DAVID H. ROMER
University of California, Berkeley

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 derive more reliable estimates, this paper examines the behavior of government expenditure following legislated tax changes that narrative sources suggest are largely uncorrelated with other factors affecting spending. The results provide no support for the hypothesis that tax cuts restrain government spending; indeed, the point estimates suggest that tax cuts 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.

I n a speech urging passage of the 1981 tax cuts, President 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.[1]

1. "Address to the Nation on the Economy," February 5, 1981, p. 2. Quotations from presidential speeches are from John T. Woolley and Gerhard Peters, *The American Presidency Project* (www.presidency.ucsb.edu), an online database of presidential documents.

139

This idea that cutting taxes will lead to a reduction in government spending is often referred to as the "starve the beast" hypothesis: the most effective way to shrink the size of government is to reduce the revenue that feeds it. This view has been embraced not just by politicians but also by distinguished economists from Milton Friedman to Robert Barro.[2]

Of course, the starve-the-beast hypothesis is not the only view of how tax cuts affect expenditure. 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 James Buchanan and Richard Wagner (1977) and by William 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 (see, for example, Gale and Orszag 2004).

The question of how tax cuts affect government spending is clearly an empirical one. And, indeed, there have been attempts to investigate the aggregate relationship between revenue and spending. However, such examinations of correlations cannot settle the issue. Changes in revenue occur 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. 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 economic growth.

The relationship between revenue and spending is surely not independent of the causes of changes in revenue. For example, if spending-driven tax changes are common, a regression of spending on revenue 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 counter-cyclical tax changes are common, one might expect to see expenditure rising

2. See, for example, Milton Friedman, "Fiscal Responsibility," *Newsweek,* August 7, 1967, p. 68; Robert J. Barro, "There's a Lot to Like about Bush's Tax Plan," *Business Week,* February 24, 2003, p. 28; Gary S. Becker, Edward P. Lazear, and Kevin M. Murphy, "The Double Benefit of Tax Cuts," *Wall Street Journal,* October 7, 2003, p. A20.

after declines in revenue, because spending on unemployment insurance and other relief measures typically rises in bad economic times. In this case, both revenue and spending are being driven by an omitted variable: the state of the economy. These examples suggest that looking at the aggregate relationship between revenue 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 2009), we identified all significant legislated tax changes in the United States over the period 1945–2007. 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 motivations to isolate those 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 enacted 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, intended to promote various long-run objectives, such as spurring productivity growth, improving efficiency, or, as in the case of the 1981 Reagan tax cut, shrinking the size of government. Examining the behavior of government spending following these long-run tax changes should provide a relatively unbiased test of the starve-the-beast hypothesis.

We examine the relationship between real government expenditure and our measure of long-run tax changes in a variety of specifications. We find no support for the hypothesis that a relatively exogenous decline in taxes lowers future government spending. In our baseline specification, the estimates in fact suggest a substantial and marginally significant *positive* impact of tax cuts on government spending. The finding of a lack of support for the starve-the-beast hypothesis is highly robust. The evidence of an opposite-signed effect, in contrast, is not particularly strong or robust.

The result that spending does not fall following a tax cut raises an obvious question: how then does the government budget adjust in response to the cut? One possibility is that what gives is not spending but the tax cut itself. To investigate this possibility, we examine the response of both tax revenue and tax legislation to long-run tax cuts. We find that revenue falls in response to a long-run tax cut in the short run but then recovers after about two years. Most of this recovery is due to the fact that a large part

of a long-run tax cut is typically counteracted by legislated tax increases within the next several years. As we discuss, the fact that policymakers are able to adjust on the tax side helps to explain why they do not adjust on the spending side.

Although there have been numerous long-run tax changes spread fairly uniformly over the postwar era, four stand out as the largest and best known: the tax cut passed over President 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 tax cuts passed (along with some countercyclical actions) under President George W. Bush 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 important, 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 that concurrent developments, namely wars, account for some of the rise in spending in these episodes. But other concurrent developments caused measured spending changes to understate the effects of the spending decisions made in these episodes. In particular, three of the four episodes featured decisions to expand entitlement programs that had only modest short-term effects on spending but very large long-term effects. As a result, it appears unlikely that the failure of total expenditure to fall after these tax cuts was due to chance or unobserved factors.

As mentioned above, ours is not the first study to investigate the starve-the-beast hypothesis. The most common approach is some variation of a regression of spending on lagged revenue; examples include the studies by William Anderson, Myles Wallace, and John Warner (1986) and by Rati Ram (1988). More sophisticated versions of this methodology are pursued by Henning Bohn (1991) and Alan Auerbach (2000, 2003). Bohn, focusing on a long sample period dominated by wartime budgetary changes, examines the interrelationships between revenue and spending in a vector autoregression (VAR) 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).

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 revenue, 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 changes in the 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 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 identifies the types of tax actions best suited for testing the starve-the-beast hypothesis. Section II analyzes the relationship between tax changes and government expenditure and includes a plethora of robustness checks. Section III examines how changes in taxes affect future tax revenue and tax legislation. Section IV discusses spending and taxes in the four key episodes. Section V presents our conclusions and discusses the limitations of our analysis.

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 our classification of motivations and to discuss which types of tax changes are likely to yield informative estimates of the effects of tax changes on government spending. We also provide a brief overview of our identification

---

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. Although 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.

of the motivations for tax changes and of our findings about the patterns of legislated tax changes in the postwar era.

I.A. Classification of Motivations

Our classification and identification of the motivations for postwar legislated tax changes are described in detail in Romer and Romer (2009). That paper shows that the motivations for almost all tax changes have fallen into four broad categories.

One type of tax change consists of those motivated by contemporaneous changes in spending. Often, policymakers will introduce a new program or social benefit and raise taxes at about the same time to pay for it. This was true, for example, in the mid-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 change is the impetus for the tax change. Typically, such changes are tax increases, but spending-driven tax cuts are not unheard of.

A second type of tax change encompasses those made because policymakers believe that economic growth in the near term will be above or below 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 slow or unusually rapid expected growth.

A third type of tax change consists of those made to reduce an inherited budget deficit. By definition, these changes are all increases. A classic example is the 1993 tax increase under President Bill Clinton. This increase was undertaken not to finance a contemporaneous rise in spending, but to reduce a persistent deficit caused by past developments.

The fourth type consists of tax changes intended to raise long-run economic growth. This is a broad category that includes changes motivated by a range of factors. What unites these changes is that they are all designed to improve the long-term functioning of the economy. The most common motivation is a belief that lower tax rates will improve incentives and thereby spur long-run growth. Another motivation is a belief in the benefits of small government and a desire to return the people's money to them; a third is a desire to improve the efficiency and equity of the tax system. 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 general heading of tax changes aimed at raising long-run growth.

Most of these changes are cuts, but some of the tax reforms included in this category increased revenue.

I.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. What are needed are tax changes that are not systematically correlated with other factors influencing government spending. An obvious implication is that spending-driven tax changes are not appropriate observations to use. Causation in these episodes runs from the desired change in spending to the change in taxes. There is an omitted influence on spending—the prior decision to change spending—that is strongly correlated with these tax changes. Thus, if we have classified spending-driven tax changes correctly, there will be a positive correlation between these changes and spending changes by construction. Including spending-driven tax changes in a regression of spending changes on tax changes would therefore bias the results toward finding a starve-the-beast effect.

Similar reasoning suggests that examining spending changes following countercyclical and deficit-driven tax changes could also be problematic. In these cases, however, the likely bias is against the starve-the-beast hypothesis. In both cases there may be spending changes that are negatively correlated with the tax changes but not caused by them. Rather, both the tax and the 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 economic 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 revenue and spending decisions.

In the case of deficit-driven tax changes, 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 packages in a regression of spending changes on tax changes 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 increases. Therefore, although 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 tax changes are 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 alter government spending.

To the degree that focusing on this type of tax change may lead to bias, it is likely to be in the direction of finding a positive effect of taxes on spending. The ideal experiment for testing the starve-the-beast hypothesis would be a tax change resulting from factors that have no direct impact on spending. Our long-run tax changes, however, include tax cuts for which the possible induced reduction in future spending is sometimes cited as a motivation. As a result, there is a potential correlation between spending and tax changes in these episodes driven by a third factor: a desire for smaller government. Policymakers, in addition to cutting taxes to starve the beast, may take other actions to achieve this goal. Because this possible omitted variable bias works in the direction of supporting the starve-the-beast hypothesis, a finding of a positive relationship between taxes and spending would have to be treated with caution. Since we in fact find a negative relationship, there is less cause for concern. Also, our narrative analysis suggests that this potential bias is likely to be small. The desire for smaller government is rarely the primary motivation for long-run tax changes; a belief in the incentive effects of lower taxes is considerably more common, for example.

I.C. *Overview of the Narrative Analysis*

The implementation and results of our narrative analysis of postwar tax changes are described in Romer and Romer (2009). We use a detailed examination of a wide range of policy documents to identify all significant legislated tax changes over the period 1945–2007. We then identify the motivations policymakers gave for each action. We find that policymakers were usually 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.[4]

Figure 1 shows legislated postwar tax changes classified by motivation, measured by their expected revenue effects as a percent of nominal GDP.[5] The top panel shows the long-run changes, which are the key actions for our purposes. The graph makes clear that the vast majority of long-run tax actions are cuts. It also makes clear that long-run tax changes have been fairly evenly distributed over the postwar era. The largest were the 1948 tax cut, the Kennedy-Johnson tax cut in the mid-1960s, the Reagan tax cut in the early 1980s, and the two Bush tax cuts in the early 2000s.

The bottom panel shows the other types of tax changes. Although the first half of the postwar era saw a number of small, deficit-driven tax increases, the vast majority took place in the 1980s and early 1990s. Most of the deficit-driven increases were passed to deal with the long-run solvency of the Social Security and Medicare systems. Spending-driven changes are typically tax increases, and these were both frequent and relatively large in the first half of the postwar era. By far the largest were those in the Revenue Act of 1945 following the end of World War II, and those in the early 1950s to pay for the Korean War. Many of the other changes in this category were related to expansions of Social Security. Finally, explicitly counter-cyclical tax changes were confined to the fairly short period 1966–75 until they were resurrected as the reason for portions of the tax cuts in 2001 and 2002.

4. Tax actions are often retroactive for a quarter or two. Such changes have a much larger effect on liabilities in the initial quarter than in subsequent ones. In terms of 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 ignore the retroactive changes in forming our baseline estimates. Including the retroactive changes has almost no impact on any of the results, however.

5. The nominal GDP data are from the National Income and Product Accounts, table 1.1.5 (downloaded February 17, 2008). Quarterly nominal GDP data are available only after 1947. We therefore normalize the one tax change in 1946 using the annual nominal GDP figure for that year.

Figure 1. Legislated Tax Changes Classified by Motivation, 1945–2007

Long-run tax changes

Percent of GDP

Other tax changes

Percent of GDP
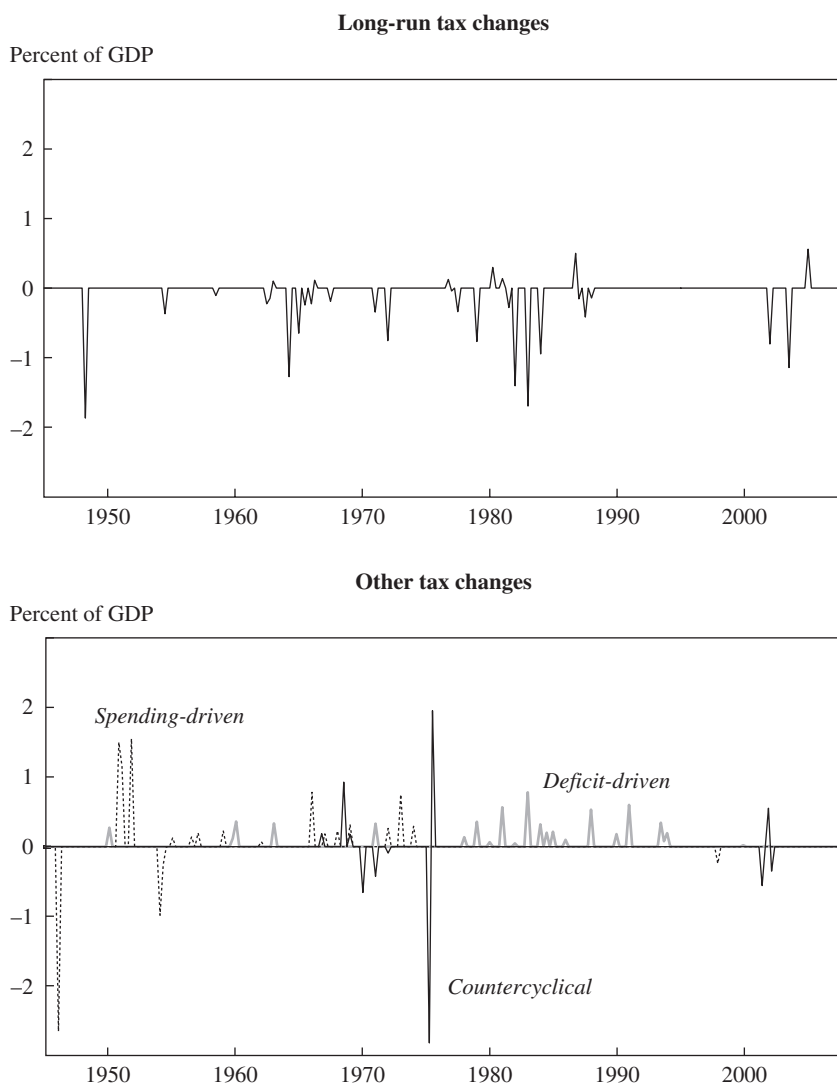
Source: Romer and Romer (2009).

II. The Effect of Tax Changes on Expenditure

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 sub-sequent changes in government spending. It includes a detailed analysis

of the robustness of the results. We also investigate the behavior of spending following other types of tax changes to see if there is evidence of bias when these changes are included.

II.A. Specification and Data

To estimate the effects of tax changes on government spending, we begin by estimating, using quarterly data, a simple reduced-form regression of the form

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

where ΔE is the change in the logarithm of real government expenditure and ΔT is our measure of long-run tax changes (specifically, the expected revenue effects, as a percent of nominal GDP, of the tax changes we identify as motivated by long-run considerations).

The key feature of long-run tax changes as we have defined them is that they are based on actions motivated by considerations largely unrelated to current spending, current macroeconomic conditions, or an inherited budget deficit. Our discussion above of why such long-run changes provide the best test of the starve-the-beast hypothesis suggests that they are unlikely to have a substantial systematic correlation 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, or that the dynamics of the relationship between tax changes and spending are more complicated than is expressed in equation 1. We therefore also consider a wide range of control variables and a variety of more complicated specifications.

We include a number of lags of the tax variable to allow for the possibility that the response of spending to tax changes is quite delayed or gradual. In our baseline specification we set the number of lags 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 the spending response, we focus on the cumulative effect at various horizons. We summarize the regression results by reporting the implied impact of a tax cut of 1 percent of GDP on the path of expenditure (in logarithms). For our baseline specification, the cumulative impact after n quarters is just the negative of 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 expenditure should be negative if the hypothesis is correct.

We use quarterly data on government expenditure from the National Income and Product Accounts (NIPA). Our series on long-run tax changes refers only to federal legislation. Therefore, we consider only the behavior of federal expenditure. What the Bureau of Economic Analysis (BEA) calls "total expenditures," however, includes two components that are not appropriate to include in considering the response of spending to tax changes. One is a deduction for the consumption of fixed capital (that is, depreciation). This largely reflects spending decisions in the distant past and so almost surely cannot show a starve-the-beast response. Thus, 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. We therefore exclude this type of spending. The resulting aggregate that we consider is thus total gross expenditure less interest. For simplicity, we refer to this as total expenditure in what follows.[6]

The NIPA expenditure data are expressed in nominal terms. Deflators exist for some components, such as defense and nondefense purchases, but not for others, especially those involving transfers. We therefore deflate total gross expenditure less interest by the price index for GDP (NIPA table 1.1.4, downloaded February 22, 2008).

Our data on tax changes begin in 1945Q1, and the data on expenditure in 1947Q1. Therefore, in the baseline specification, where we include 20 lags of the tax variable, the earliest starting date for the regression is 1950Q1. 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 forthcoming). We therefore also report estimates for regressions starting in 1957Q1. In both cases we carry the regressions through 2007Q4.

II.B. The Effect of Long-Run Tax Changes on Total Expenditure

Table 1 shows the results of estimating equation 1 for total expenditure using 20 lags of the long-run tax variable over the full sample. The coefficient estimates for the individual lags fluctuate between positive and

---

6. Data on total expenditures, consumption of fixed capital, and interest payments are from NIPA table 3.2 (downloaded February 17, 2008). Because the BEA does not have data on "net purchases of nonproduced assets" (which are normally a trivial component of total expenditures) until 1959Q3, before then we estimate total gross expenditure less interest as the sum of current expenditure, gross government investment, and capital transfer payments, minus interest payments.

Table 1. Estimated Impact of Tax Changes on Total Expenditure[a]

| Variable	Coefficient
Constant	0.72 (0.25)
Tax change:	
Lag 0	0.24 (0.85)
Lag 1	0.40 (0.85)
Lag 2	−0.11 (0.85)
Lag 3	−0.28 (0.83)
Lag 4	−0.92 (0.87)
Lag 5	−1.50 (0.87)
Lag 6	0.31 (0.87)
Lag 7	−1.42 (0.75)
Lag 8	2.63 (0.75)
Lag 9	2.52 (0.75)
Lag 10	−0.98 (0.75)
Lag 11	−1.53 (0.74)
Lag 12	−2.19 (0.76)
Lag 13	−2.13 (0.76)
Lag 14	−1.11 (0.76)
Lag 15	0.47 (0.76)
Lag 16	0.02 (0.76)
Lag 17	−0.11 (0.74)
Lag 18	0.51 (0.78)
Lag 19	0.86 (0.78)
Lag 20	0.20 (0.78)
R^2	0.20
Durbin-Watson statistic	1.90
Standard error of the estimate	2.72

Source: Authors' regression.
a. The table reports estimates of equation 1 in the text using data for long-run tax changes only and defining expenditure as total gross expenditure less interest payments. The sample period is 1950Q1–2007Q4. Numbers in parentheses are standard errors.

