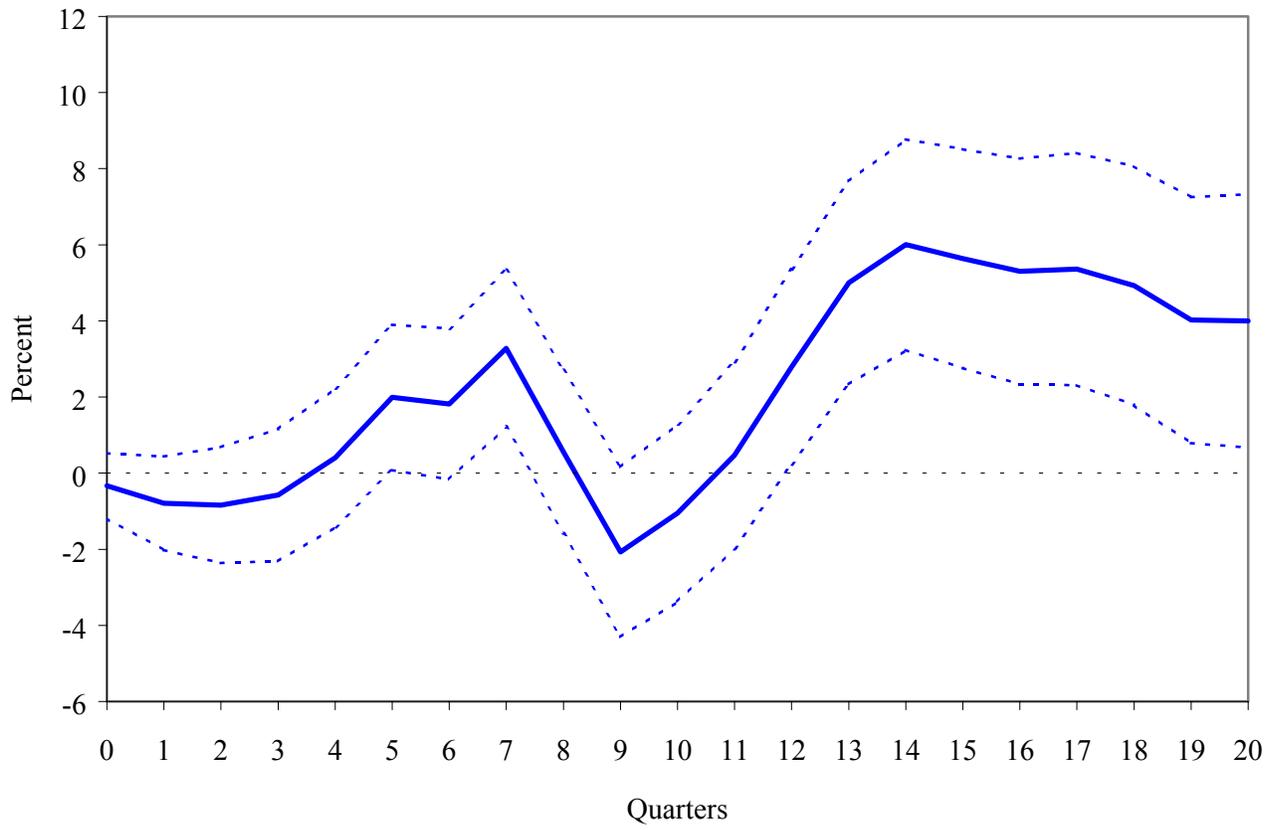
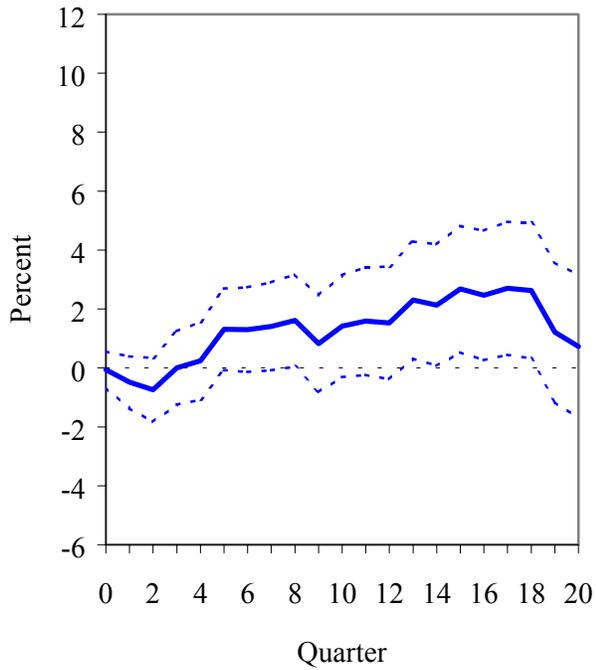
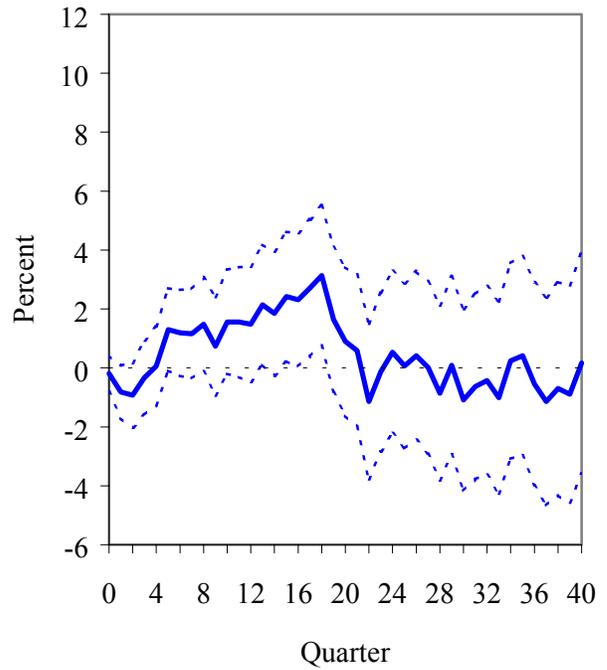


Figure 3
Estimated Impact of a Tax Cut of 1% of GDP on Total Expenditures
Using Different Specifications

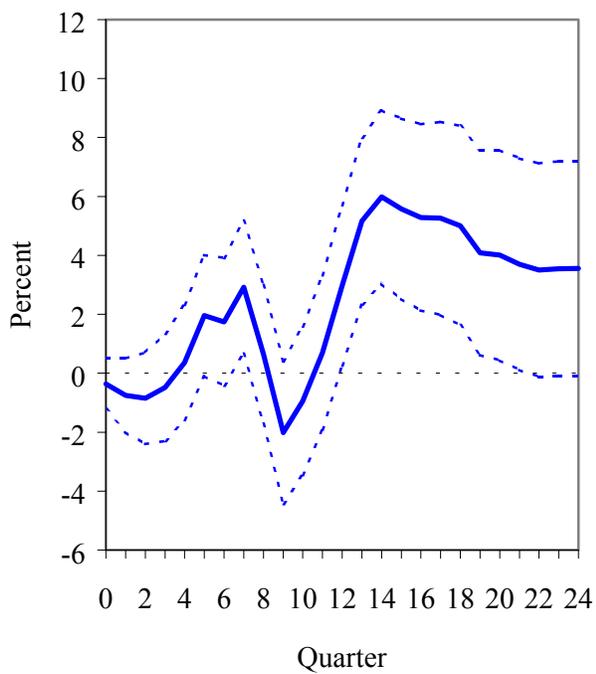
a. Starting Sample in 1957Q1



b. Including 40 Lags of Tax Changes



c. Including 11 Lags of Total Expenditures



d. Using a Two-Variable VAR

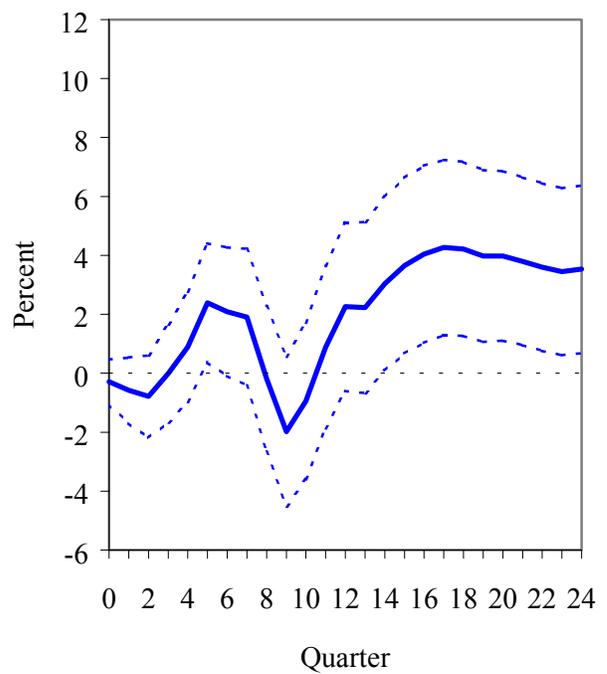
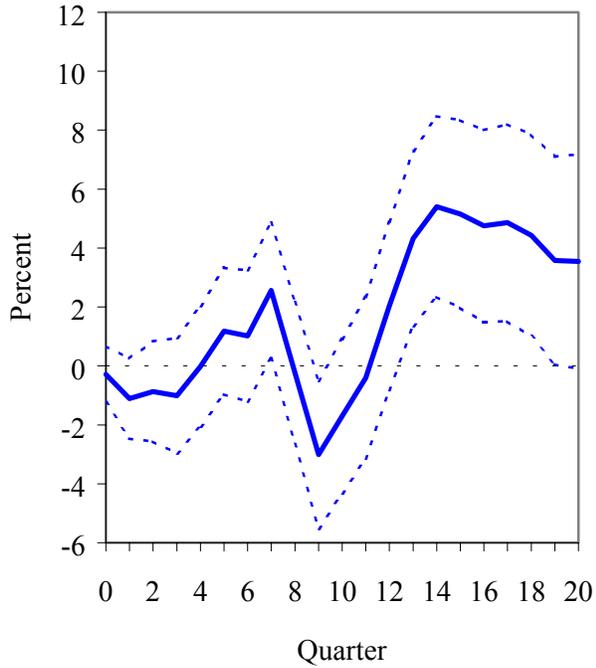
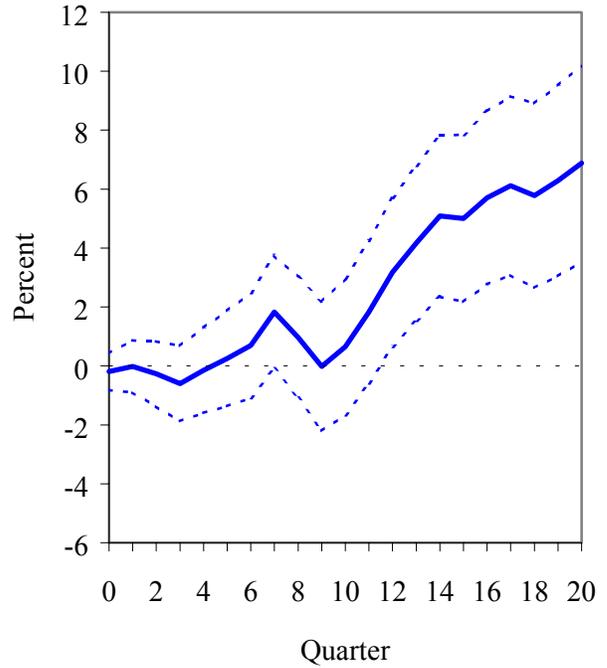


Figure 3 (continued)
 Estimated Impact of a Tax Cut of 1% of GDP on Total Expenditures
 Using Different Specifications