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 expenditure to a long-run tax cut of 1 percent of GDP, together with 1-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 0.6 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 20 percent of GDP in our sample, a tax cut of 1 percent of GDP is equal to about 5 percent

Figure 2. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Total Expenditure, Baseline Specification[a]

Percent

Quarters after tax change
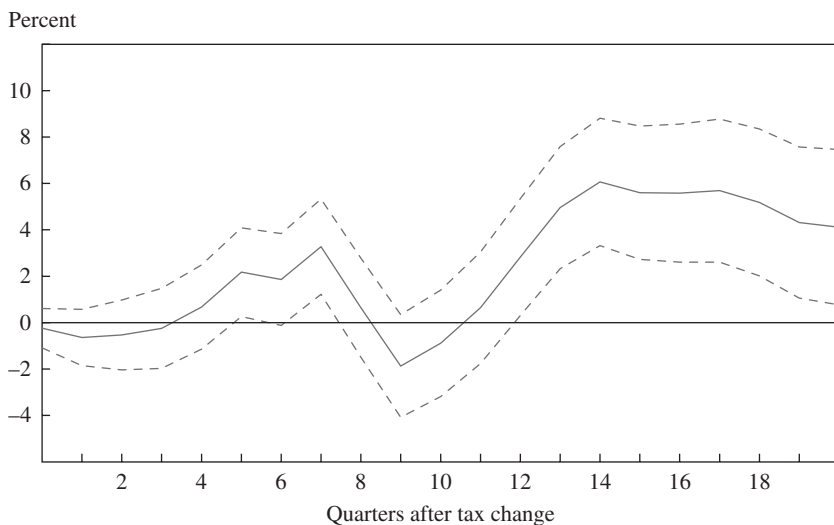
Source: Authors re gressions.
a. Based on an ordinary least squares (OLS) regression of the quarterly change in the logarithm of real total gross federal expenditure less interest payments on the contemporaneous value and 20 quarterly lags of the measure of long-run tax changes; the sample period is 1950Q1–2007Q4. Dashed lines indicate 1-standard-error bands.
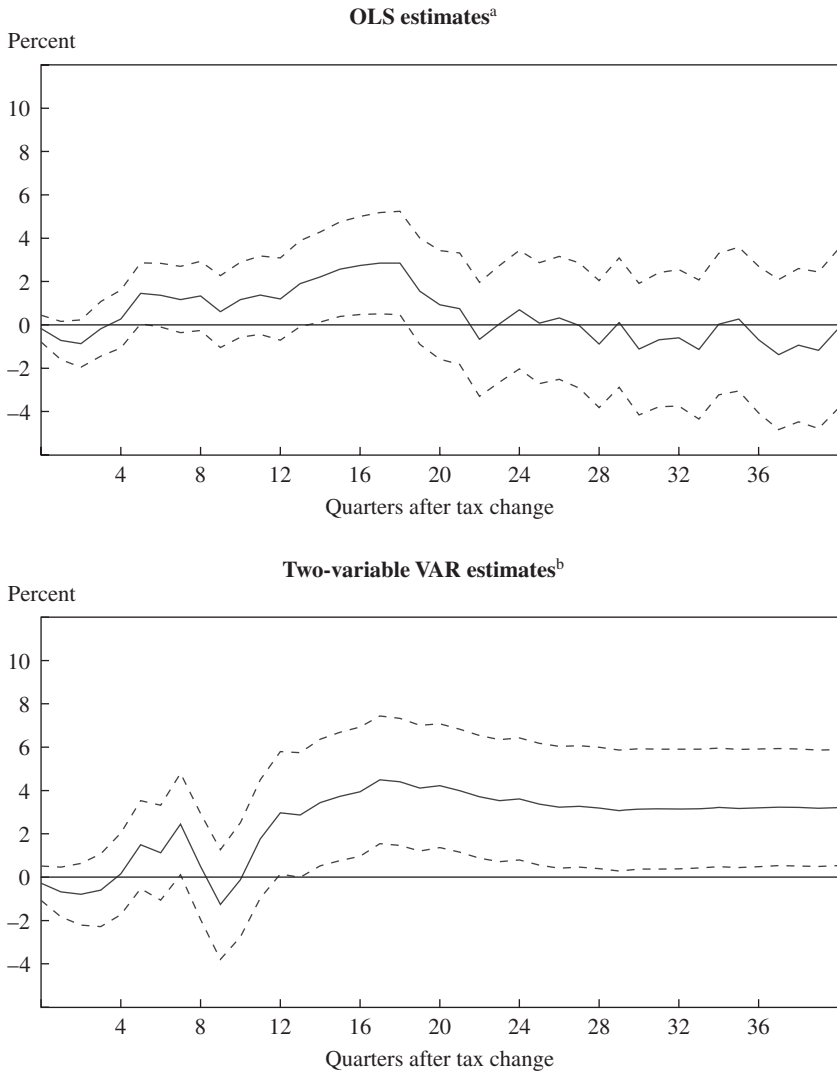
of government spending. The point estimates suggest that a tax cut of that magnitude raises spending by 4 percent or more in quarters 13 through 20. That is, they suggest that spending eventually rises by almost the amount of the tax cut. However, the estimates are not very precise. The *t* statistics for the cumulative impact of the tax cut on spending at horizons of more than three years are generally between 1.5 and 2, exceeding 2 for only one horizon (quarter 14, for which the *t* statistic is 2.21).

II.C. Richer Dynamics

Our baseline results suggest that there is no discernable starve-the-beast effect, and some evidence of shared fiscal irresponsibility, over a five-year horizon. But perhaps the main effects of tax changes occur with longer lags. Here we consider several approaches to allowing for more delayed effects.

ADDITIONAL LAGS. The most straightforward approach to examining whether tax changes have important effects at longer horizons is to include additional lags in equation 1. Of course, including more lags requires shortening the sample period and estimating additional parameters. The top panel of figure 3 shows the results of including 40 lags of the tax variable in

Figure 3. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Total Expenditure, Estimates over Longer Horizons

OLS estimates[a]

Percent

Quarters after tax change

Two-variable VAR estimates[b]

Percent

Quarters after tax change

Source: Authors' estimates.

a. Regression is specified as in figure 2 but with 40 lags of the measure of long-run tax changes; the sample period is 1955Q1–2007Q4.

b. Impulse response function from a vector autoregression (VAR) using the logarithm of total expenditure as defined in figure 2 and the measure of long-run tax changes; there are 12 lags, and the tax measure is ordered first.

equation 1 and estimating the regression over the longest feasible sample (1955Q1–2007Q4). For horizons beyond five years, the estimated cumulative impact of a tax cut of 1 percent of GDP on total expenditure is always small, fluctuates between positive and negative, and is never remotely close to statistically significant. Thus, this specification provides no evidence that tax cuts reduce government spending, but also fails to support the hypothesis that they increase it.

A TWO-VARIABLE VECTOR AUTOREGRESSION. Our second approach to allowing for more complicated and potentially longer-lasting dynamics is to estimate a VAR with our series for long-run tax changes and total expenditure. This approach allows spending to depend on its own lags as well as on the tax changes, and so allows for dynamics beyond the number of lags of the tax variable that are included.

For consistency with the earlier regressions, we put the tax changes first and expenditure second, so that tax changes can affect spending within the quarter. We enter expenditure in logarithms; given the availability of the data, we can include 12 lags while still using our baseline sample. The bottom panel of figure 3 shows that the estimated response of spending to an innovation of −1 percent of GDP to our series on long-run tax changes is similar to that for a long-run tax cut of 1 percent of GDP in the baseline specification.[7] The point 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, however. Thus, again there is no support for the starve-the-beast hypothesis. Another finding from the VAR is that the estimated response of the tax series to an innovation to government spending is very small and highly insignificant at all horizons. This indicates that the actions we classify as long-run tax changes are not responses to spending developments.[8]

7. Note that this experiment is slightly different from that considered in summarizing the results from the baseline specification. There we consider a one-time tax cut of 1 percent of GDP with no further tax changes. Here, following the innovation to our tax measure in the VAR, there are on average additional long-run tax cuts of about one-fifth of a percent of GDP over the next several years. 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.

8. We also estimated the bivariate VAR with 20 lags for the period 1952Q1–2007Q4. 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 3.97 percent after 18 quarters ($t = 1.93$).

LARGER SYSTEMS. Another way that a starve-the-beast effect could occur at longer horizons is if tax cuts affect other variables that in turn affect government spending. We therefore consider VARs with additional variables. This, however, requires either estimating more parameters in each equation or including fewer lags. Thus, rather than just include a long list of variables that might be relevant, we consider various combinations of variables.

One way that tax cuts could create pressures for reduced government spending is by increasing government debt. Thus, our first multivariable VAR uses three variables: our series on long-run tax changes, log real spending, and log real debt.[9]

We also consider two four-variable VARs. In one, we add the log of real federal total receipts as the fourth variable, so that the system includes both the spending and the revenue sides of the government budget. In the other, the fourth variable is log real GDP. Our reason for including this variable is that tax cuts have large short-run effects on output (Romer and Romer forthcoming), which could in turn affect the dynamics of spending in response to a tax cut.[10]

Finally, the nominal interest rate and inflation also affect the government budget constraint. Our last system is therefore a VAR with seven variables: our long-run tax series, log real spending, log real debt, log real revenue, log real GDP, the three-month Treasury bill rate, and the log of the price index for GDP.[11] In all of the VARs we put the tax series first, so that it can affect the other variables within the quarter. We include 12 lags and use the full sample (1950Q1–2007Q4).

---

9. From 1970Q1 to the end of the sample, we use quarterly data on the stock of federal debt held by the public. From the beginning of the sample to 1969Q4, we use the available series on gross federal debt held by the public for the second quarter of each year, and we interpolate linearly between the annual observations. Both series are taken from the St. Louis Federal Reserve Bank's FRED database, series FYGFDPUN and FYGFDPUB (www.stls.frb.org, downloaded March 24, 2008). We ratio-splice the two series in 1970Q2 and deflate the resulting series by the price index for GDP. Note that since it is likely to be the level of debt, rather than the change, that affects spending, the errors caused by the interpolation in the first part of the sample should have only minor effects on the estimates.

10. For receipts we use the federal total receipts series from NIPA table 3.2 (downloaded April 6, 2009), deflated by the price index for GDP from NIPA table 1.1.4. Our real GDP series is the quantity index for GDP from NIPA table 1.1.3 (downloaded February 17, 2008).

11. Data on the three-month Treasury bill rate are from the Board of Governors, series H15/H15/RIFSGFSM03_N.M (monthly data for secondary market rates on a discount basis, downloaded February 15, 2008).

Figure 4. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Total Expenditure, Multivariate VAR Estimates
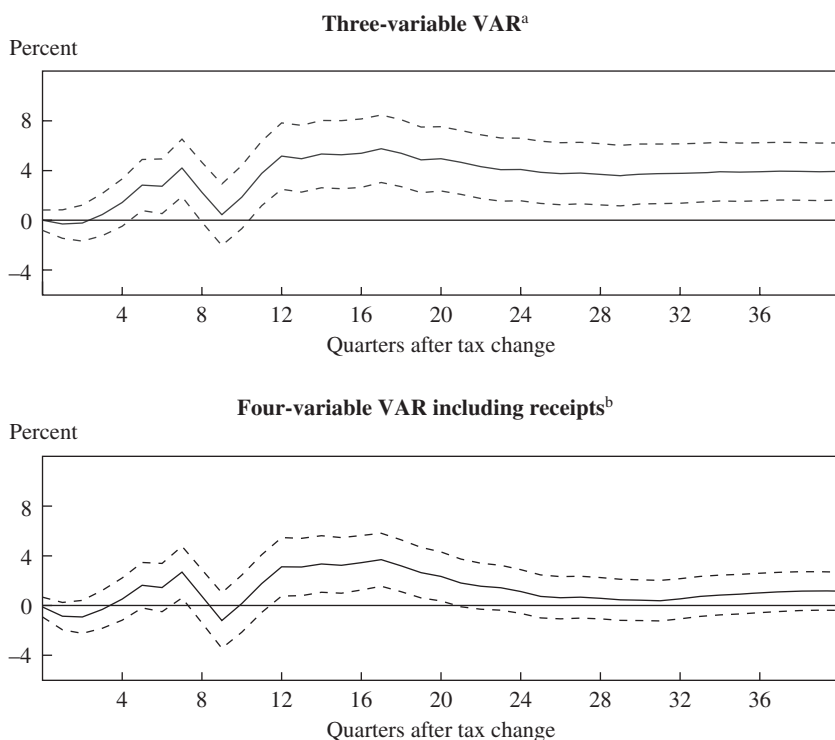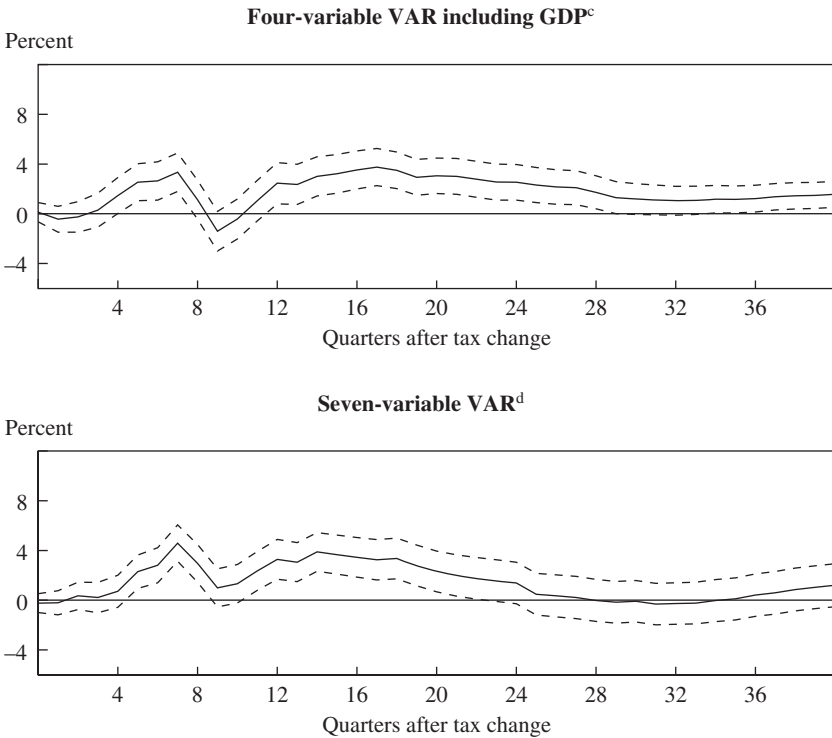
Three-variable VAR[a]

Percent

Quarters after tax change

Four-variable VAR including receipts[b]

Percent

Quarters after tax change

Figure 4 displays the response of government spending to an innovation of −1 percent of GDP to our series on long-run tax changes in each of the VARs.[12] The results consistently fail to support the starve-the-beast hypothesis. In every specification, the estimated effect of a tax cut on spending is negative at only a few horizons. And in every case, those estimates are small and insignificant: at no horizon is the *t* statistic for the spending response negative and greater than 1 in absolute value. Adding debt to the baseline VAR (first panel) in fact moves the estimates further in the direction of suggesting fiscal illusion. The estimated maximum effect of the tax cut is an increase in spending of 5.75 percent ($t = 2.12$) after 17 quarters, and the estimated effect after 10 years is an increase of 3.93 percent ($t = 1.70$). In the four-variable and seven-variable systems, the point estimates suggest a slightly weaker fiscal illusion effect, although it is more precisely

12. In each of the VARs, following the innovation to the tax series, there are modest additional long-run tax cuts over the next year that are largely offset over the following few years. There is never an important response of the tax variable to the other variables.

Figure 4. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Total Expenditure, Multivariate VAR Estimates (*Continued*)

Four-variable VAR including GDP[c]

Percent

Quarters after tax change

Seven-variable VAR[d]

Percent

Quarters after tax change

Source: Authors' estimates.

a. VAR includes the measure of long-run tax changes, the logarithm of total expenditure as defined in figure 2, and the logarithm of real federal debt held by the public. All VARs include 12 lags and order the tax measure first, and all use the full sample period (1950Q1–2007Q4).

b. VAR includes the variables in the previous VAR plus the logarithm of real federal total receipts.

c. VAR includes the variables in the first VAR plus the logarithm of real GDP.

d. VAR includes the variables in the first VAR plus the logarithm of real federal total receipts, the logarithm of real GDP, the three-month Treasury bill rate, and the logarithm of the price index for GDP.

estimated than in the two-variable VAR. In all three of those systems, the estimated maximum effect is an increase in spending of between 3.6 and 3.9 percent after about four years (except for a spike to 4.6 percent after seven quarters in the seven-variable system). In the four-variable VAR with receipts (second panel of figure 4), the effect is not significant ($t = 1.73$), but in the other two it is: the t statistic for the maximum effect is 2.51 in the four-variable VAR with GDP (third panel) and 2.49 in the seven-variable VAR (fourth panel). Finally, in all three of these specifications, the estimated effect after 10 years is in the direction predicted by fiscal illusion but is small and not significant.

II.D. Other Robustness Checks

The next step is to examine the robustness of the findings along other dimensions. The most important of these checks are summarized in figure 5, which shows the implied response of total expenditure to a long-run tax cut of 1 percent of GDP for a number of variants of the baseline regression (equation 1). For comparison, panel A of the figure repeats the baseline estimates from figure 2.