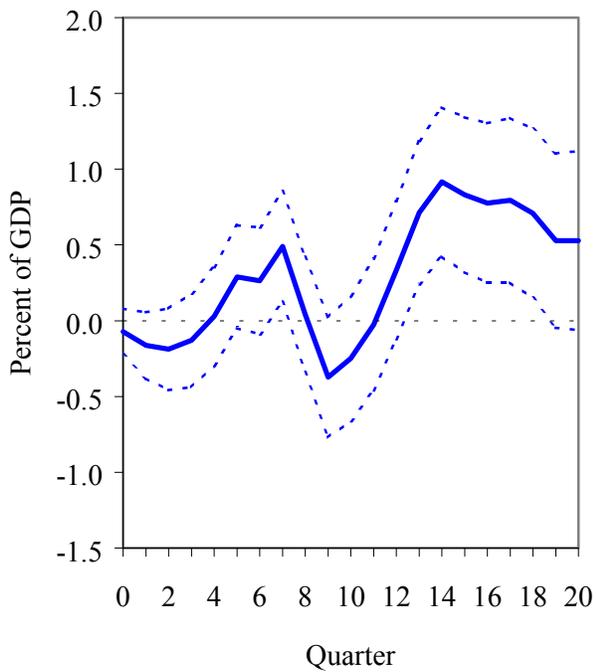
e. Including Democratic Dummy Variable



f. Using PDV Measure of Tax Changes



g. Expenditures as a Share of Trend GDP



h. Expenditures as a Share of Actual GDP

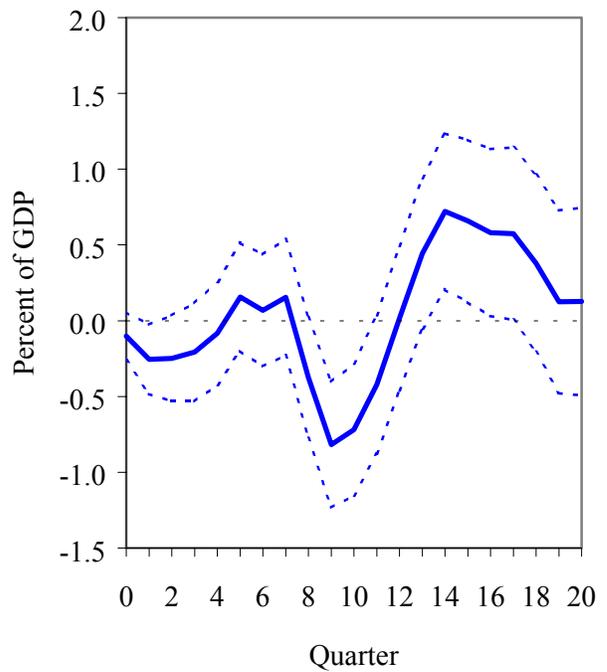
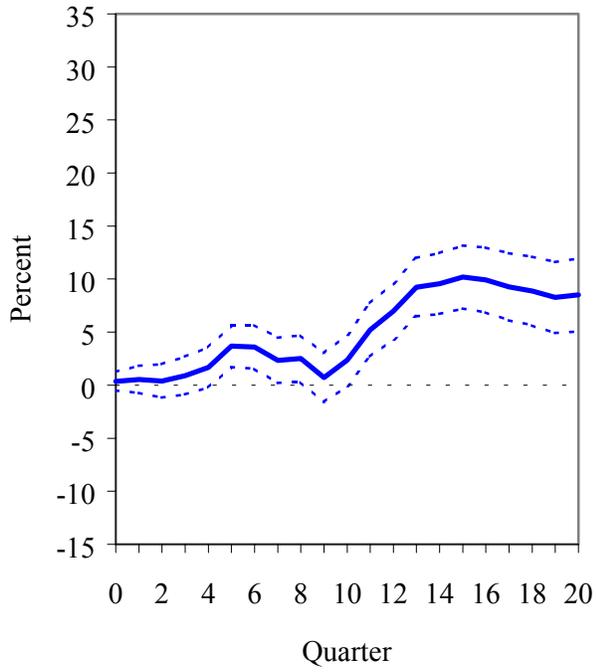
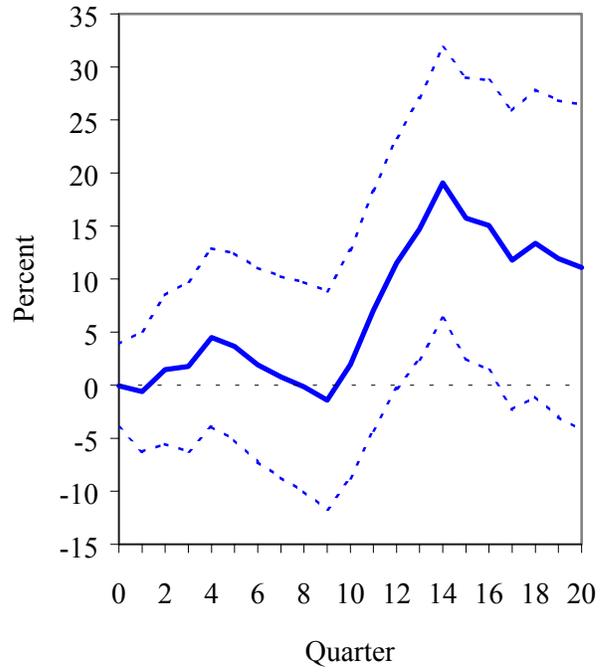


Figure 4
Estimated Impact of a Tax Cut of 1% of GDP on Expenditures by Type

a. Consumption Expenditures



b. Total Government Investment



c. Current Transfer Payments

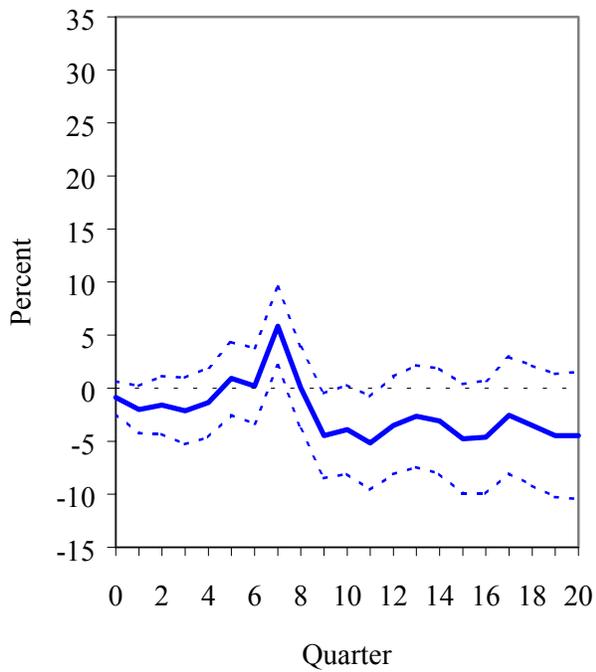
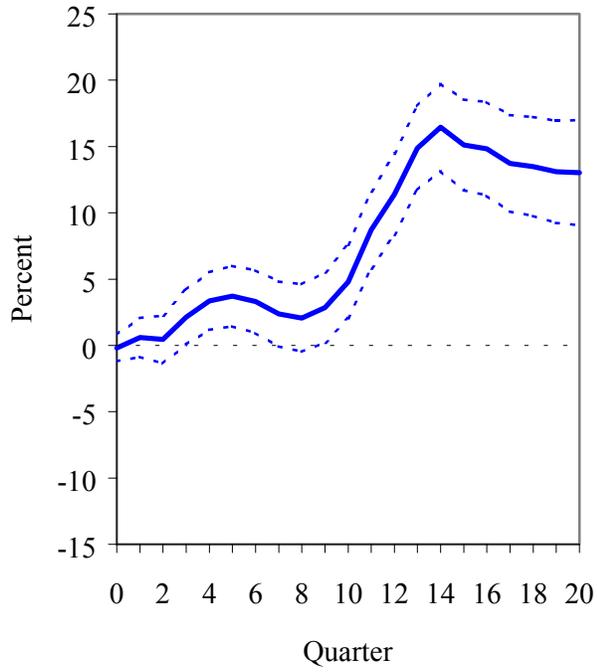


Figure 5
Estimated Impact of a Tax Cut of 1% of GDP on Expenditures by Sector

a. National Defense Purchases



b. Nondefense Purchases

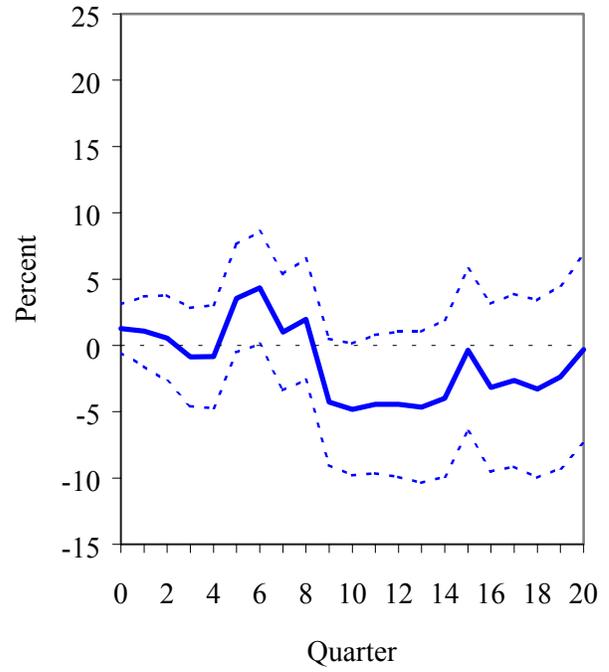


Figure 6
 Estimated Impact of a Tax Cut of 1% of GDP on Total Expenditures by Type of Tax Change

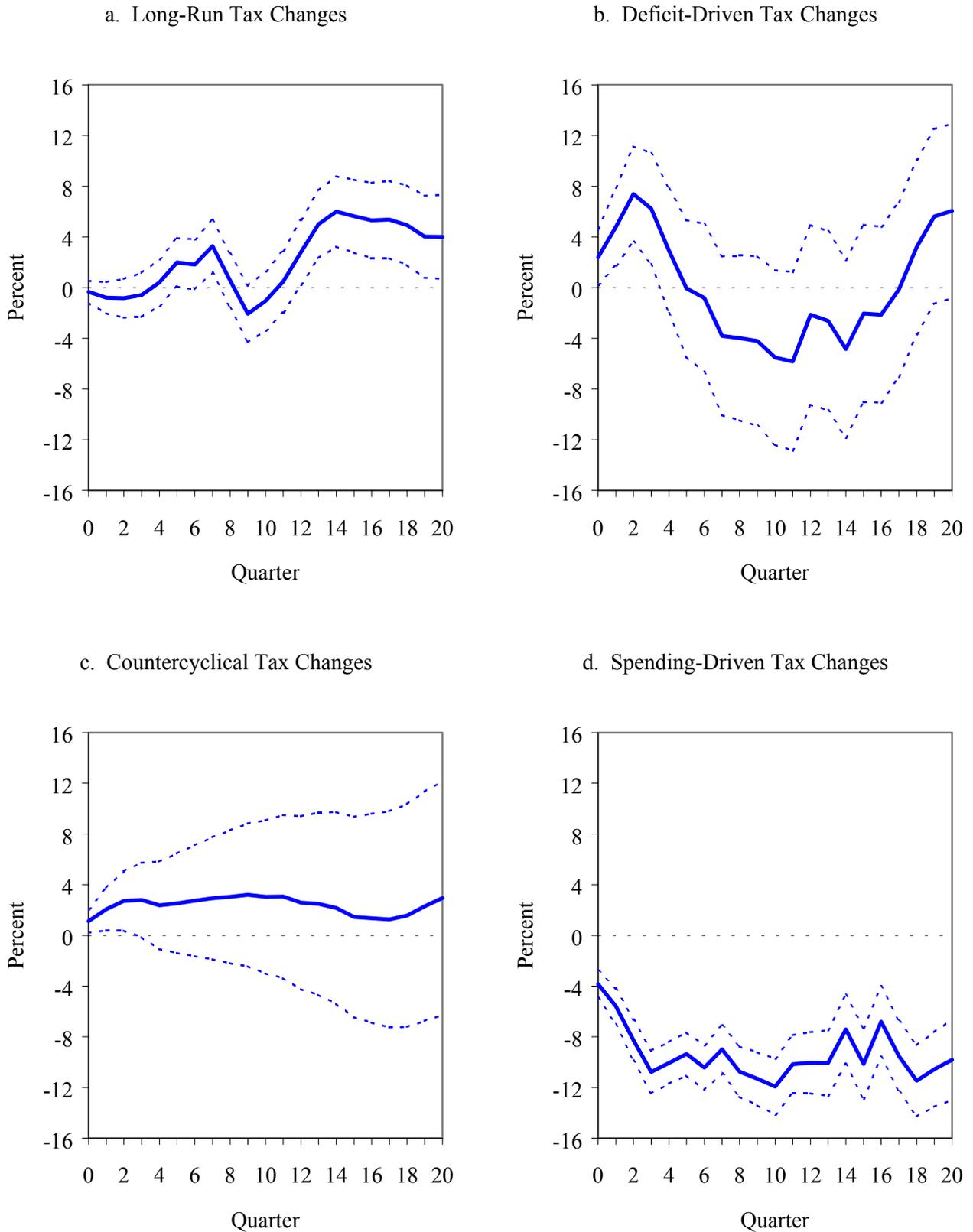
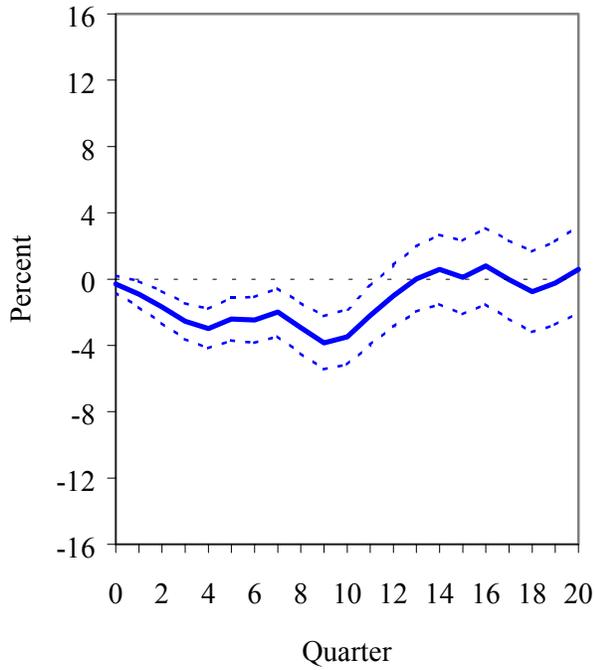
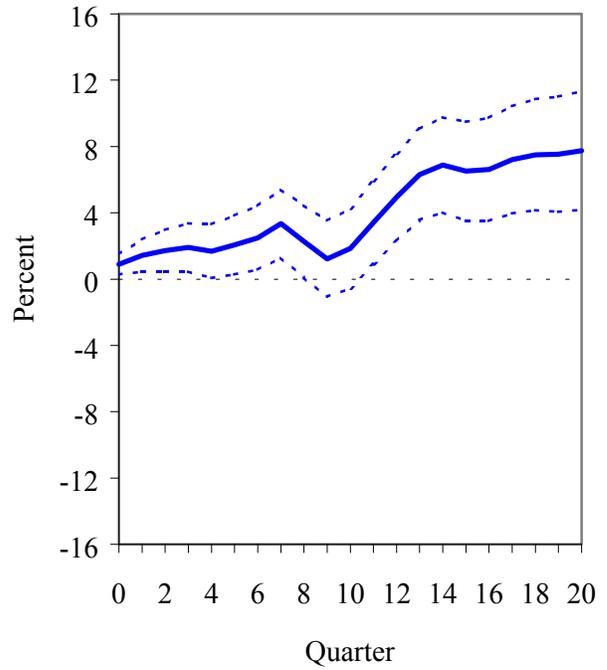