SAMPLE PERIOD AND OUTLIERS. One obvious concern is the possible importance of the sample period and of outliers. As described above, fiscal policy was very unusual in the Korean War period. Panel B of figure 5 shows that considering only the post–Korean War 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 3 through 20 is always positive, but con siderably 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 large long-run tax cuts discussed in the case studies in section IV.[13] 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.[14]
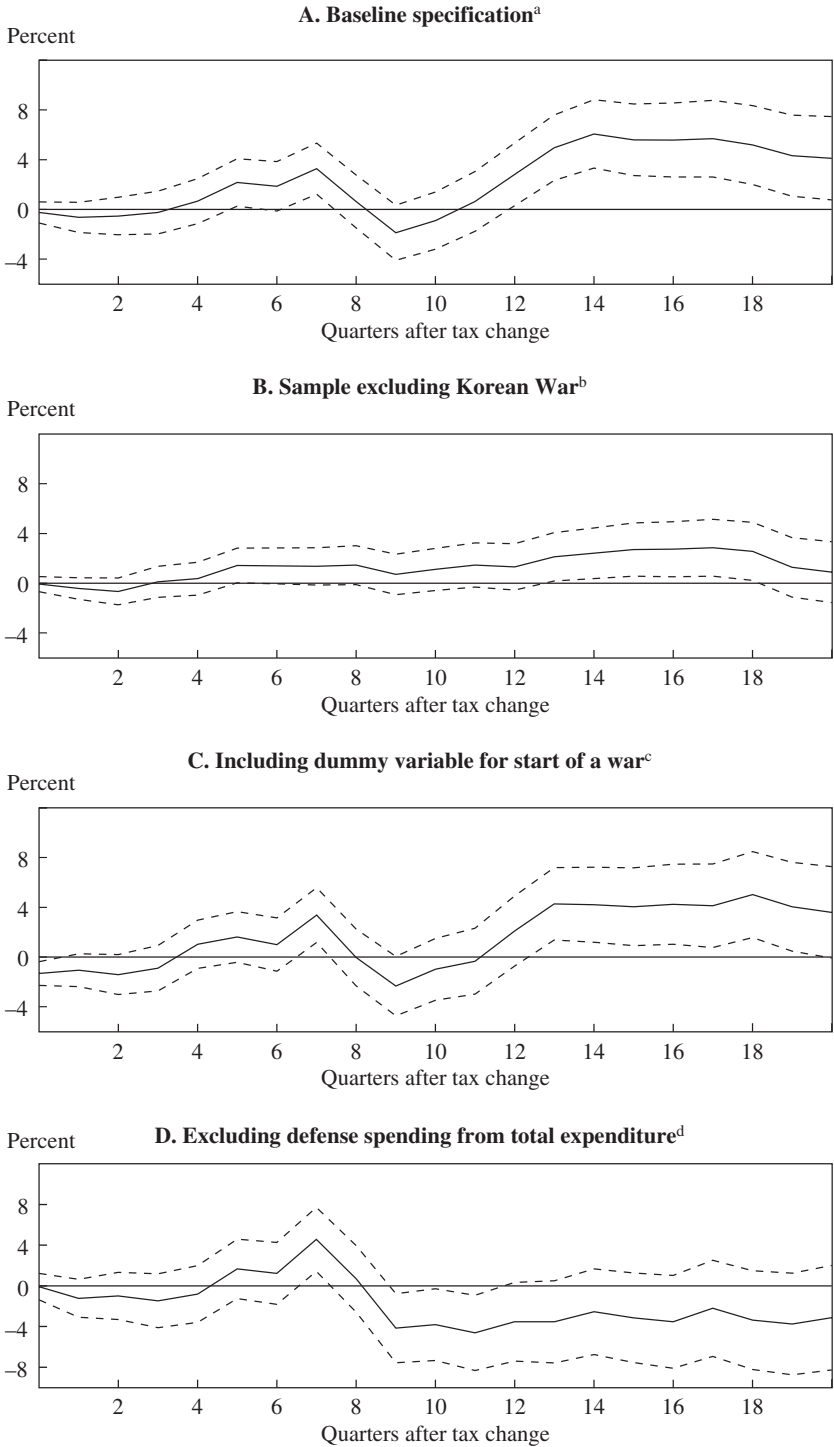
MILITARY ACTIONS. A second concern is the role of military actions in driving spending. As discussed in the case studies, many of the largest long-run tax cuts were followed by wars. The wars could have caused federal spending to rise after the tax change just by chance, thus obscuring any starve-the-beast effect. To test for this possibility, we consider two alternative specifications of our baseline regression.

The first adds an indicator of military actions to equation 1. Valerie Ramey (2008) suggests an updated list of the exogenous military actions identified by Ramey and Matthew Shapiro (1998) from narrative sources. This list dates military actions as beginning in 1950Q3 (Korean War),

---

13. 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.
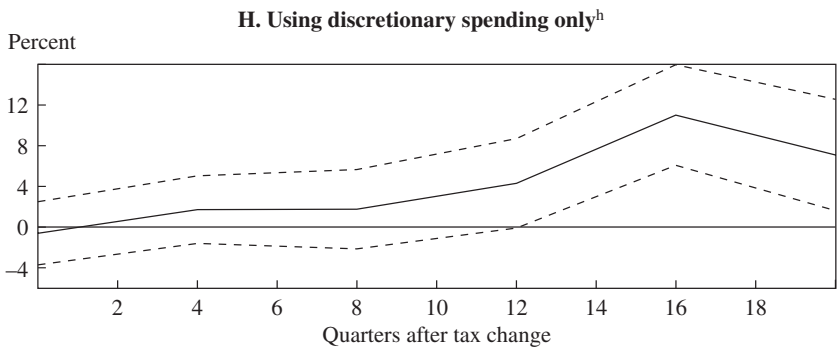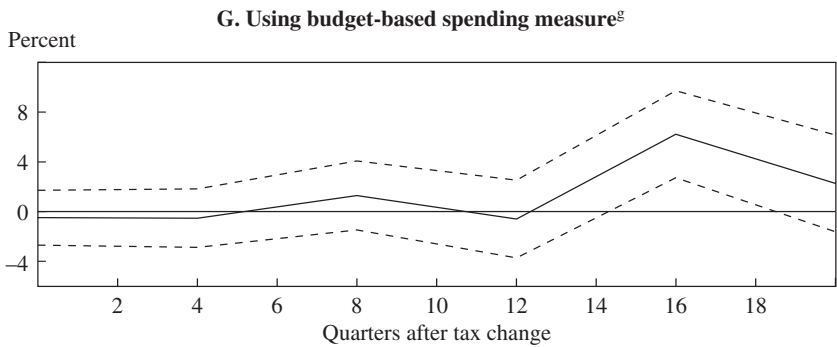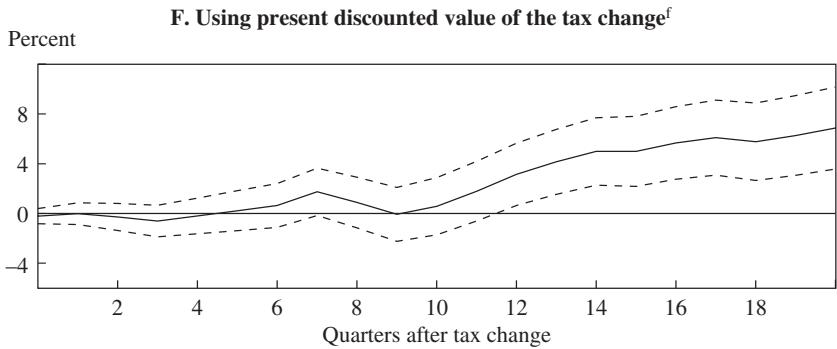
14. In a related exercise along these lines, we split the sample in 1980Q4. For the period 1950Q1–1980Q4, the estimates suggest a large and statistically significant positive effect of tax cuts on spending. For the period 1981Q1–2007Q4, the estimated effects are again virtually always positive, but consistently small and far from significant.

Figure 5. Cumulative Spending Impact of a Tax Cut of 1 Percent of GDP, Alternative Specifications

A. Baseline specification[a]

Percent

Quarters after tax change

B. Sample excluding Korean War[b]

Percent

Quarters after tax change

C. Including dummy variable for start of a war[c]

Percent

Quarters after tax change

D. Excluding defense spending from total expenditure[d]

Percent

Quarters after tax change

(*continued*)

Figure 5. Cumulative Spending Impact of a Tax Cut of 1 Percent of GDP, Alternative Specifications (*Continued*)

E. Including dummy variable for Democratic administrations[e]

Percent

Quarters after tax change

F. Using present discounted value of the tax change[f]

Percent

Quarters after tax change

G. Using budget-based spending measure[g]

Percent

Quarters after tax change

H. Using discretionary spending only[h]

Percent

Quarters after tax change
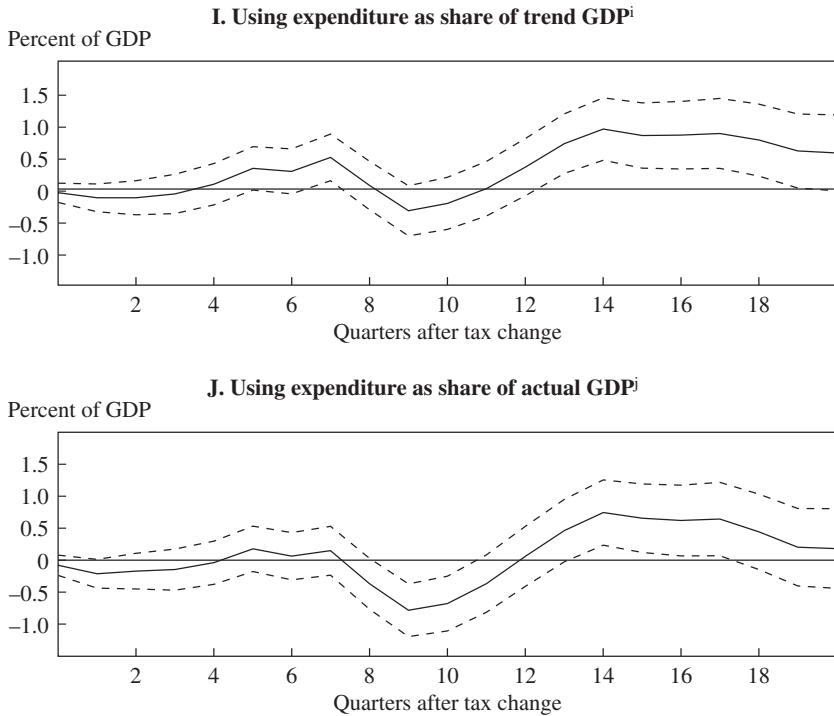
Figure 5. Cumulative Spending Impact of a Tax Cut of 1 Percent of GDP, Alternative Specifications (*Continued*)

I. Using expenditure as share of trend GDP[i]

Percent of GDP

Quarters after tax change

J. Using expenditure as share of actual GDP[j]

Percent of GDP

Quarters after tax change
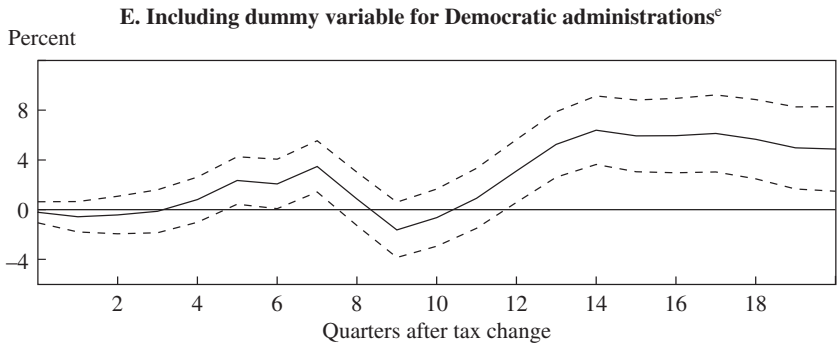
Source: Authors' regressions.

a. Repeated from figure 2. The other panels differ from this specification only as noted below.

b. Regression omits observations from the beginning of the sample period through 1956Q4.

c. Regression adds the contemporaneous value and 20 lags of a dummy variable set equal to 1 in each of the following quarters: 1950Q3, 1965Q1, 1980Q1, and 2001Q3.

d. Regression replaces the change in the logarithm of real total expenditure with the change in the logarithm of real total expenditure less national defense purchases.

e. Regression includes a dummy variable set equal to 1 in quarters when a Democrat is president.

f. Regression replaces the measure of tax changes based on the quarters in which liabilities changed with the present discounted value of all revenue changes called for by a given piece of legislation, dated as occurring in the quarter it was passed.

g. Regression replaces the NIPA measure of total expenditure with official budget data.

h. Regression replaces the NIPA total expenditure measure with discretionary spending only, from official budget data.

i. Regression uses as the expenditure measure the change in the ratio of NIPA real total expenditure to trend real GDP, calculated by fitting a Hodrick-Prescott filter ($\lambda = 1600$) to real GDP for the full sample period (1947Q1–2007Q4).

j. Regression uses as the expenditure measure the change in the ratio of real total expenditure to actual real GDP.

1965Q1 (Vietnam War), 1980Q1 (the Carter-Reagan military buildup in response to the Soviet invasion of Afghanistan), and 2001Q3 (the wars in Afghanistan and Iraq following the September 11 terrorist attacks). We expand the baseline regression to include the contemporaneous value and 20 lags of a dummy variable set equal to 1 in each of these four quarters. This specification shows the effect of tax cuts on total expenditure allowing for the possibility that wars have a separate effect on spending.

Panel C of figure 5 shows the cumulative impact of a tax cut of 1 percent of GDP in this specification. The estimates are very similar to those in the baseline specification. The effect of tax cuts on total spending controlling for military actions is largely positive, although not statistically significant. Thus, accounting for military actions does not reveal a starve-the-beast relationship. This is true even though wars exert a strong independent upward force on spending: the maximum cumulative impact of a military action on total expenditure is an increase of 15.83 percent ($t = 2.77$). The lack of a relationship between taxes and spending in this alternative specification is equally strong in the post-1957 sample.

The second alternative specification looks only at the response of non-defense spending to long-run tax cuts. In place of the log difference in total federal expenditure in equation 1, we use the log difference in total expenditure less national defense purchases (from NIPA table 3.9.5, downloaded March 25, 2008), deflated by the price index for GDP. This test almost surely biases the results in favor of the starve-the-beast hypothesis, for two reasons. First, the case studies show some correlation in our sample between support for tax cuts and support for shifting spending toward defense. Most notably, Ronald Reagan, who presided over the largest long-run tax cut in the postwar period, strongly advocated such a reallocation. Thus, nondefense spending could fall in the wake of long-run tax cuts not because of the effects of the cuts themselves but because of other actions. Second, to the degree that defense spending rises following a tax cut because of war, nondefense spending may decline for the same reason. Wartime tends naturally to lead policymakers to reallocate spending away from other purposes and toward defense. Therefore, chance correlation between wars and long-run tax cuts could cause the regression to find a starve-the-beast effect for nondefense spending when none exists.

Panel D of figure 5 shows the results of this exercise. (Note that the vertical scale differs from that in most of the other panels.) The point estimates are now generally negative, consistent with the starve-the-beast hypothesis. The effects are not statistically significant, however: the *t* statistics for the cumulative impact are almost always less than −1 and never greater than −1.3.

More important, the estimates are small and not robust. Total expenditure less defense accounts, on average, for about 10 percent of GDP over our sample. Therefore, for a tax cut of 1 percent of GDP to reduce nondefense spending by the same amount, spending would need to decline by roughly 10 percent. The estimated effect, however, is almost always a fall of less than 4 percent (or a rise). And dropping the Reagan tax cut (where, as described above, an important omitted factor seems to have acted directly to reduce nondefense spending) yields estimates that fluctuate irregularly around zero; similarly, either excluding the Korean War period or including the contemporaneous value and 20 lags of the dummy variable for military actions weakens the estimated effect considerably. Thus, there is little evidence that tax cuts have a noticeable negative effect even on nondefense spending.

POLITICAL VARIABLES. A third robustness issue concerns the role of political variables. It is certainly possible that the party of the president or the existence of unified government (that is, the same party controlling both houses of Congress and the presidency) has an influence on government spending. If such variables are 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 5 shows the effect of a tax cut on spending when a dummy variable for Democratic administrations is 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 has very little effect on the estimates, although it strengthens the evidence for fiscal illusion or shared fiscal irresponsibility slightly: both the estimated positive effects of tax cuts on spending and their statistical significance increase modestly. We also consider specifications including a dummy variable for unified government, and including separate dummies for the first quarter of a new Republican or a new Democratic administration.[15] Both specifications change the estimates only trivially, and neither provides support for the starve-the-beast hypothesis.

ALTERNATIVE TAX VARIABLE. A fourth concern involves the specification of our tax variable. Our baseline series dates revenue changes in the quarter in which liabilities actually change. An alternative measure, which emphasizes expectational effects, calculates the present discounted value of all

---

15. For the latter specification, we include both the contemporaneous value and 15 lags of the new Republican and new Democratic dummy variables.

revenue changes called for by a given piece of legislation and dates the revenue change in the quarter the law was passed.[16] Panel F of figure 5 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 generally in the opposite direction from the prediction of the hypothesis, often large, and sometimes marginally significant.

ALTERNATIVE SPENDING CONCEPTS. Our baseline specification uses a NIPA measure of total spending on the grounds that it is available quarterly and is likely to correspond most closely with economic concepts of government spending. A natural alternative is to use the official budget numbers, which may be more closely tied to policymakers' intentions. To do this, we aggregate our quarterly measure of long-run tax changes to construct a fiscal-year measure, and then reestimate equation 1 using the change in the logarithm of the budget-based real expenditure measure and the contemporaneous value and five annual lags of our tax measure.

For there to be a substantial starve-the-beast effect, tax cuts would almost certainly have to reduce not just discretionary spending, but also spending on entitlement programs. At the same time, because policymakers can change discretionary spending more quickly, it is interesting to ask whether there is a starve-the-beast effect for this type of spending. We therefore also examine the response of discretionary spending to long-run tax cuts, again using annual budget data and five annual lags of our tax measure.[17]

Panels G and H of figure 5 show the results. Once again, there is no support for the starve-the-beast hypothesis. The response of overall spending using the official budget measure (panel G) is quite similar to that using the NIPA measure in panel A. And discretionary spending (panel H, again on a different scale) rises even more than overall spending follow-

---

16. See Romer and Romer (2009) for a detailed description of how we calculate the present value of revenue changes.

17. The budget data are from *Budget of the United States Government: Historical Tables Fiscal Year 2009* (www.gpoaccess.gov/usbudget/fy09/hist.html, tables 3.1 and 8.1, downloaded March 16, 2009). We measure overall spending as total federal spending minus net interest. Discretionary spending figures are available only beginning in 1962. For the years up through 1962, we estimate the growth rate of discretionary spending as the change in the log of total spending minus the sum of Social Security, income security, veterans benefits and services, agriculture, commerce and housing credit, net interest, and undistributed offsetting receipts. The estimates constructed in this way track the official estimates for the years immediately after 1962 quite well. In aggregating our measure of long-run tax changes to fiscal-year values, we omit the transition quarter (1976Q3). We deflate both the overall spending measure and the discretionary measure by the price index for GDP.

ing a tax cut, with a maximum increase of 11.01 percent after four years ($t = 2.23$).

ALTERNATIVE SPECIFICATIONS OF THE SPENDING VARIABLE. A final robustness issue involves the appropriate way to enter the spending variable. In all of the specifications discussed so far, we examine the response of the growth rate of real government expenditure to long-run tax changes. The cumulative impact therefore shows the effect of a tax change on the level of real expenditure. 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 reduces expenditure as a percent of GDP. In this view, a tax cut could lower the share of spending in GDP not by changing policymakers' spending decisions, but by changing output growth.

To test this alternative version, we reestimate equation 1 using two different specifications of the dependent variable. The more sensible of the two expresses real total expenditure 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.[18] Detrending real GDP is reasonable because, to the extent that a tax cut causes a temporary boom, it will inherently tend to reduce real expenditure 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 reestimate equation 1 using the change in the ratio of total real expenditure to actual real GDP.

Panels I and J of figure 5 show the results of these two exercises. (These two panels are on a different scale than the others in figure 5 because the dependent variable is now a percent of GDP, not a percent of total expenditure.) Panel I 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 1 percent of GDP generally raises the share of spending in GDP. The estimated maximum effect is large (0.94 percent of GDP) but only marginally significant ($t = 1.92$). 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.

---

18. We again calculate real expenditure by dividing nominal expenditure by the price index for GDP. Real GDP is constructed by dividing nominal GDP by the same price index. We fit a Hodrick-Prescott filter ($\lambda = 1600$) to log real GDP for the full sample (1947Q1–2007Q4).

Panel J 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 (forthcoming), 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 expenditure as a share of GDP.

The robustness checks in this section yield two conclusions. First, and more important, the lack of support for the starve-the-beast hypothesis is very robust: with the possible exception of the examination of nondefense spending, which appears to be biased in favor of the starve-the-beast hypothesis and for which the results are mixed, none of the specifications we consider provide evidence that tax cuts reduce government expenditure. Second, although we find evidence for the alternative view of fiscal illusion or shared fiscal irresponsibility, it is only modest. The point estimates consistently suggest that tax cuts raise government expenditure, but they are only occasionally significantly different from zero, and then usually only marginally so.

II.E. *The Relationship between Other Types of Tax Changes and Total Expenditure*

As discussed above, we focus on the response of government spending to long-run tax changes because this is likely to provide the least biased 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. This analysis can reveal whether 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 be positively correlated with spending changes.

For this exercise we reestimate equation 1 using the contemporaneous value and 20 lags of a particular type of tax change as the independent variable. We estimate a separate regression for each type of tax change, using data from the full postwar sample period. The results are again summarized by calculating the implied cumulative response of spending to a tax cut (of a given type) of 1 percent of GDP. Figure 6 presents the results for each type

of tax action.[19] To facilitate comparisons, the first panel repeats our base-line results for long-run tax actions from figure 2.
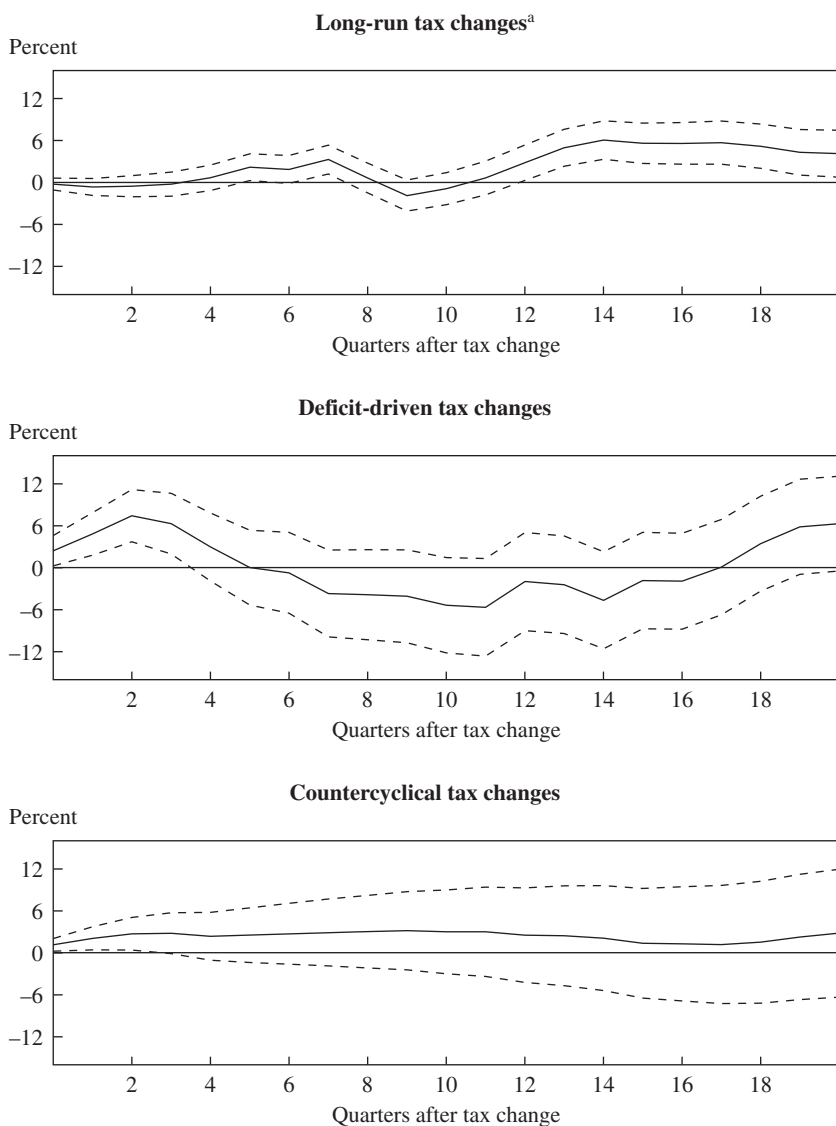
DEFICIT-DRIVEN TAX CHANGES. Of the three additional types of tax changes, those driven by deficits are likely to be the most informative about the starve-the-beast hypothesis. Like the long-run changes, these actions are not taken in response to current or prospective short-run macroeconomic conditions 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 parts of deficit reduction packages that include spending reductions. These observations might therefore bias the results against the starve-the-beast hypothesis. The estimated impact of deficit-driven tax changes on total expenditure (second panel 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 put it in terms of the realistic case, following a deficit-driven tax increase, spending falls sub-stantially. This is exactly the sort of inverse relationship one would expect if deficit reduction packages were common. The effects, although large, are not precisely estimated. The *t* statistic for the maximum impact is 1.98.

After the first few quarters, the estimated effects of a deficit-driven tax cut turn negative for several years but return to 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. The lack of a consistent pattern to the estimates at longer horizons 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 do not support the starve-the-beast hypothesis.
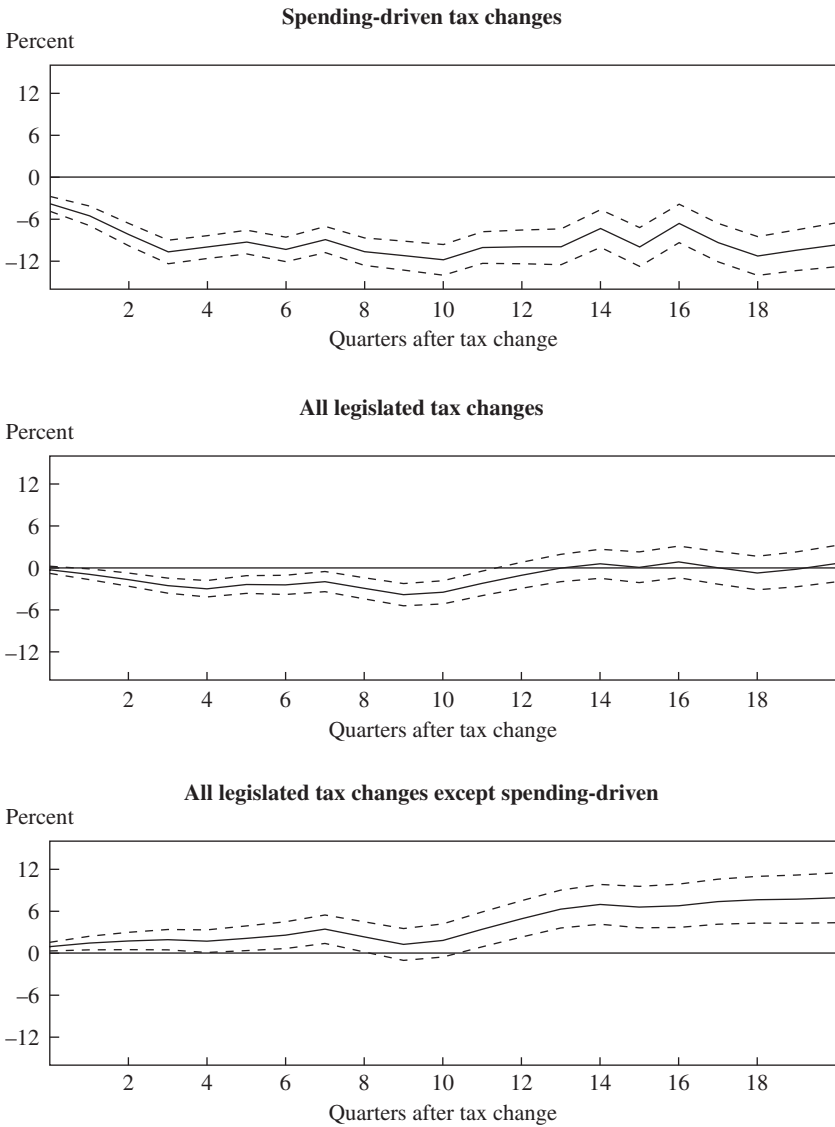
COUNTERCYCLICAL TAX CHANGES. The third panel of figure 6 shows the implied impact on spending of a countercyclical tax cut. We exclude such tax changes from our baseline regression because the state of the economy could tend to influence spending and taxes in opposite directions, and so again bias the 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.

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 interpre-tation 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.

Figure 6. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Total Expenditure, by Type of Tax Cut

Long-run tax changes[a]

Percent

Quarters after tax change

Deficit-driven tax changes

Percent

Quarters after tax change

Countercyclical tax changes

Percent

Quarters after tax change

Figure 6. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Total Expenditure, by Type of Tax Cut (*Continued*)

Spending-driven tax changes

Percent

Quarters after tax change

All legislated tax changes

Percent

Quarters after tax change

All legislated tax changes except spending-driven

Percent

Quarters after tax change

Source: Authors' regressions.
a. Repeated from figure 2.

SPENDING-DRIVEN TAX CHANGES. The fourth panel of figure 6 shows the behavior of government spending following a spending-driven tax cut. In this case the relationship is negative, large in absolute terms, and highly statistically significant.[20] This is exactly the result one would expect: if we have classified spending-driven tax changes correctly, there should be a positive correlation between them and spending. That the relationship persists is consistent with the spending changes associated with these spending-driven actions being permanent. The findings for spending-driven tax changes both confirm our classification 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.

ALL LEGISLATED TAX CHANGES. One way to see how much bias would result from including spending-driven tax changes in our analysis is to define a tax variable that sums all four types of legislated tax changes and then use this as the explanatory variable in equation 1. The fifth panel of figure 6 shows the implied impact on total expenditure of a legislated tax cut of any motivation of 1 percent of GDP. The estimated response is strongly negative, and often statistically significant, for the first three years after a tax cut. The point estimate for the maximum cumulative effect is -3.82 percent ($t = -2.41$). Since none of the other types of tax changes show a consistent negative response, this implied negative effect of the aggregate tax variable must reflect 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. The last panel of figure 6 shows the cumulative response of total expenditure to a non-spending-driven legislated tax cut of 1 percent of GDP. The effects are consistently positive, suggesting that, if anything, tax cuts appear to be followed by increases in government spending, not decreases as the starve-the-beast hypothesis predicts. And, for horizons beyond three years, these positive effects are significantly different from zero.

THE CHANGE IN CYCLICALLY ADJUSTED REVENUE. These results suggest that the inclusion of spending-driven tax changes in the sample may explain why much of the previous literature has found evidence for the starve-the-beast hypothesis. This possibility can be investigated further by considering a

20. These findings are somewhat sensitive to the sample period. Some of the largest spending-driven tax changes occurred during the Korean War. When the post-1957 sample period is used, the maximum impact of a spending-driven tax cut of 1 percent of GDP is large (-6.65 percent) but not statistically significant ($t = -1.60$).

Figure 7. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Total Expenditure, Estimates Using Cyclically Adjusted Revenue[a]

All changes in cyclically adjusted revenue

All changes in cyclically adjusted revenue except spending-driven
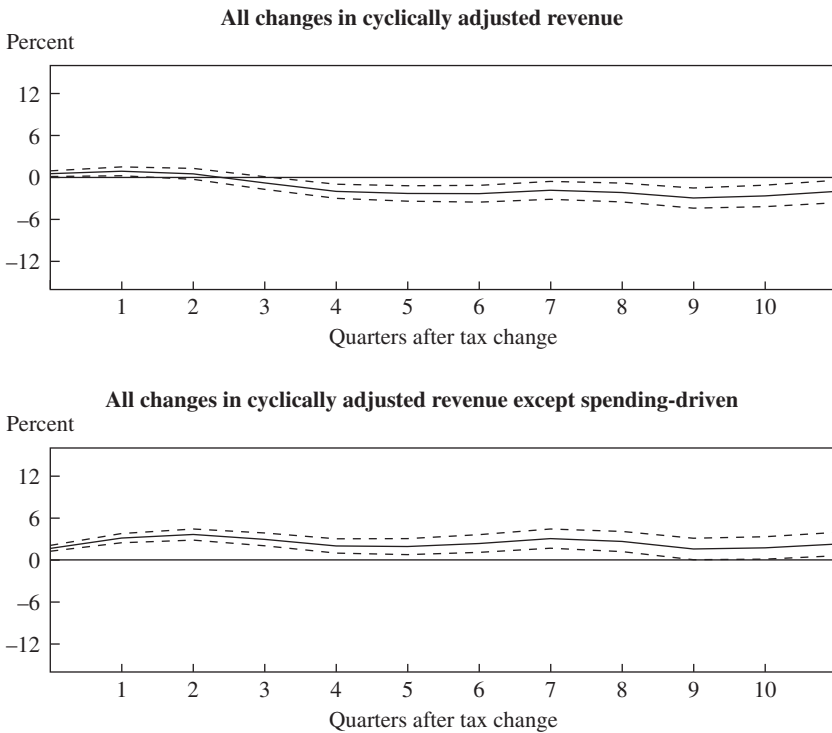
Source: Authors' regressions.

a. Based on regressions using revenue changes measured as the change in real cyclically adjusted revenue as a percent of real GDP; the contemporaneous value and 11 lags are entered for the period 1950Q1–2007Q4.

more standard measure of tax changes. A typical test of the starve-the-beast hypothesis uses the change in cyclically adjusted revenue, which includes all changes in revenue not related to short-run fluctuations in income, as the measure of tax changes. Data on the change in cyclically adjusted revenue are available beginning in 1947Q2. We therefore investigate the effects of using the contemporaneous value and 11 lags of this variable as the tax measure for the period 1950Q1–2007Q4.[21] When we use this conventional tax variable, the results indeed seem to support the starve-the-beast hypothesis. The top panel of figure 7 shows that the estimated cumulative effect

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

of a decline in real cyclically adjusted revenue of 1 percent of GDP starts out positive but then turns negative. The maximum impact is a change in government expenditure of −2.94 percent ($t = −2.04$).

If spending-driven tax changes are driving this result, subtracting these changes from the change in cyclically adjusted revenue should cause the effect to disappear.[22] Indeed, the results using such a series (bottom panel of figure 7) are dramatically different from those using the total change in cyclically adjusted revenue. The estimated impact of a 1-percent-of-GDP decline in cyclically adjusted revenue less spending-driven changes is strongly positive in the short run: the maximum impact is 3.63 percent ($t = 4.56$). It then gradually declines toward zero, but it never turns negative over the 11-quarter horizon we consider. Thus, the results provide no support for the starve-the-beast hypothesis and, indeed, are somewhat supportive of shared fiscal irresponsibility. This supports the view 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 revenue.[23]

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 naturally raises another question: how then does the government budget adjust to the cuts? 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 revenue and tax legislation to long-run tax changes.[24]

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

23. The importance of spending-driven tax changes in biasing the results toward finding a starve-the-beast effect 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, whereas those such as von Furstenberg, Green, and Jeong (1986), which use data starting in 1954, do not.

24. Bohn (1991) also examines the degree to which deficits caused by falls in revenue are eliminated by subsequent tax increases. But because he does not account for the sources of changes in revenue, his estimates may suffer from important omitted variable bias. This is particularly true because many of the most important revenue changes in his sample are associated with wars.
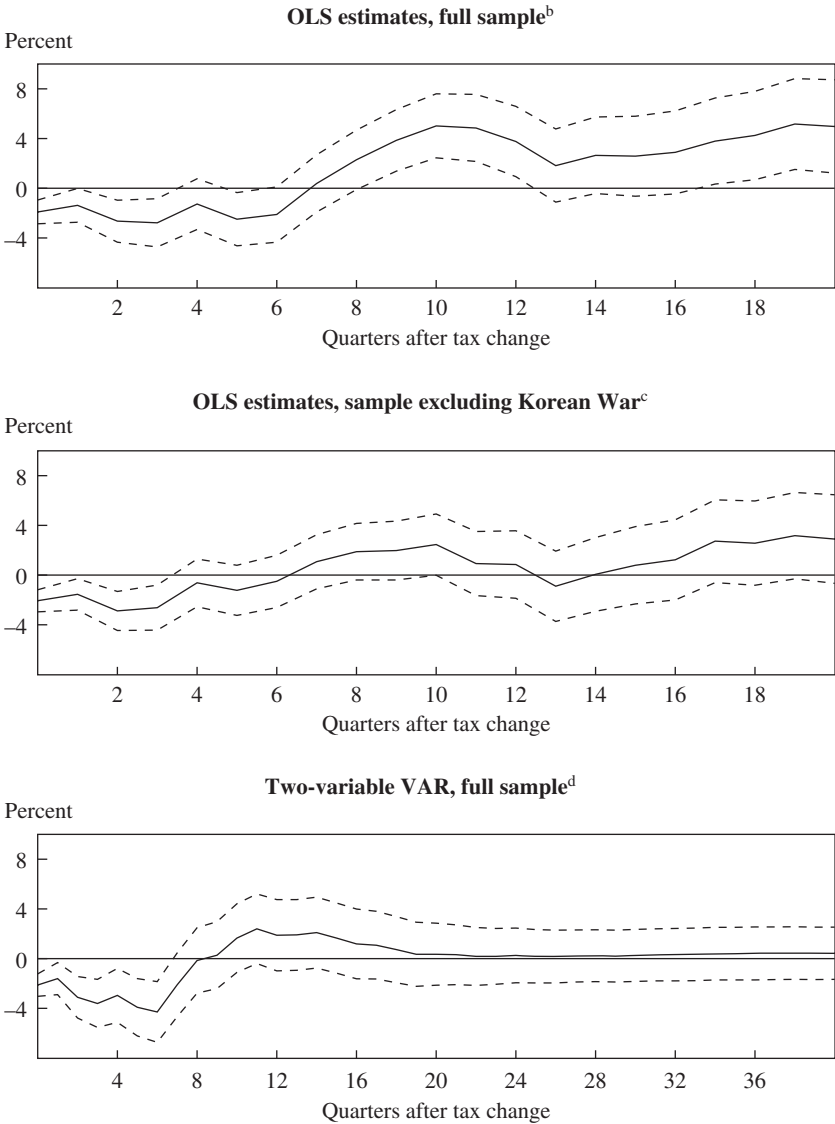
III.A. Response of Tax Revenue

To investigate how revenue responds to long-run tax changes, we first reestimate equation 1 using a measure of the change in real tax revenue as the dependent variable. That is, we regress the percentage change in real revenue on a constant and on the contemporaneous value and 20 lags of our measure of long-run tax actions. As in the VARs in section II, we measure revenue using NIPA federal total receipts, deflated by the price index for GDP. We estimate the revenue response over both the full postwar sample period (1950Q1–2007Q4) and the post–Korean War sample (1957Q1–2007Q4).