Figure 6 (continued)
 Estimated Impact of a Tax Cut of 1% of GDP on Total Expenditures by Type of Tax Change

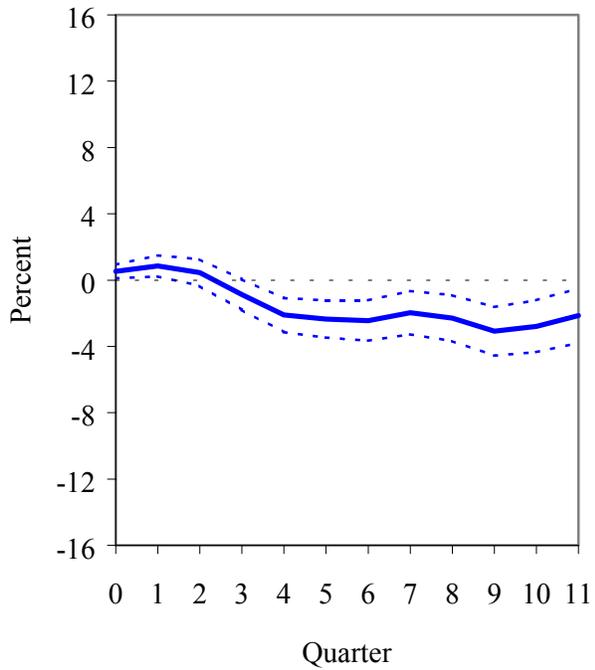
e. All Legislated Tax Changes



f. All Legislated Less Spending-Driven



g. Cyclically Adjusted Revenues



h. C. A. Revenues Less Spending-Driven

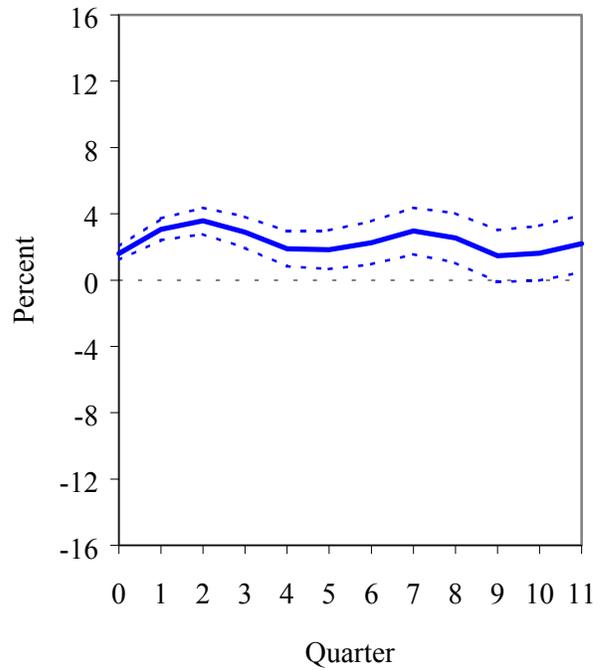
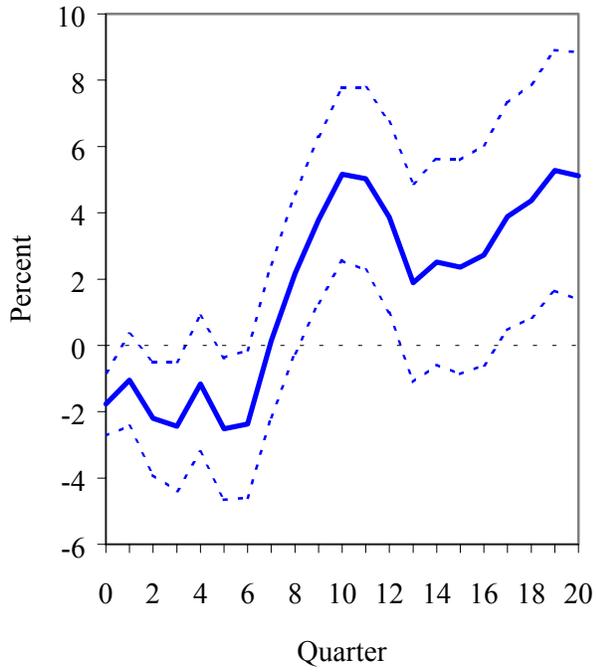