The top and middle panels of figure 8 show the implied cumulative response of total receipts to a long-run tax cut of 1 percent of GDP in each sample period. Tax receipts decline strongly in the short run in response to a tax cut. The contemporaneous effect is a change in receipts of −1.90 percent in the full sample ($t = -2.00$) and −2.06 percent in the post–Korean War sample ($t = -2.33$). 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 revenue 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. Three major tax increases were passed during 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–Korean War 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 statistic for the positive effect does not rise above 1).

To further investigate the response of receipts to tax shocks, we also estimate a bivariate VAR using our measure of long-run tax changes and the log of real total receipts. We include 12 lags of each series, which allows us to use our baseline sample period of 1950Q1–2007Q4. The bottom panel of figure 8 shows the response of real receipts to a long-run tax cut of 1 percent of GDP in this specification. Receipts fall markedly following a long-run tax cut, and the effects are significant, or nearly so, for the first year and a half. Receipts then turn positive nine quarters after the shock. However, even though this specification uses the full sample,

Figure 8. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Total Receipts[a]

OLS estimates, full sample[b]

Percent

Quarters after tax change

OLS estimates, sample excluding Korean War[c]

Percent

Quarters after tax change

Two-variable VAR, full sample[d]

Percent

Quarters after tax change

Source: Authors' estimates.

a. Based on regressions of the percentage change in real revenue (NIPA federal total receipts deflated by the price index for GDP) on a constant and the contemporaneous value and 20 lags of the measure of long-run tax changes.

b. Regression is estimated for the full postwar sample period (1950Q1–2007Q4).

c. Regression is estimated for the post–Korean War sample period (1957Q1–2007Q4).

d. Impulse response function from a VAR using the logarithm of real total receipts and the measure of long-run tax changes estimated for the full postwar sample period; there are 12 lags, and the tax measure is ordered first.

the positive effects are extremely small in absolute terms and not statistically significant.[25]
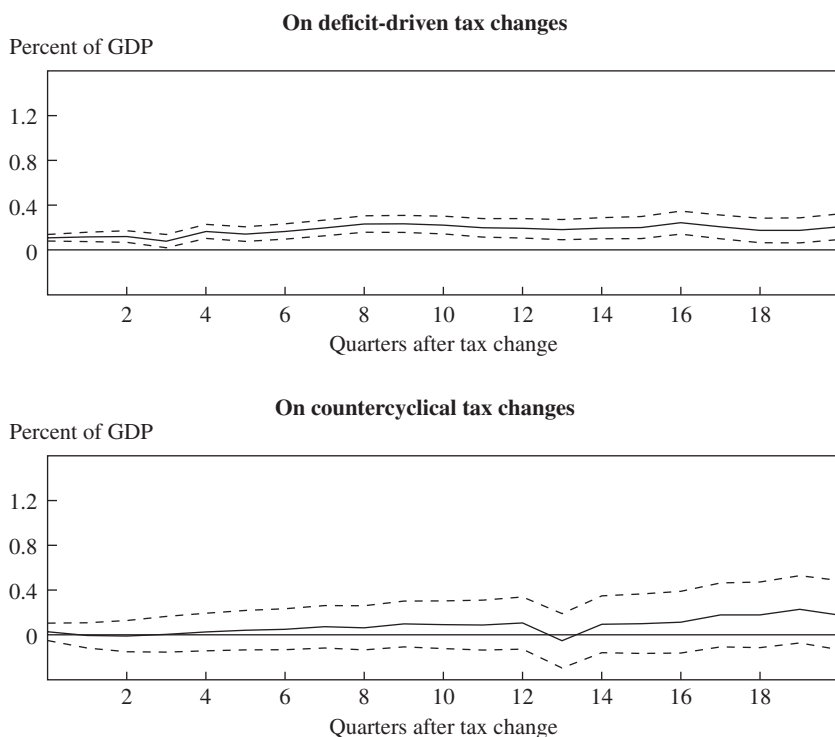
III.B. Response of Tax Legislation

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

In our single-equation analyses of spending and revenue, we consider the experiment of a tax cut intended to spur long-run growth 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 this 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-you-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 basic 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 on the contemporaneous and lagged values of our measure of long-run tax changes. In our baseline specification we again use 20 lags, but we also

---

25. The response of total receipts to a long-run tax cut is even more negative when the bivariate VAR includes 20 lags of each variable and is estimated over the shorter sample period 1952Q2–2007Q4. For this specification, tax revenue does not turn consistently positive until four years after the tax cut. The results for the behavior of revenue using the multivariate VARs described in section II are broadly similar to those from the bivariate VAR. For example, in the four-variable VAR that includes our measure of long-run tax changes, government expenditure, debt, and tax receipts, the effect of a long-run tax change of 1 percent of GDP on receipts is negative for the contemporaneous quarter and the six quarters after the shock and then turns positive. The positive effects are somewhat larger than in the bivariate VAR, but still small in absolute terms and not significant.
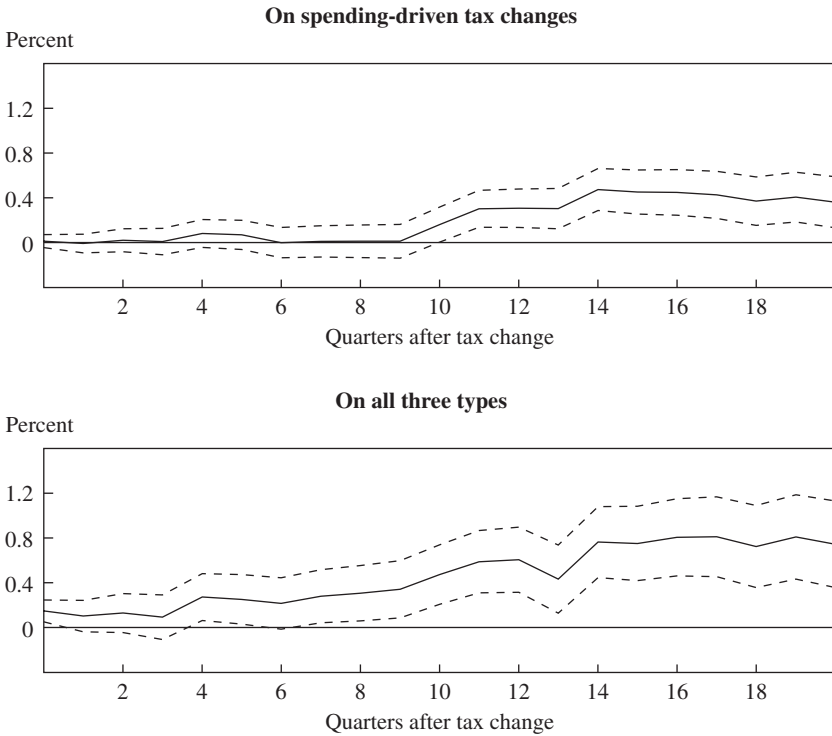
Figure 9. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Subsequent Tax Changes, by Type[a]

On deficit-driven tax changes

Percent of GDP

Quarters after tax change

On countercyclical tax changes

Percent of GDP

Quarters after tax change

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 1 percent of GDP. A positive impact implies that subsequent tax actions counteracted the long-run tax cut. Because the other tax variables are also expressed as a percent of nominal GDP, the cumulative impact can be interpreted as the fraction of the long-run tax cut that is undone over the horizon considered.

Figure 9 shows the estimated impacts of a long-run tax cut of 1 percent of GDP on tax changes of various types. The first panel shows that the impact on deficit-driven tax actions 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 0.23 percentage point ($t = 3.06$) after

Figure 9. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Subsequent Tax Changes, by Type[a] (*Continued*)

On spending-driven tax changes

Percent

Quarters after tax change

On all three types

Percent

Quarters after tax change

Source: Authors' regressions

a. Based on regressions of the measure of legislated tax changes of the indicated type on a constant and the contemporaneous and 20 lagged values of the measure of long-run tax changes. Estimates are for the full postwar sample period (1950Q1–2007Q4).

8 quarters and 0.24 percentage point ($t = 2.39$) after 16 quarters.[26] 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

26. The contemporaneous impact is substantial (0.11 percentage point; $t = 3.73$). 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 that quarter. Concern about current and prospective deficits, however, led to passage of the Tax Equity and Fiscal Responsibility Act of 1982, which raised revenue mainly by modifying some features of the 1981 act that had already taken effect (Romer and Romer 2009). 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.

the sample in 1957 has virtually no impact, and increasing the number of lags to 40 and carrying out the simulations for 10 years strengthen the results. Ten years after the long-run tax action, 44 percent of the action has been undone by deficit-driven tax increases ($t = 2.53$).

The second panel in figure 9 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 20 quarters, countercyclical tax actions have counteracted 18 percent of a long-run tax cut ($t = 0.57$). Starting the sample in 1957 has virtually no impact, because there were no countercyclical tax actions in the early 1950s. Including more lags suggests that the response diminishes at longer horizons. The estimated effect after 10 years is 0.11 percentage point ($t = 0.21$).[27]

The third panel of figure 9 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 and then turn strongly positive. The maximum cumulative impact is 0.47 percentage point ($t = 2.53$) after 14 quarters. The impact after 20 quarters is 0.36 percentage point ($t = 1.58$). 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 tend to give rise to pay-as-you-go policies.

More than with the other tax changes, there is reason to be concerned that the results for spending-driven actions are influenced by the observations from the Korean War. Starting the sample in 1957 does indeed weaken the link substantially. The strongest impact of a long-run tax cut is now a rise in spending-driven taxes of 0.14 percentage point after eight quarters ($t = 2.03$). Likewise, including 40 lags reduces the impact substantially for the full sample, but this effect is due entirely to the required shortening of the sample period.

The last panel of figure 9 shows the effect of a long-run tax cut on the other three types of legislated tax changes combined. The effect is positive, large, and significant: 0.61 percentage point ($t = 2.08$) after 12 quarters, 0.81 percentage point ($t = 2.34$) after 16, and 0.74 percentage point ($t = 1.92$) after 20. This suggests that roughly three-quarters of a long-run tax cut is typically undone by legislated tax increases of various sorts within five years.

---

27. We also experiment with leaving out the 1975 tax rebate, which is a huge outlier among countercyclical actions, because it 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 is to cut the standard errors by more than half.

Figure 10 reports the results of three robustness checks for the effect of a long-run tax cut on this composite of other tax changes. The top panel 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 0.60 percentage point ($t = 1.66$) after 19 quarters. The middle panel shows the effect of including 40 lags of long-run tax changes. The required shortening of the sample reduces the estimated response over the first 20 quarters somewhat. Thereafter it moves irregularly upward. The response after 40 quarters is large (0.77 percentage point) but not precisely estimated ($t = 1.39$). Although they weaken the evidence slightly, these two robustness checks tend to confirm that a large fraction of a long-run tax cut is typically reversed by legislated tax increases within the next few years.

Our final robustness check allows for more complicated dynamics. We estimate a bivariate VAR that includes both our measure of long-run tax changes and the composite measure of the three other types of legislated tax changes. We include 12 lags of each series and estimate the VAR over our baseline sample period of 1950Q1–2007Q4.[28]

The bottom panel of figure 10 shows the response of other legislated tax changes to a long-run tax cut of 1 percent of GDP in this specification. The results are again very similar to those from the single-equation specification. The response of other tax changes is strongly positive: the maximum effect is 0.78 percentage point ($t = 2.22$) 18 quarters after the shock. The effect diminishes slightly thereafter but levels off at around 0.65 percentage point. Thus, the VAR specification confirms that long-run tax cuts tend to be substantially counteracted by other types of tax increases over the next several years.
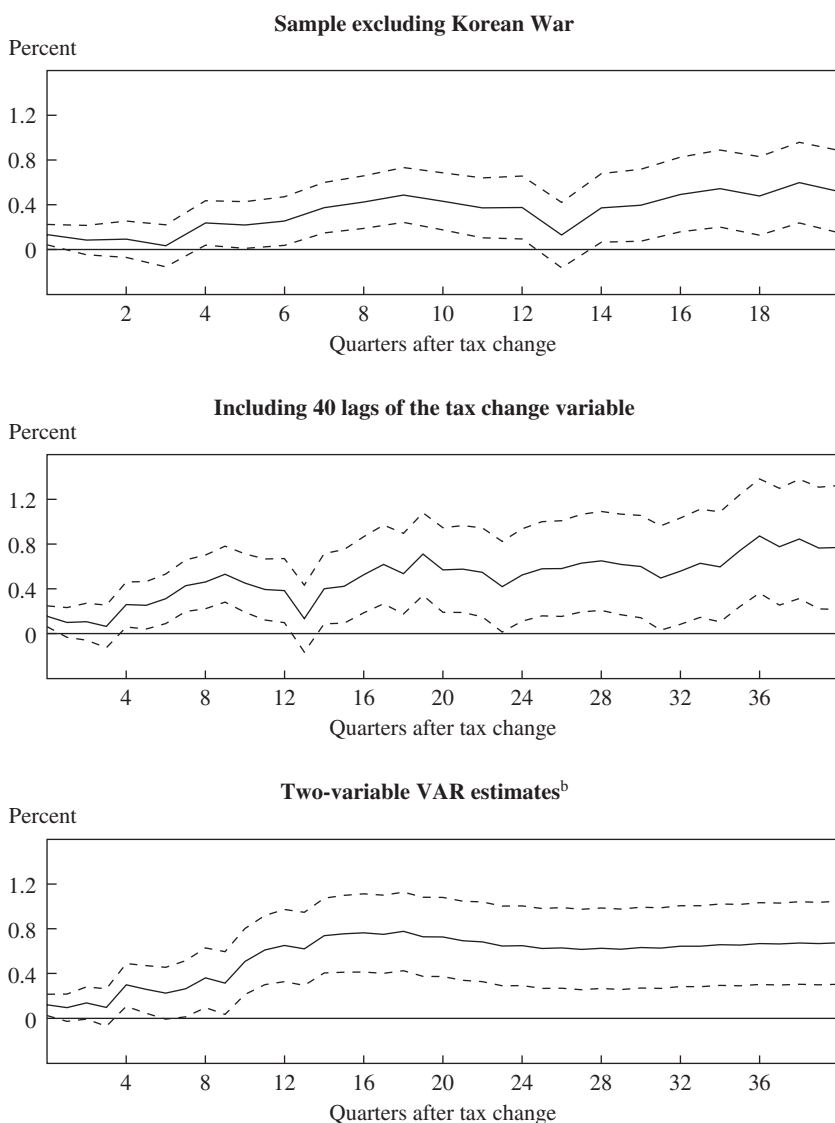
III.C. Discussion

The fact that policymakers have been able to largely reverse tax cuts helps to explain why the cuts have not reduced spending.[29] To see this connection, note that a tax cut could reduce future spending in either of two ways. The first is through debt: by bequeathing greater debt to future policymakers,

28. The experiment we can consider in this framework is again slightly different from that in the single-equation specification. When we look at the effect of an innovation to long-run tax changes in the VAR specification, we are no longer assuming that the tax change is not followed by other long-run tax changes. Rather, we let the data say how long-run tax changes respond to the innovation. The cumulative response of long-run tax changes to a long-run tax cut of 1 percent of GDP levels off at around −1.2 percentage points. This suggests that a long-run tax change is typically followed by subsequent long-run tax changes in the same direction. This is consistent with the fact that many long-run tax changes are legislated to take effect in a series of steps.

29. We are grateful to our discussant Steven Davis for this point.

Figure 10. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Subsequent Tax Changes, Alternative Specifications[a]

Sample excluding Korean War

Percent

Quarters after tax change

Including 40 lags of the tax change variable

Percent

Quarters after tax change

Two-variable VAR estimates[b]

Percent

Quarters after tax change

Source: Authors' estimates.

a. Based on regressions of the measure of legislated tax changes of all types other than long-run tax changes on a constant and the contemporaneous and lagged values of the measure of long-run tax changes. Regressions include 20 lags of the tax variable and are estimated for the full postwar sample except where indicated otherwise.

b. Impulse response function from a VAR using the measure of long-run tax changes and the composite measure of the three other types of legislated tax changes; there are 12 lags, and the long-run tax measure is ordered first.

current policymakers restrict future policymakers' choice set, which is likely to lead to some combination of higher taxes and lower spending. This is the mechanism emphasized in standard models of strategic budget deficits (for example, Tabellini and Alesina 1990 and Persson and Svensson 1989). The second is by leaving future policymakers with less tax revenue. If increasing taxes is costly, this further reduces spending. This mechanism appears important in informal discussions of the starve-the-beast effect (see, for example, the quotation from Ronald Reagan at the beginning of the paper, which seems to suggest a permanent reduction in government revenue).

If the costs of reversing a tax cut are small relative to the costs of cutting spending, then only the first channel is relevant. And that channel is likely to be quantitatively small. Suppose, for example, that a policymaker cuts taxes by 2 percent of GDP for five years. The result will be a deficit that is larger than it otherwise would have been by about 2 percent of GDP for five years, and thus a stock of debt that is larger by about 10 percent of GDP after five years. If the difference between the real interest rate and the economy's growth rate is 2 percentage points, then the interest costs associated with maintaining the debt-to-GDP ratio at its higher level are about 0.2 percent of GDP (2 percent times 10 percent). Thus, policymakers can keep the tax cut from raising the debt-to-GDP ratio further by first undoing the tax cut and then enacting a permanent spending reduction of 0.1 percent of GDP and an additional permanent tax increase of 0.1 percent of GDP. Since spending is about 20 percent of GDP, this corresponds to a spending reduction of about 0.5 percent—a quite small starve-the-beast effect.

If, however, undoing the tax cut is difficult, the effect is much stronger. In the extreme case where none of the tax cut can be reversed, satisfying the government budget constraint requires a spending cut equal to the tax cut (or by even more if there is a delay between the tax cut and the spending reduction, so that the amount of debt increases before the spending reduction). The result is a spending reduction of about 10 percent.

Our results concerning the behavior of tax legislation following tax cuts suggest that the truth is closer to the first case than to the second. This suggests a critical reason for our failure to find a substantial starve-the-beast effect: adjustment on the tax side, although presumably not costless, appears feasible, making large adjustments on the spending side unnecessary.

We also find that the overall rebound in revenue exceeds the portion due to legislated changes. The key source of the nonlegislated change in revenue is almost certainly the effect of the tax cut on economic activity.

In Romer and Romer (forthcoming), we find that a tax cut of 1 percent of GDP increases real output by approximately 3 percent over the next three years. Since revenue is a function of income, this growth raises revenue.

There is, however, an important caveat to this finding that tax cuts partly 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. To the extent that this is the case, some of the rebound in revenue is also temporary. As a result, without 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 budget constraint adjusts following a tax cut. What we can say is 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 the largest long-run tax cuts. This examination serves several purposes. The first is to see whether the narrative record suggests that the tax cuts affected spending decisions. We examine the reasoning that policymakers gave for their spending behavior, and so check whether tax cuts appear to have had an important effect on the decisionmaking process. 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.

The second purpose is to check whether our regression results reflect consistent patterns in the data. Specifically, we look at the behavior of overall spending and its two broad components, defense purchases and nondefense spending, in each episode. This allows us to investigate whether the relationships shown by the regressions appear in the key episodes.

Our third purpose is to examine whether any omitted variables or idiosyncratic shocks account for the failure of spending to fall after a tax cut.

---

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 Woolley and Peters, *The American Presidency Project* (www.presidency.ucsb.edu).

We ask whether any unusual developments in the episodes had important impacts on spending. This analysis can suggest whether the regression results overstate (or understate) the evidence against the starve-the-beast hypothesis.

The final purpose is to address a similar set of issues concerning 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 important, 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 each episode.

IV.A. The Revenue Act of 1948

The Revenue Act of 1948 was passed over President Harry Truman's veto in April 1948. The bill was projected to reduce revenue by 1.9 percent of GDP beginning in 1948Q2. The primary motivation for the cut was a desire to improve economic efficiency by reducing marginal tax rates.[31]

The tax cut was followed by a substantial reduction in revenue. Truman's view, however, was that government spending should be determined by considerations other than the level of revenue and that tax policy should be adjusted accordingly. The 1950 *Economic Report* provides a clear statement of this belief:

> 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)

Consistent with this view, Truman's 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" (1950 *Budget,* p. M5). This increase would have offset 80 percent of the 1948 cut.

---

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 (2009). The revenue estimates exclude the effects of retroactive features of the bills.

Nonetheless, the fall in revenue appears to have had a marginal effect on Truman's spending policies. 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 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 (see, for example, his Annual Message to the Congress on the State of the Union, January 5, 1949, p. 3), this points to at least some effect of the tax cut on spending decisions.

After North Korea invaded South Korea on June 25, 1950, taxes and the deficit essentially disappeared from Truman's discussions of spending. Even more than it had been 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 of 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 sought 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. As a result, 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," August 15, 1948, p. 3). Thus, there is no evidence of a starve-the-beast effect operating through congressional actions in this episode.

The first panel of figure 11 shows the behavior of real government spending in this episode. It plots, in logarithms, both our measure of total expenditure and the two categories of spending, national defense purchases and nondefense spending. As in section II, we define nondefense spending as the difference between our measure of total expenditure and national defense purchases; the two main components of this measure are nondefense purchases and current transfer payments. The vertical line indicates the quarter in which the tax cut took effect. Several things are apparent. First and most important, there was no discernable slowdown in overall
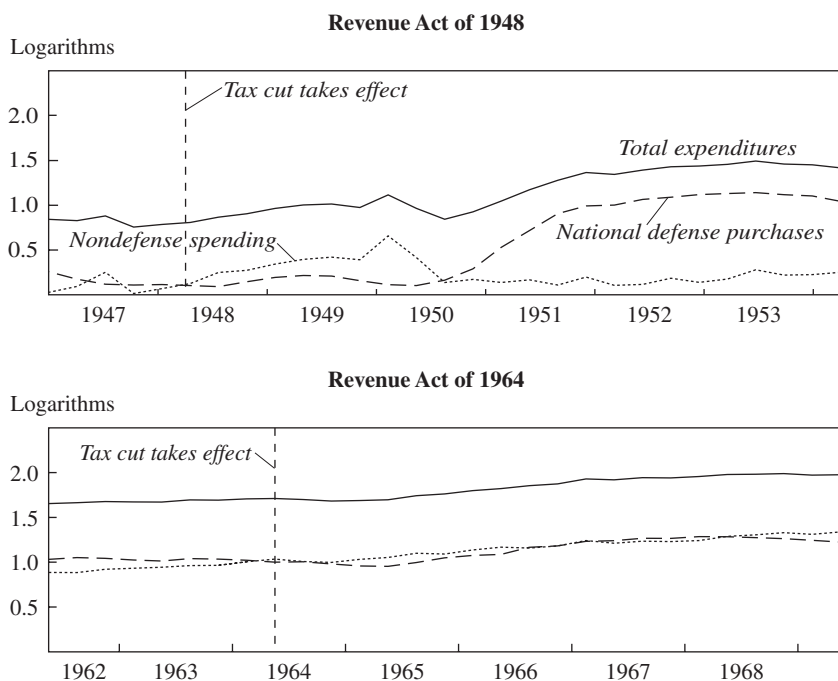
spending or in either of the two categories of spending. Indeed, growth in overall spending increased after the tax cut. Total expenditure, which had been essentially flat before the tax cut, rose by 16 percent (calculated using the change in logarithms) in the two years between the cut and the start of the war. Second, there was a substantial one-time spike in nondefense spending in 1950Q1, reflecting a one-time dividend payment from the trust fund for National Service Life Insurance (the government insurance program for military personnel). 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, both defense and overall spending rose sharply after the outbreak of the war.

Both the National Service Life Insurance dividend payment 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. Thus, they tend to cause our regressions to overstate the impact of tax cuts on subsequent spending increases.

Another important unusual development operated 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 benefits were initially small, these changes had little immediate impact on overall spending. Nonetheless, the rise in the benefit base and the expansion of coverage contributed significantly to the growth of spending over time. The fact that these delayed spending effects are not captured by our regressions tends to make them 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 that were largely spending driven. The first, and least important, was an increase in Social Security taxes of 0.3 percent of GDP in 1950Q1, which had been legislated before the tax cut was passed. Larger tax actions followed. The Social Security Amendments of 1950 increased the base of the payroll tax from $3,000 to $3,600, effective at the beginning of 1951, and called for a gradual increase in the combined (employer plus employee) 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 4.1 percent of GDP.[32]
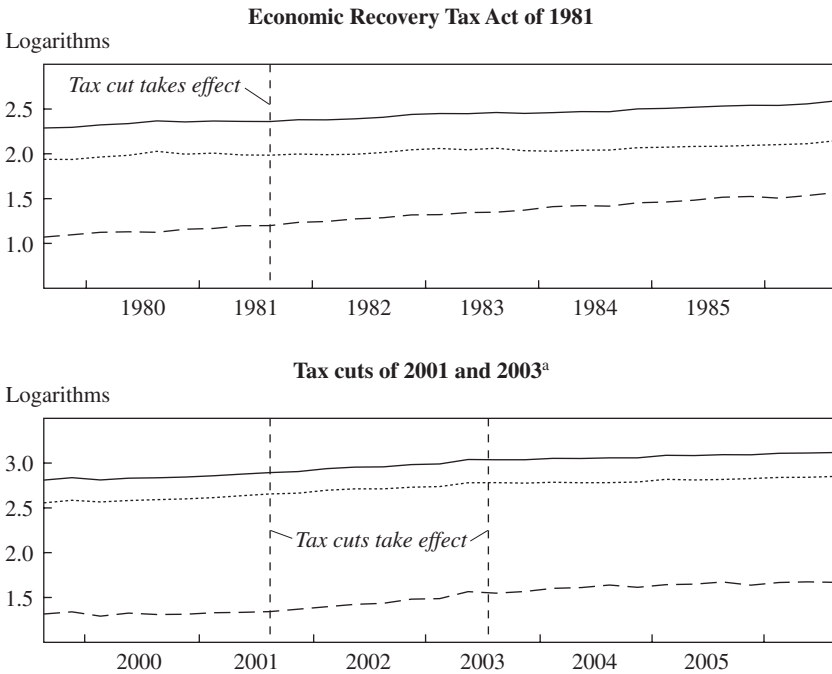
---

32. We measure the effect of a series of tax changes by finding the share of each one in nominal GDP in the quarter in which it took place, and then summing the shares.

Figure 11. Real Expenditure Following Major Long-Run Tax Cuts

Revenue Act of 1948

Logarithms

Tax cut takes effect

2.0

1.5

1.0

0.5

Total expenditures

Nondefense spending

National defense purchases

1947   1948   1949   1950   1951   1952   1953

Revenue Act of 1964

Logarithms

Tax cut takes effect

2.0

1.5

1.0

0.5

1962   1963   1964   1965   1966   1967   1968

The move to spending-driven tax increases in the early 1950s was clearly a policy decision. In the case of Social Security, policymakers were grappling with 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 revenue. Instead, however, the 1950 amendments repealed the provision of the Social Security Act that permitted financing from general revenue, 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). However, we have found no direct evidence that the 1948 tax cut played a causal role in this decision.

The extent of the government's reliance on contemporaneous tax increases to finance the Korean War is remarkable: total government expenditure rose by 6.0 percent of GDP 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

Figure 11. Real Expenditure Following Major Long-Run Tax Cuts (*Continued*)

Economic Recovery Tax Act of 1981

Logarithms

Tax cut takes effect

[Chart showing three lines (solid, dotted, dashed) with y-axis values 1.0, 1.5, 2.0, 2.5 and x-axis years 1980, 1981, 1982, 1983, 1984, 1985, with a vertical dashed line at 1981]

Tax cuts of 2001 and 2003[a]

Logarithms

Tax cuts take effect

[Chart showing three lines (solid, dotted, dashed) with y-axis values 1.5, 2.0, 2.5, 3.0 and x-axis years 2000, 2001, 2002, 2003, 2004, 2005, with two vertical dashed lines at 2001 and 2003]

Source: Authors' calculations using National Income and Product Accounts data.

a. The first vertical line marks the date of the Economic Growth and Tax Relief Reconciliation Act of 2001, and the second that of the Jobs and Growth Tax Relief Reconciliation Act of 2003.

this heavy reliance on tax finance. Soon after the war began, 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," July 25, 1950, p. 1)

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

IV.B. The Revenue Act of 1964

In February 1964 President Lyndon Johnson signed the Revenue Act of 1964. The act reduced revenue 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 economic growth following the tax cut was very rapid, revenue recovered quickly, and budget deficits that could have triggered a starve-the-beast response did not emerge immediately. Nevertheless, policymakers' statements and behavior 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," February 10, 1964). His "Great Society" speech followed in May 1964, calling for the elimination of poverty, urban renewal, pollution reduction, and expansion of education ("Remarks at the University of Michigan," May 22, 1964). Over the next year, a number of spending increases directed at achieving these goals were passed. The most significant were the dramatic expansion of benefits and the introduction of Medicare contained in the Social Security Amendments of 1965.

The Johnson administration believed that spending should be determined by necessity and efficiency. For example, 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. . . . It is preferable to emphasize changes in tax rates (suitably coordinated with changes in monetary policy) for stabilization purposes" (p. 68). The narrative record in this episode is striking in the degree to which revenue was *not* mentioned as a determinant of expenditure.

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

> 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," January 12, 1966, p. 2)

Congress went along with his calls for increased spending. For example, the Social Security Amendments of 1967 brought about another substan-

tial increase in benefits and a significant increase in coverage. Thus, the rise in spending following the tax cut was not just the consequence of the war.

Beginning in early 1966, policymakers began to worry that the economy was overheating, and by late that year the budget deficit had increased substantially. Nevertheless, the administration did not call for substantial spending reductions. Federal expenditure was expected to rise by $15 billion in 1968 (1968 *Economic Report,* p. 54). Instead, the administration concluded that "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 projections) in the 1968 bill imposing a 10 percent temporary tax surcharge. Congress pressed for the spending cuts not because revenue had declined, but because members 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 those of Senator Robert Byrd of West Virginia, 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, 2d session, volume 114, part 7, April 2, 1968, p. 8561). The tax cut was surely one factor contributing to the overheating that motivated the tax surcharge. Therefore, although policymakers did not explicitly draw a direct link between the tax cut and the spending reduction, the reduction is the one development in this episode that could suggest some connection between tax cuts and subsequent spending decisions.

The actual behavior of spending following the 1964 tax cut is completely consistent with policymakers' stated positions. The second panel of figure 11 shows that total expenditure was basically constant during the first year after the tax cut but then rose strongly. Total expenditure increased by 27 percent in the five years after the tax cut, noticeably more than the 18 percent in the five years before the cut.[33] The rise in defense purchases was one source of the increase, but nondefense spending, fueled by a large increase in transfer payments, increased even more rapidly.

Special factors clearly played a role in the behavior of spending. Much of the rise in defense expenditure was related to the Vietnam War. To the extent

33. These changes are computed as the change (in logarithms) of our measure of real total gross expenditure less interest over the periods 1959Q2–1964Q2 and 1964Q2–1969Q2. The other figures for spending growth reported in this section are computed similarly.

that defense spending truly was nondiscretionary, 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 understates in a fundamental way 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. Thus, in present value terms, the increase in spending 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, passed in January of that year. The administration viewed this tax cut as a continuation of the 1964 action. In this case, then, the serial correlation of tax changes reflects continuity in views about appropriate policy. The second was the Social Security Amendments of 1965, which included a substantial increase in payroll taxes to help pay for a large 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 by raising taxes, because the decision had been made in 1950 that the Social Security system should be self-financing.[34]

The overheating of the economy beginning in 1966 led policymakers to advocate tax increases. The Tax Adjustment Act of 1966 (enacted in March) rescinded the excise tax reduction of the previous January, and Public Law 89-800 (enacted in November) suspended the investment tax credit. Together these two tax increases were expected to raise revenue by 0.3 percent of GDP.[35]

By far the largest tax increase in the immediate post-1964 period was the 1968 surcharge. The administration first proposed a 6 percent surcharge in January 1967. In August 1967 Johnson stated, "If left untended, this deficit could cause . . . a spiral of ruinous inflation" and "brutally higher interest rates" ("Special Message to the Congress: The State of the Budget

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

35. Public Law 90-26, enacted in June 1967, restored the investment tax credit. As discussed in Romer and Romer (2009), 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 by short-run macroeconomic conditions.

and the Economy," August 3, 1967, p. 1). He requested that the surcharge be increased to 10 percent, 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 explicit in saying that the surcharge was undoing part of the 1964 tax cut. In his signing statement he said, "This temporary surcharge will return to the Treasury about half the tax cuts I signed into law in 1964 and 1965" (June 28, 1968, p. 1). This action, combined with the continued rise in expenditure, is a vivid example that what typically gives in response to a tax cut is not spending but the tax cut itself.

IV.C. The Economic Recovery Tax Act of 1981

A very large long-run tax cut was enacted in August 1981, shortly after President 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 "spending 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," February 18, 1981, pp. 1, 5). Similarly, in his first budget message, in February 1982, he listed "reducing 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 1981 tax cut was followed by a substantial fall in revenue and a sharp rise in the 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 evidence that the decline in revenue due to the 1981 tax cut affected later spending decisions.

The third panel of figure 11 plots government spending before and after the 1981 tax cut. The vertical line is drawn at 1981Q3, the date of the first of the series of cuts. Despite what the narrative evidence suggests, growth in overall spending did not slow but actually quickened. In the five years following the tax cut, total expenditure grew by 23 percent, substantially above the 14 percent growth in the five years before the cut. This acceleration in overall spending reflects a combination of a large rise in the growth of defense spending and a more moderate rise in the growth of nondefense spending.

Two important unusual spending developments marked 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, President 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). The balance of political power in Congress also swung sharply toward advocates of spending restraint at the time of Reagan's election. Thus, there was clearly an omitted variable acting to reduce spending in this episode.[36]

Second, the heightening of the cold war prompted policymakers to increase defense spending. Ramey and Shapiro (1998), for example, identify the Soviet invasion of Afghanistan at the end of 1979 as an exogenous positive shock to defense spending. This factor operated in the opposite direction of the political shift toward supporters of lower spending.

The tax cuts were followed by two types of tax increases. First, the Social Security Amendments of 1983 called for a series of payroll tax increases from 1984 to 1990 to improve the solvency of the Social Security system. These increases appear to have been largely a continuing consequence of the 1950 decision to make the Social Security program self-financing.

---

36. Although Reagan supported spending reduction in general, he favored higher defense spending. He 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).

Second, a series of income tax increases were explicitly motivated by a desire to reduce the budget deficits that emerged following 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," August 16, 1982, p. 4). Similarly, the 1989 *Budget* reported that the 1987 act was enacted "in 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. Thus, these tax increases were a fairly direct response to the earlier tax cut.

IV.D. The Tax Cuts of 2001 and 2003

Two long-run tax cuts were passed early in the administration of President George W. Bush. The Economic Growth and Tax Relief Reconciliation Act of 2001, enacted in June, included a long-run tax cut of 0.8 percent of GDP in 2002Q1, as well as a large countercyclical tax cut in 2001Q3. The Jobs and Growth Tax Relief Reconciliation Act of 2003, enacted in May, 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 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 the national 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," February 28, 2001, p. 7).[37]

---

37. 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 www.gpoaccess.gov/usbudget/fy02/index.html.