Figure 7
Estimated Impact of a Tax Cut of 1% of GDP on Total Receipts

a. Full Sample



b. Starting Sample in 1957Q1

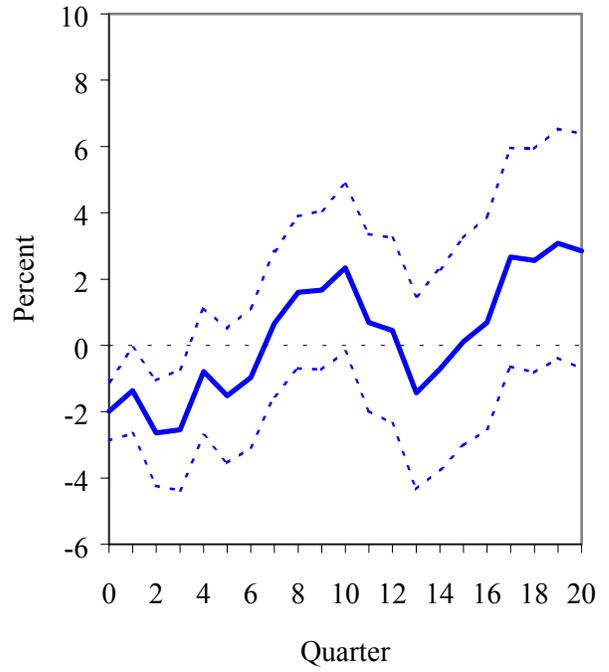
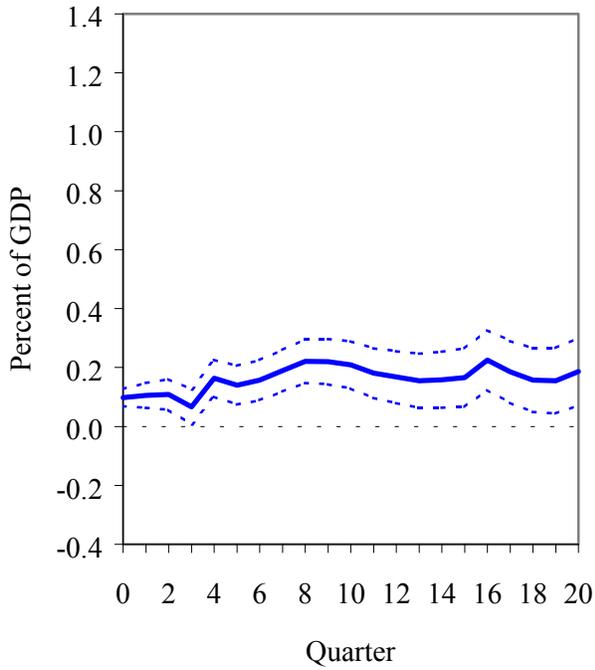
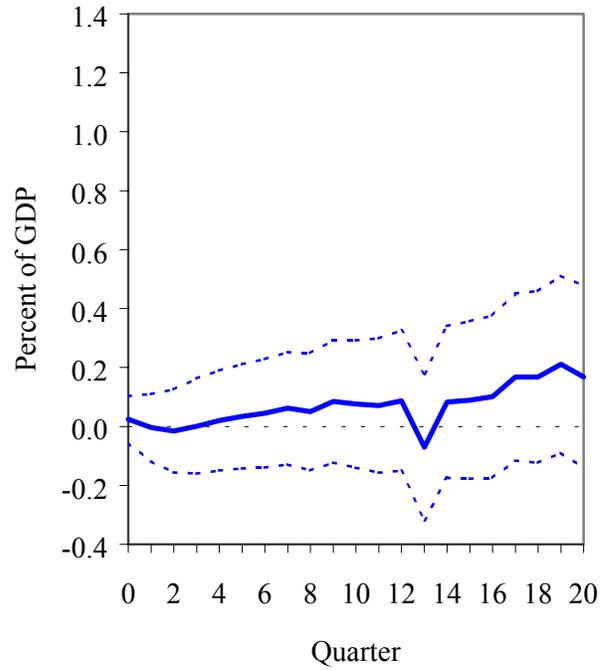


Figure 8
 Estimated Impact of a Long-Run Tax Cut of 1% of GDP on Tax Changes by Type