It also said that "the 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 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 terrorist attacks of September 11, 2001, discussions of budget policy placed less emphasis on spending restraint (see, for example, Bush's "Address before a Joint Session of the Congress on the State of the Union," January 29, 2002, pp. 3–4). 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 5 years" ("Address before a Joint Session of the Congress on the State of the Union," January 20, 2004, p. 4). Similarly, in his 2007 State of the Union Address, Bush 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," January 23, 2007, p. 1). Although these statements were very similar to those Bush had made before the tax cuts, actual budget conditions had changed substantially: revenue 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 little link between the level of revenue and the perceived need for spending restraint.

The last panel of figure 11 plots the major categories of spending in this episode. The two vertical lines show the dates that the two tax cuts first took effect. As in the other episodes, overall spending growth did not slow. In the five years following the first cut in 2001Q3, spending grew by 22 percent, substantially more than the 14 percent in the five years before the cut. The growth in spending following the tax cut was greatest in defense: national defense purchases rose by 33 percent in the five years after the tax cut, while nondefense spending rose by 19 percent.

The events of September 11, 2001, were clearly an important outside influence on spending. Some of the behavior of total expenditure surely reflects the impact of this development rather than the effect of the tax cuts. On the other hand, one important spending action 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 at the end of 2004. Thus, the issue of how the government will eventually deal with the loss of revenue from the 2001 and 2003 tax cuts remains open.

IV.E. Assessment

Examination of these four episodes of major long-run tax cuts reinforces the findings from our statistical analysis: there is little evidence of a starve-the-beast effect. The one aspect of the episodes that is at times 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 revenue in formulating their budget policies, the presidents in the other two (Truman and Reagan) cited the level of revenue as a consideration in budget policy. Even in these cases, however, other factors were clearly much more important, and to a considerable extent the concern over revenue led not to advocacy of spending reductions, but to support for (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, all of the episodes saw an acceleration of spending. This is similar to the overall statistical finding of a positive (although only marginally significant) effect of tax cuts on spending, and it suggests that the regression results reflect a consistent pattern in the data rather than the effects of outliers.

Examination of other influences on spending in the episodes does not change these conclusions. On the one hand, there was an important external development in each episode that acted to raise defense spending. 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.[38] First, the largest of the tax cuts (that of 1981) coincided with the election of a president who had a strong commitment to reducing the size of government. This suggests that the positive impact of tax cuts on spending might be even larger than implied by the regressions. Second, significant actions were taken in three of the four 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. Since our regressions miss much of the effects of these actions, 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 four episodes suggests that, on net, the regressions do not overstate the evidence against the starve-the-beast hypothesis.

Tax policy in these episodes is also consistent with the regression results. In three of the four episodes, substantial tax increases followed the initial tax cut within five years, offsetting a substantial fraction of it. Perhaps more striking 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—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 the tests that have been done have important limitations.

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 were 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, govern-

---

38. In addition, recall that our statistical results are robust to controlling for a measure of exogenous shocks to defense spending, and that even excluding defense spending entirely provides little evidence for the starve-the-beast hypothesis.

ment spending does not fall. Indeed, if anything, spending rises, providing some support for the alternative view of fiscal illusion or shared fiscal irresponsibility. The lack of support for a starve-the-beast effect is 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 revenue 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 in these cases, however, runs from the decision to raise spending to the tax change. For the full postwar sample, this type of tax change is sufficiently common that it causes the overall relationship between tax revenue and spending to be significantly positive. Excluding these spending-driven changes makes the relationship negative and marginally significant.

The fact that tax cuts do not lead to spending reductions raises the question of how the government budget constraint is ultimately satisfied. We find that long-run tax cuts are offset by legislated and nonlegislated tax increases over the next several years. The fact that policymakers are able to make changes on the tax side helps to explain why they do not appear to make large changes on the spending side.

Of course, failing to find support for the starve-the-beast hypothesis is not the same as definitively refuting it. There are several ways in which our results are not inconsistent with the presence of at least some starve-the-beast effect. First, our failure to find such an effect for the postwar U.S. federal government does not mean it is not important in other times and places. Second, the case that focusing on tax changes taken for long-run purposes yields unbiased estimates is not airtight. As we explain, however, the most likely direction of bias is in favor of the starve-the-beast hypothesis, not against it. Third, because our estimates are not highly precise, the hypothesis that tax cuts exert some restraining influence on spending usually cannot be rejected. Fourth, some of our evidence (the statistical examination of nondefense spending and the narrative evidence for the 1948 and 1981 episodes) provides some hints of support for a small starve-the-beast effect.

Finally, although we find that the fall in revenue 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. 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. It is possible that some of the remaining adjustment will take place on the spending side.

Taken together, these caveats imply that one cannot necessarily conclude that tax cuts do not restrain government spending at all. But it remains the case that, over the period we consider, there is virtually no evidence of such an effect.

The finding that tax cuts do not appear to substantially 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 reduce their estimates of the present value of their present and future liabilities, and so to 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 Torsten Persson and Lars Svensson (1989), Guido Tabellini and Alberto Alesina (1990), and others. If decisionmakers understand that tax cuts do not in fact lead to substantial reductions in government spending, these mechanisms are much less important. Thus, better estimates of the effects of tax cuts on spending may require changes to the modeling of a wide range of issues.

References

Anderson, William, Myles S. Wallace, and John T. Warner. 1986. "Government Spending and Taxation: What Causes What?" *Southern Economic Journal* 52, no. 3: 630–39.

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, no. 1: 9–23.

———. 2003. "Fiscal Policy, Past and Present." *BPEA,* no. 1: 75–122.

Becker, Gary S., and Casey B. Mulligan. 2003. "Deadweight Costs and the Size of Government." *Journal of Law and Economics* 46, no. 2: 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, no. 4: 1329–68.

Bohn, Henning. 1991. "Budget Balance through Revenue or Spending Adjustments? Some Historical Evidence for the United States." *Journal of Monetary Economics* 27, no. 3: 333–59.

———. 1992. "Endogenous Government Spending and Ricardian Equivalence." *Economic Journal* 102, no. 412: 588–97.

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, no. 1: 119–31.

Gale, William G., and Peter R. Orszag. 2004. "Bush Administration Tax Policy: Starving the Beast?" *Tax Notes* 105, no. 8: 999–1002.

Hines, Frank T. 1943. "National Service Life Insurance." *Annals of the American Academy of Political and Social Science* 227: 83–93.

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, no. 1: 221–29.

Niskanen, William A. 1978. "Deficits, Government Spending, and Inflation: What Is the Evidence?" *Journal of Monetary Economics* 4, no. 3: 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, no. 2: 325–45.

Ram, Rati. 1988. "Additional Evidence on Causality between Government Revenue and Government Expenditure." *Southern Economic Journal* 54, no. 3: 763–69.

Ramey, Valerie A. 2008. "Identifying Government Spending Shocks: It's All in the Timing." University of California, San Diego (June).

Ramey, Valerie A., and Matthew D. Shapiro. 1998. "Costly Capital Reallocation and the Effects of Government Spending." *Carnegie-Rochester Conference Series on Public Policy* 48: 145–94.

Romer, Christina D., and David H. Romer. 2009. "A Narrative Analysis of Postwar Tax Changes." University of California, Berkeley (June).

———. Forthcoming. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review.*

Tabellini, Guido, and Alberto Alesina. 1990. "Voting on the Budget Deficit." *American Economic Review* 80, no. 1: 37–49.

Von Furstenberg, George M., R. Jeffery Green, and Jin-Ho Jeong. 1986. "Tax and Spend, or Spend and Tax?" *Review of Economics and Statistics* 68, no. 2: 179–88.

Comments and Discussion

COMMENT BY

STEVEN J. DAVIS   In this paper Christina Romer and David Romer investigate the hypothesis that tax cuts curtail government spending. To do so, they study the experience of the federal government since 1945. They stress, quite rightly, that the empirical relationship between tax changes and spending changes depends greatly on why the changes occurred. Some tax change episodes are potentially informative about the hypothesis, and others are not.

This observation underlies their two-step empirical strategy. First, Romer and Romer use contemporaneous narrative sources to determine the motives for legislated tax changes. The goal is to identify tax changes that aim to spur productivity growth or promote other long-run objectives. They argue that such tax changes are less likely to be correlated with other factors that drive government spending and, hence, are more informative about the effect of tax changes on government spending. In the second step, they examine the response of government spending to these informative tax change episodes. They consider a variety of statistical specifications, and they supplement the statistical analysis with a detailed examination of four large tax changes.

The authors execute this empirical strategy with considerable care and skill.[1] They conclude that the results provide "virtually no evidence" that tax cuts restrain government spending. Instead, the results suggest that tax cuts motivated by long-run objectives are largely offset in the ensuing years by tax increases. They provide a balanced summary of these and other results in their concluding section.

---

1. I encourage the reader to consult their closely related paper (Romer and Romer 2009) to gain a fuller appreciation for the care and skill that they bring to the first step of their empirical strategy.

In my view, legislated tax cuts have done little to restrain U.S. government spending in the postwar era. I reach this view based mainly on the arguments sketched in Romer and Romer's section III.C. These arguments rely on economic reasoning about the force of the mechanisms that link current tax cuts to future government spending. I place less weight on the results of the two-step empirical strategy outlined above. The strategy is a sensible one, but it does not yield sharp inferences in a sample focused on the postwar U.S. experience. This fact shows up as large standard errors for the estimated spending responses to tax cuts. In addition, and despite the authors' careful effort, it is hard to fully dispel concerns about the classification of tax change episodes and concurrent developments that influence the estimates.

Section III.C describes two mechanisms whereby tax cuts might curtail future government spending. One mechanism works through the link between current tax cuts and future debt-servicing costs. In particular, a deficit-financed tax cut today means higher debt-servicing costs in the future, leading future policymakers to choose a lower level of noninterest government spending than otherwise. A second mechanism rests on the political and economic costs of reversing a tax cut.

To assess the force of the first mechanism, assume linear marginal schedules for the costs and benefits of government spending:

$$MC = 1 + cg, \quad c > 0,$$

$$MB = m - bg, \quad m > 0 \text{ and } b \geq 0;$$

where g is the ratio of government spending to GDP, and c, b, and m are parameters. Treating output as exogenous and equating benefits and costs at the margin, the policymaker chooses $g^* = (m - 1)/(b + c)$ for the size of government. This outcome need not be optimal from the perspective of the median voter or a utilitarian social welfare criterion. It simply reflects the policymaker's preferred outcome in light of budgetary and political pressures.

When a policymaker implements a deficit-financed tax cut, this raises the MC schedule facing future policymakers. In the example offered in section III.C, the policymaker cuts taxes by 2 percent of GDP for five years, raising the debt-to-GDP ratio by about 10 percentage points. Given a real interest rate that exceeds the output growth rate by 2 percentage points a year, the implied rise in debt-servicing costs amounts to about 0.2 percent of GDP and 1.0 percent of government spending. Accounting for this upward

shift in the MC schedule, the effect is to lower future government spending by $c/(c + b)$ multiplied by 0.2 percent of GDP, that is, by at most 0.2 percent of GDP. This is a very small starve-the-beast effect. Relaxing the assumption of exogenous output and allowing for tax cuts to stimulate growth yields an even smaller restraint on government spending.

Since the example is similar in size to the largest tax cut episodes in the postwar U.S. experience, this analysis implies that tax cuts have not, through their effects on debt-servicing costs, significantly restrained government spending. It also implies that the mechanism is much too weak to be detected in a sample of postwar U.S. tax changes. Of course, the mechanism operates with greater force when there is a bigger rise in the debt-to-GDP ratio or the government faces a higher real interest rate. In the postwar U.S. setting, however, the first mechanism has little force.

Now consider the second mechanism. If tax cuts are hard to reverse for political or economic reasons, it is easy to see that they exercise more restraint on future government spending. Building on the previous example, if it takes 5 years for a new policymaker to reverse a previous tax cut, so that it remains in effect for 10 years rather than 5, the starve-the-beast effect roughly doubles. In the extreme case where tax cuts cannot be reversed, government spending cuts must eventually absorb the entire adjustment. Clearly, then, tax cuts can produce large starve-the-beast effects if they are sufficiently sticky. Thus, the force of the second mechanism depends on the difficulty of reversing tax cuts in practice.

Romer and Romer address this issue in their section III.B. Figures 9 and 10 provide strong evidence that tax hikes usually follow in the wake of tax cuts motivated by long-run concerns. The final panel of figure 9 suggests that about three-quarters of the tax cut is reversed within five years, and it provides little evidence against the hypothesis of full reversal. This evidence, coupled with the analysis above, indicates that tax cuts of the sort that dominate the postwar U.S. experience are not sticky enough to generate large starve-the-beast effects.

In short, neither mechanism operates with much force under the conditions that have prevailed in the postwar United States. This conclusion has important implications for economic policymaking and for models of fiscal behavior, as the authors discuss. However, the conclusion also has limited scope. In particular, it does not apply to tax changes or other fiscal policy actions that are hard to reverse. My remaining remarks develop this point.

Most developed economies rely on a national value added tax (VAT) as a major source of government revenue. The United States is a large outlier

in this respect. Many, perhaps most, economists look on the VAT with favor because of its broad tax base, ease of administration, and pro-saving incentive effects. These observations motivate many proposals to introduce a national VAT or other broad-based consumption tax in the United States. In contrast, Gary Becker and Casey Mulligan (2003), among others, question the desirability of introducing a broad-based consumption tax, which in their view would lead to substantial increases in federal spending. I share this view, and I see it as fully consistent with the evidence produced by Romer and Romer's two-part empirical strategy and with my analysis of the mechanisms whereby tax cuts restrain government spending.

Two observations are important in this regard. First, I expect that a new national consumption tax, once introduced, would be hard to reverse. In all likelihood, it would become a permanent feature of the U.S. fiscal landscape. In this respect, U.S. experience with "routine" tax changes in the postwar era is not a good guide to the reversibility of a new national consumption tax. Second, I agree with most other economists that the VAT and other broad-based consumption taxes rank highly on standard economic efficiency criteria. In addition, the VAT is less visible and less salient to taxpayers than the personal income tax and hence less likely to generate political pressure for lower taxes. For this reason, as well, the VAT generates lower marginal costs of government revenue as perceived by the policymaker.

To parameterize the effects of introducing a broad-based consumption tax, rewrite the marginal cost schedule for government revenues as

$$\mathrm{MC}' = 1 + (1 - \gamma)cg.$$

The new parameter γ captures the effect of introducing the VAT on the marginal cost of funds, again as perceived by the policymaker. Comparing outcomes under MC and MC′, it is easy to show that the introduction of a VAT increases the size of government by

$$\frac{\Delta g}{g} = \frac{b + c}{b + (1 - \gamma)c}.$$

As an example, suppose $\gamma = 0.2$, which corresponds to a reduction in the marginal cost of funds from 1.5 to 1.4 with $c = 0.5$. Using the formula above and $\gamma = 0.2$, the introduction of a VAT causes government spending to rise by 25 percent when $b = 0$, and by 11 percent when $b = c$. Obviously, these are large effects on the size of government.

There is certainly room to improve and deepen this analysis by embedding it in a fuller model and by grounding the choice of parameter values. The analysis is sufficient, however, to support two conclusions. First, there are good reasons to anticipate that the introduction of a national consumption tax would lead to a large expansion in the size of government. Second, this first conclusion is fully consistent with the evidence in this paper and with my analysis of the mechanisms that link current tax cuts to future government spending.

As a final remark, it should be clear that a similar analysis applies to other new sources of government revenue that lower the marginal cost of government revenue from the perspective of policymakers. Cap-and-trade proposals to limit carbon emissions and other pollutants are a good case in point. These proposals have the potential to raise large amounts of government revenue in ways that are opaque to most taxpayers and that will make it easy for politicians to deflect the blame for higher energy costs onto energy producers, electric utilities, and others. These features of cap-and-trade proposals are likely to lower the marginal cost of government revenue from the perspective of policymakers and to lead to higher government spending as a result.

REFERENCES FOR THE DAVIS COMMENT

Becker, Gary S., and Casey B. Mulligan. 2003. "Deadweight Costs and the Size of Government." *Journal of Law and Economics* 46, no. 2 (October): 293–340.
Romer, Christina D., and David H. Romer. 2009. "A Narrative Analysis of Postwar Tax Changes." University of California, Berkeley (June).

COMMENT BY
JEFFREY A. MIRON   I was delighted to be asked to discuss this paper, in part because I enjoy reading anything by Christina Romer and David Romer, and in part because I believe this is an important topic. Although I had not spent a significant amount of time thinking about the starve-the-beast hypothesis before taking up the paper, my hunch had always been that the standard version was probably correct. I think my gut instinct, however, came from thinking about the hypothesis in terms that are the reverse of the way Romer and Romer state it: that is, my guess was that if some event provides policymakers with additional tax revenue, they will spend it, not save it. If one assumes that the effect is symmetric, then the standard starve-the-beast conclusion follows. So, implicitly assuming symmetry, I took the hypothesis as at least plausible.

The paper thus initially presented me with a dilemma, since I am hard pressed to think of a paper by either or both of these authors that I did not find convincing. In particular, I liked the precursor to this paper (Romer and Romer 2009), for two reasons. On the one hand, that paper made a solid case for their approach to identifying the effects of tax cuts. On the other, that paper's result was consistent with my prior, which is that tax cuts should increase output because, on average, tax cuts mean lower tax rates, and that means improved incentives.

My goal in reviewing the current paper, therefore, is to determine whether some aspect of their interpretation might not be the whole story, or whether instead my instincts about the starve-the-beast hypothesis were just wrong. In the end, my conclusion merges a bit of both possibilities. I will explain this by first discussing the aspects of the paper that I do not wish to dispute, and then by presenting a modified interpretation of certain key results that I think can reconcile their results and my priors.

OVERALL EVALUATION. The first aspect of the paper that I do not wish to challenge is the authors' strategy for identifying the effects of tax cuts. This is not to say that I regard that strategy as beyond all possible quibbling. For example, policymakers' stated reasons for a particular tax change might differ from their actual reasons, and even their stated intentions might be ambiguous in some cases. Nevertheless, no approach to identification is beyond reproach. On the whole, I find the authors' strategy far more convincing than most of those commonly used.

The second aspect of their paper that I find myself unable to challenge is the thoroughness of their empirical investigation. That is, I have not identified ways in which some aspect of that analysis seems inappropriate or incomplete. On the contrary, every time I thought I had discovered a possible weakness, such as some alternative specification that might yield a different answer, I discovered a page or two later that they had already addressed the issue and that it did not make much difference to their overall results.

One such issue might be worth mentioning, however, since I actually missed their treatment of it the first time through and therefore spent some effort, courtesy of their data, examining it on my own. I have long had the hunch that divided government (gridlock) might be a significant factor in slowing expenditure, reducing the deficit, and even improving output growth. I thought the authors' failure to find a starve-the-beast effect might be due to omission of this factor. In fact, I could not find any gridlock effect, and the authors had in fact tested this hypothesis themselves and come to the same conclusion.

So, given this assessment, it might seem that reconciliation of their results with my priors requires me to update my priors. That will be part of the resolution, but not the whole story. To show this, I will examine two specific results in more detail.

INTERPRETING THE RESULTS ON LONG-RUN TAX CHANGES. The first of the authors' results that I think bears additional scrutiny is their baseline result, reported in their table 1 and figure 2, which indicates that exogenous tax cuts (what they call long-run tax changes) do not appear to lead to reductions in expenditure. Indeed, the authors find mild evidence that these tax cuts lead to increased expenditure over the 5-year horizon, although this effect seems to disappear over the 10-year horizon (see their figure 3).

A possibly relevant objection, however, is that virtually all the exogenous changes in taxes in their data are tax cuts, not tax increases. The top panel of their figure 1, which plots the exogenous tax variable, shows mainly decreases in taxes throughout the sample period, with only a few examples of increases. This makes sense, since Romer and Romer identify exogenous tax changes as those motivated by a desire to shrink government or improve incentives, and it is not obvious why these motivations would favor tax increases.

One can confirm that their main result is dominated by the exogenous tax cuts rather than the exogenous tax increases by rerunning their baseline regression using only those tax changes that are decreases. Figure 1 below, which is virtually the same as their figure 2, shows the results. Tax cuts do not appear to starve the beast and may even feed it.

So, given that their results are dominated by episodes of tax cuts, it is clear that they do not necessarily address my prior that a windfall tax increase might cause expenditure to increase. One could assume that the relationship is symmetric, in which case the latter proposition follows from the former, but there is no a priori reason why the effect has to be symmetric. Given sufficient observations on exogenous tax increases, one could examine the possibility of asymmetry directly. It seems unlikely that such an exercise would be fruitful in their dataset, however, because there are so few exogenous increases in their sample period. More generally, given the classification system they have used, it seems unlikely that one could ever examine this asymmetry, since it is not obvious that policymakers would ever announce that their intention is to make incentives worse.

The bottom line on this first result is therefore the following: I take the authors' result as convincing when stated as they state it, that is, that exogenous tax cuts do not starve the beast. The results are silent, however, on whether exogenous tax increases feed the beast.

Figure 1. Cumulative Impact of a Tax Cut of 1 Percent of GDP on Total Expenditure, Sample Excluding Tax Increases[a]

Percent

Quarters after tax change
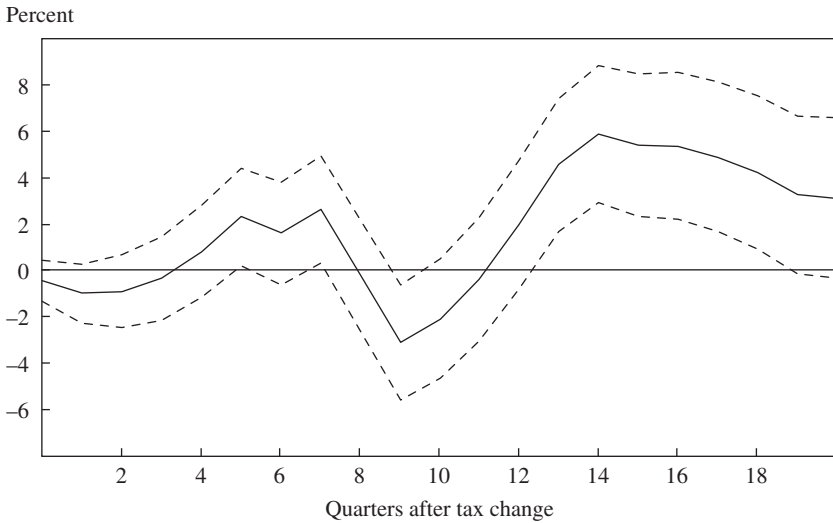
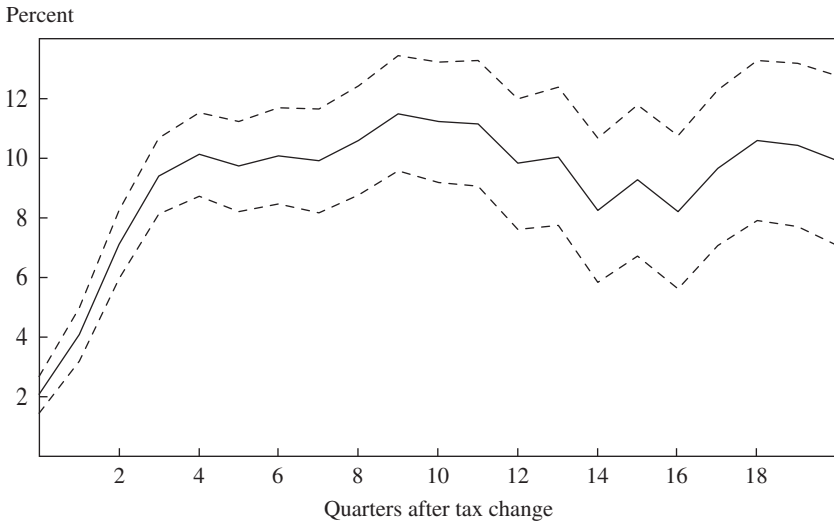Source: Author's regression using data from Romer and Romer, this volume.
a. The sample includes only those long-term tax changes identified by Romer and Romer that are tax decreases.

INTERPRETING THE RESULTS ON SPENDING-DRIVEN TAX CHANGES.   The second result I want to consider in more detail is the finding that spending-driven tax cuts are followed by noticeable reductions in expenditure (see the panel labeled "Spending-driven tax changes" in the authors' figure 6). Romer and Romer argue that this should not be taken as evidence in favor of the starve-the-beast hypothesis, because the correlation confounds a missing, unmeasured variable, namely, prior decisions to change spending. Such decisions plausibly move spending and taxes in the same direction, independent of any causal impact of taxes on spending.

The authors' argument for not regarding this as evidence for the starve-the-beast hypothesis is appropriate given the way that its advocates have typically stated the hypothesis, arguing that any tax cut is good because it helps shrink government. This view suggests an independent effect of tax cuts, but one can only estimate that effect by controlling for other factors, like antigovernment sentiment, that might also reduce spending.

Again, however, it is useful to examine this result a bit more carefully, and to pose the question as the reverse of the way the authors present it. In their sample, most spending-driven tax changes are increases, not decreases

Figure 2. Cumulative Impact of a Spending-Driven Tax Increase of 1 Percent of GDP on Total Expenditure[a]

Percent

Quarters after tax change

Source: Author's regression using data from Romer and Romer, this volume.
a. Figure inverts the fourth panel in figure 6 in Romer and Romer, this volume, so that it depicts the effect of an increase rather than a decrease in taxes by 1 percent of GDP.

(again see their figure 1). Hence, their result is mainly saying that when taxes *increase* because policymakers want to increase spending, expenditure in fact *goes up.* Figure 2 above shows this explicitly simply by presenting the mirror image of the analogous graph in the paper.

Even more important, this figure shows that for an expenditure-driven tax increase, expenditure increases by well more than one for one. Specifically, a tax cut of 1 percent of GDP equals about 5 percent of government spending, and the estimates suggest that even 20 quarters out, a spending-driven tax increase of that magnitude raises government expenditure by 10 percent. Thus, the long-term increase in spending is about twice the initial increase in taxes.

Why might this occur? The obvious explanation is that initial estimates of program costs are systematically below the eventual costs. Congress, for example, might systematically underestimate costs in order to get programs adopted, or political forces might lead to the expansion of programs once they have been adopted, whether or not the initial costs were fair estimates of the future costs. As a result, if the size of the tax increase was chosen to match the initial estimate of program costs, the actual costs incurred will

far exceed the tax increase. Whatever the mechanism, the implication is that spending-driven tax increases feed the beast, or at least allow the beast to feed itself.

Thus, my interpretation of these results is more nuanced than the authors' interpretation. I agree with their assessment that exogenous tax cuts do not starve the beast. Their evidence would still appear to be consistent, however, with my prior and with the broader concern of small-government advocates, which is that when policymakers have ready access to tax revenue, they spend it.

A simple story to account for this combination of results goes as follows. Politicians want to spend money because that helps them get reelected. The kind of spending they seek differs from politician to politician according to the political preferences of their districts, but logrolling and earmarking allow everyone to be happy when money is free and easy. Thus, if politicians are flush with cash, the temptation to spend is huge. If instead politicians are pushed to reduce spending, they resist, because they usually get more benefit from higher spending than from tax cuts, and so they find ways to raise taxes back up when they can. This simple "model" does not validate the claim that all tax cuts are good tax cuts because they starve the beast, but it does suggest that concerns over letting children play with matches—that is, giving politicians access to increased tax revenue—are valid. Thus, advocates of small government would seem to have good reason to oppose tax increases.

HOW SHOULD ADVOCATES OF SMALL GOVERNMENT RESPOND TO THESE RESULTS? One final issue is whether advocates of small government should be unhappy or happy with the authors' results, taking them as correct. The fact that attempts to shrink government through tax cuts do not seem to work might at first blush strike small-government types as frustrating. Much of the citizenry has some interest in tax cuts, and politicians are sometimes interested in running on a tax-cutting platform, so this might appear an easy way to accomplish the goal of shrinking government, if the starve-the-beast hypothesis were correct.

Further reflection, however, should make advocates of small government fully comfortable with these results. The cut-taxes-first approach is at some level dishonest; it tries to shrink government while avoiding discussion of the fact that lower taxes mean less government. Advocates of small government should pride themselves on being honest about their intentions and have confidence that their criticisms of government are sufficiently convincing to carry the day without resort to trickery. That means reducing government by debating specific policies and programs on their merits.

The result that tax cuts are not sufficient to reduce government is also consistent with the view that institutional "tricks" are rarely successful at producing substantial and sustained changes in the way governments operate. Balanced-budget amendments are one such trick, but they founder on the fact that governments have access to innumerable accounting gimmicks for appearing to balance a budget while not really doing so (for example, by providing off-budget subsidies to Fannie Mae and Freddie Mac). Similarly, laws that allegedly establish central bank independence do not seem to bind in practice (Campillo and Miron 1997). This is not to say that institutions are irrelevant or to deny that having institutions that nudge in the right direction might help generate better outcomes. Institutions and tricks nevertheless do not seem to fundamentally change outcomes by themselves.

Finally, advocates of small government need not shed their view that tax cuts are desirable. After all, the very same methodology that invalidates the starve-the-beast hypothesis also suggests that tax cuts stimulate output substantially. What advocates of tax cuts presumably should do, however, is focus their attention not on any and all tax cuts, independent of their merit, but instead on those tax cuts that make sense from an efficiency perspective. At the same time, they need to refocus their efforts on convincing the populace that government spending is too high. If they can do that, lowering taxes should be easy.

REFERENCES FOR THE MIRON COMMENT

Campillo, Marta, and Jeffrey A. Miron. 1997. "Why Does Inflation Differ across Countries?" In *Reducing Inflation: Motivation and Strategy,* edited by Christina D. Romer and David H. Romer. University of Chicago Press.
Romer, Christina D., and David H. Romer. 2009. "A Narrative Analysis of Postwar Tax Changes." University of California, Berkeley (June).

GENERAL DISCUSSION   George Perry suggested that the introduction of inflation indexing of income tax brackets about halfway through the authors' sample period should have had a noticeable effect on spending if starve-the-beast effects were in fact important. In the years before indexing, politicians had the luxury of deciding what to do with the "fiscal dividend" that gradually arose. In the early 1960s, it provided fiscal room for a major tax cut without the need to restrain spending, whereas in the early 1970s it permitted an outrageous enhancement of Social Security benefits. Once tax brackets were indexed—a feature not captured by the authors' tax cut measure—discretionary tax cuts or spending increases should have been more constrained, and if starve-the-beast effects were significant, they should

have been more evident in this period. That they were not strengthens the authors' results.

Robert Shiller questioned the paper's implicit assumption that the starve-the-beast impulse takes the same form in all periods, suggesting instead that it was a Reagan invention. He noted that the largest long-run tax cut other than Reagan's during the sample period came in 1948 and could be attributed to postwar demobilization. The subsequent increase in spending could be explained by the Korean War. Both factors might offset the paper's results.

Robert Hall argued for analyzing the relationship between spending and taxation in the context of the level of U.S. national debt. Unlike some European countries whose debt is large enough to be in danger of falling below investment grade, the United States has maintained a persistently low debt-to-GDP ratio and a credit rating well above triple-A. Spending could indeed be much higher than it is, given the fiscal headroom provided by a small national debt. He suggested that a factor that contributes to keeping spending low in the United States but not in European countries is the former's racial and ethnic diversity, which may discourage spending on social programs if such spending tends to favor one group over another.

Benjamin Friedman agreed with Hall and with Steven Davis that the level of the national debt should be included in the analysis, and he proposed another, related factor to consider, namely, the relationship between the interest rate on the debt and the growth rate of the economy. Although the theoretical literature assumes that the real interest rate will exceed the real growth rate, the opposite was true during most of the authors' sample period. If the economy grows at a rate above the real interest rate, the ratio of the national debt to GDP will decline over time, weakening the tax burden argument that underlies the supposed starve-the-beast mechanism.

Caroline Hoxby noted that reducing the standard errors on the paper's main findings would be challenging given that there are essentially only four observations of the long-run tax cut variable. She also observed that testing the starve-the-beast hypothesis becomes nearly impossible if reductions in top marginal tax rates increase the rate of economic growth. Increased growth brings increased tax revenue without an increase in tax rates. For example, during the Reagan years marginal tax rates fell yet revenue increased significantly.

Matthew Shapiro stated that even though he found the paper's narrative believable, it did not match his understanding of the stylized facts. Around 1980 the U.S. political economy changed from one in which the debt-to-GDP ratio was steadily declining to one where, except during the Clinton

administration, the debt-to-GDP ratio has been generally increasing. The fiscal restraint of the first six years of the Clinton administration clearly arose in part because of concern about inherited deficits. He wondered why the authors' regressions did not pick up these broad trends. Two possible reasons were, first, that the lags used are too short, and second, the difficulty in inferring effects from time series that consist of only a small number of very persistent policy episodes.

Ricardo Reis remarked that although he appreciated the virtue of focusing on long-run tax cuts, given their exogenous nature, he worried that they are not representative of tax cuts in general. He suggested looking at the substance of tax cuts, in addition to their motivation, to determine whether the long-run cuts are really representative. Reis also noted that long-run tax cuts have only long-run benefits and therefore tend not to create short-run political advocates. As a result, these cuts are prone to reversal after a short while, with a change in administration or in the dominant ideology. Large, immediate cuts could avoid this problem and thus allow a starve-the-beast strategy a chance to force a correction of the resulting deficit through spending.

Gregory Mankiw credited Robert Reich and Henning Bohn with making him sympathetic toward the starve-the-beast hypothesis. Reich's book *Locked in the Cabinet* documents that the Clinton administration had had great spending plans but was prevented by the inherited Reagan-Bush budget deficits from carrying them out. However, the events in the book occurred roughly 12 years (48 quarters) after the Reagan tax cuts, a lag much longer than used in the paper and possibly beyond the capability of any econometric study. Henning Bohn's 1991 paper in the *Journal of Monetary Economics* also comes to a very different conclusion than the authors, and Mankiw suggested that the authors address that paper directly and explain why they believe Bohn was wrong.

Luigi Zingales agreed with Caroline Hoxby on the limitations imposed by using, in effect, only four observations. To get around this problem, he suggested looking at data from other countries with different levels of debt and different political constraints to determine whether a starve-the-beast strategy worked. Additionally, he noted that in corporate finance there is an analogy to the starve-the-beast hypothesis, namely, the free cash flow theory, which can be tested on micro rather than macro data and has found a lot of empirical support. Steven Davis agreed with Zingales but observed that extending the data internationally would entail a large amount of additional work. He also remarked that a desire to starve the beast could motivate many tax changes yet not significantly restrain spending. For example,

a current policymaker might implement a deficit-financing tax cut to undo the strategic beast-starving efforts of its predecessor. If political power changes hands every few years, then strategic tax cuts with a starve-the-beast motive can be both frequent and largely ineffective.

William Gale noted that the real-world experience in the United States since 1980 has been the opposite of what the starve-the-beast hypothesis predicts, unless a very long term story is told. The effect, if any, of tax changes on spending appears to be inverse: Ronald Reagan cut taxes and increased spending, Bill Clinton raised taxes and lowered spending, and George W. Bush cut taxes and raised spending again. Gale also cited a study he did with Brennan Kelly (published in *Tax Notes* in 2004) of the voting behavior of members of Congress who had signed the "no new taxes" pledge. That study found that among those who had signed the pledge, nearly all voted for the 2001 and 2003 tax cuts, 86 percent voted for Medicare Part D (the most expensive new federal entitlement in decades), and 90 percent voted for the pork-laden 2005 highway bill. Essentially, those who insisted that taxes must not be raised were the very people most willing to raise spending—evidence against the starve-the-beast hypothesis. Lastly, Gale suggested looking further into which tax features change in tax cuts and in subsequent tax increases. If the changes occur via marginal tax rates, which are cut first but end up rising later, that is inconsistent with optimal public finance theory, which shows that it is more efficient to keep tax rates constant than to shift them up and down.

Henry Aaron cited several established facts of political economy that, in addition to the inflation indexing of tax brackets mentioned by Perry, would make it difficult to find any statistically significant effects from four relatively small events. The first is that government spending as a share of GDP has been nearly flat for the past 50 years. Thus, the data likely contain too little variation to allow any strong effect to emerge. Second, the composition of spending has, in contrast, changed drastically, and these changes would likely mask the effect of modest fiscal policy changes. For example, defense spending declined from over 10 percent of GDP at the time of the Korean War to only 3 percent at its lowest point in the late 1990s. Non-defense discretionary spending declined significantly during the Reagan administration and has continued to decline since then. These spending changes imply large shifts in the political consensus on government spending over time and make it unlikely that any real impact of small tax changes on total spending at different points in time could be detected.