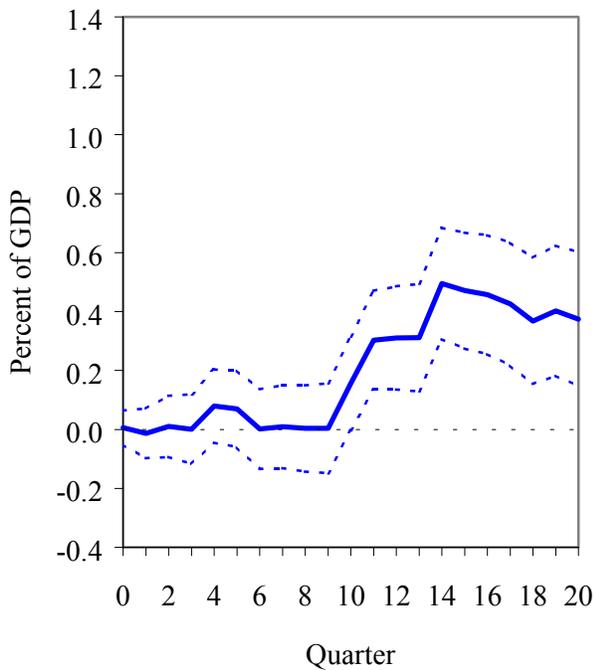
a. On Deficit-Driven Tax Changes



b. On Countercyclical Tax Changes



c. On Spending-Driven Tax Changes



d. On All Other Tax Changes Combined

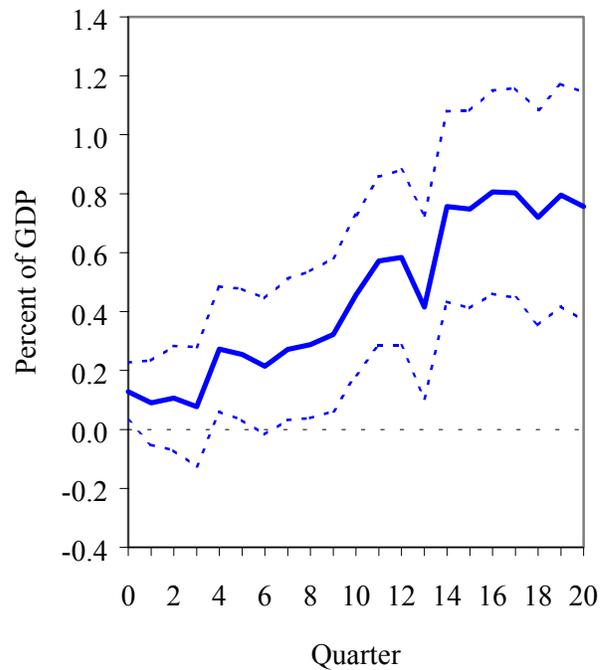
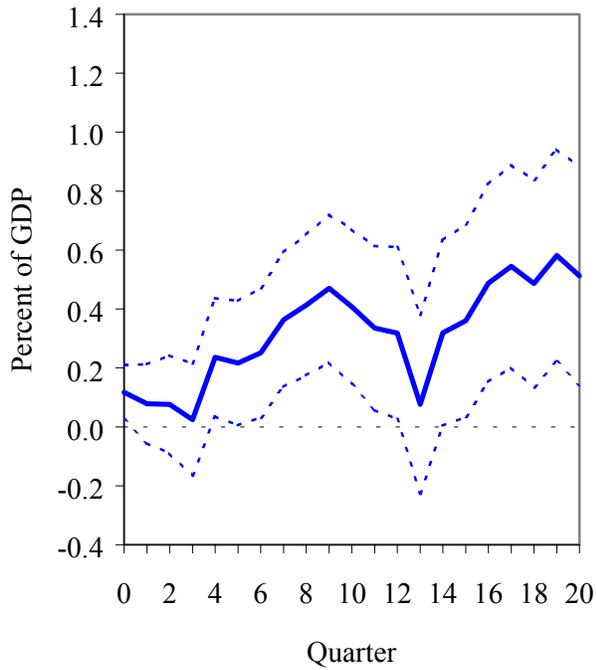


Figure 9
Estimated Impact of a Long-Run Tax Cut of 1% of GDP on All Other Tax Changes Combined
Using Different Specifications

a. Starting Sample in 1957Q1



b. Including 40 Lags

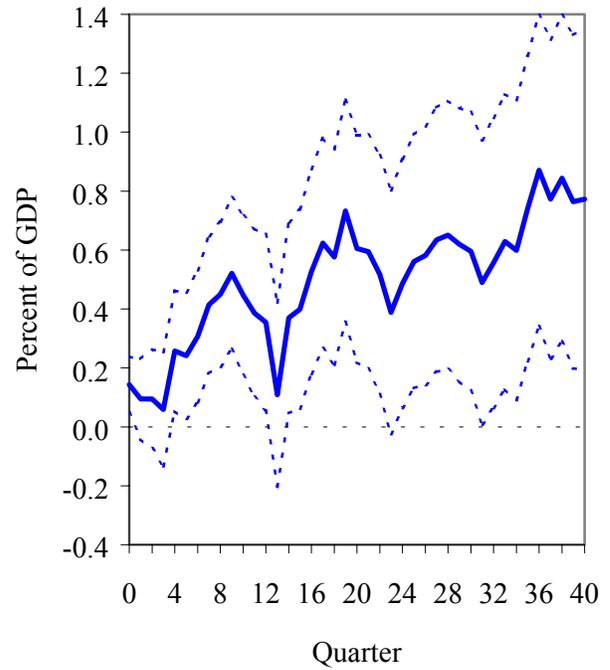
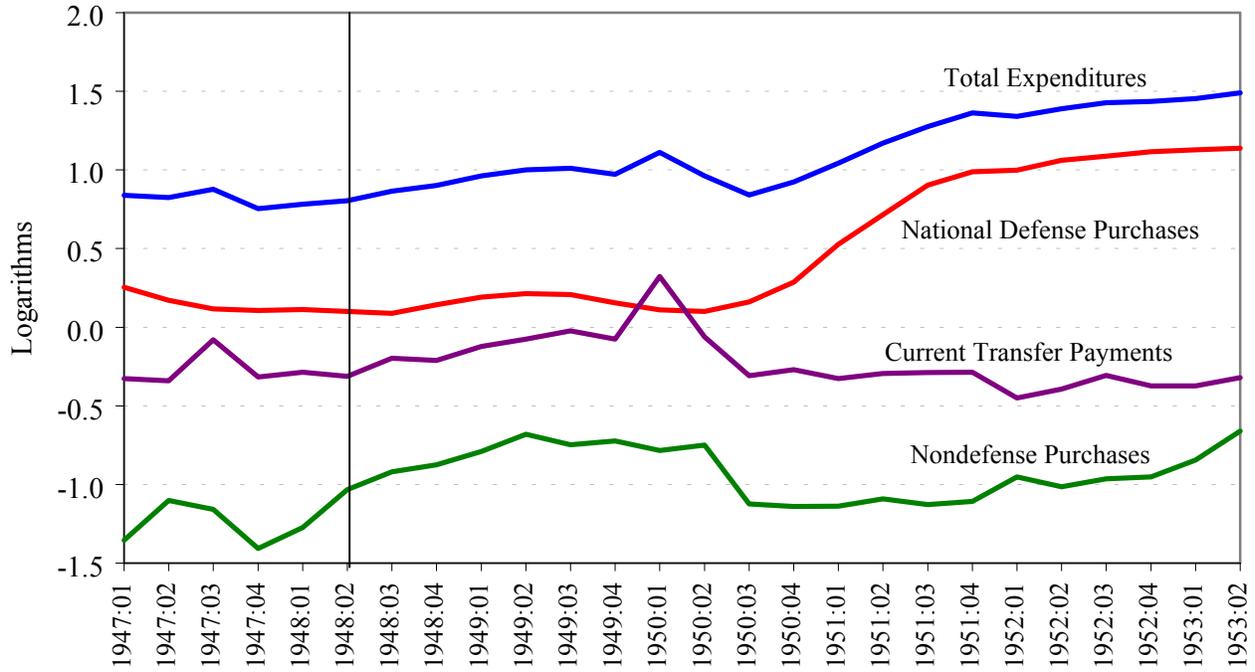


Figure 10
Real Federal Government Expenditures after Four Key Long-Run Tax Cuts

a. Revenue Act of 1948



b. Revenue Act of 1964

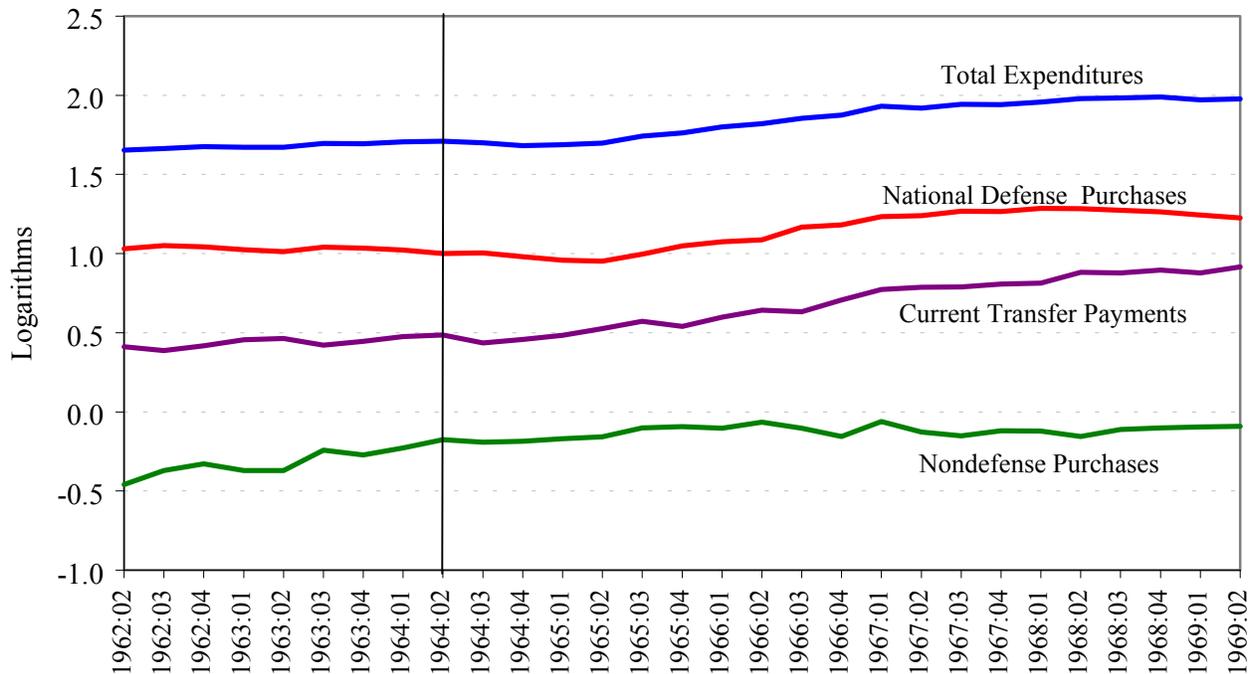
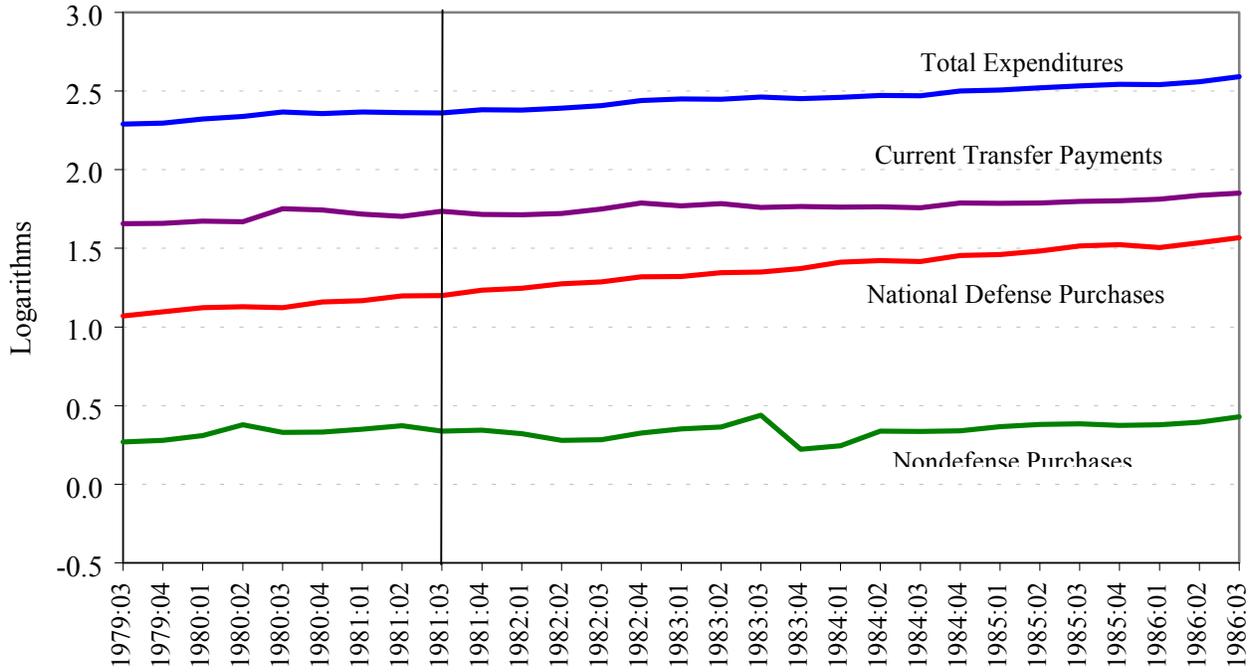


Figure 10 (continued)
 Real Federal Government Expenditures after Four Key Long-Run Tax Cuts

c. Economic Recovery Tax Act of 1981



d. EGTRRA of 2001 and JGTRRA of 2003

