Macroeconomic Effects from Government Purchases and Taxes*

Robert J. Barro and Charles J. Redlick, Harvard University

November 2009

Abstract

For U.S. annual data that include WWII, the estimated multiplier for defense spending is 0.6-0.7 at the median unemployment rate (while holding fixed average marginal income-tax rates). There is some evidence that the spending multiplier rises with the extent of economic slack and reaches 1.0 when the unemployment rate is around 12%. We cannot estimate reliable multipliers for non-defense purchases because we lack good instruments. Since the defense-spending multiplier is typically less than one, greater spending tends to crowd out other components of GDP. The largest effects are on private investment, but non-defense purchases and net exports tend also to fall. The response of private consumer expenditure differs insignificantly from zero. For a sample that begins in 1950, increases in average marginal income-tax rates (measured by a newly constructed time series) have a significantly negative effect on GDP. When interpreted as a tax multiplier, the magnitude is around 1.1. We lack reliable statistical evidence on how the responses to tax changes divide up between substitution effects from changes in tax rates versus income effects from changes in tax revenue. The combination of the estimated spending and tax multipliers implies that, as the median unemployment rate, the estimated balanced-budget multiplier for defense spending is negative.

*This research was supported by a grant from the National Science Foundation. We particularly appreciate the assistance with the marginal tax-rate data from Jon Bakija and Dan Feenberg. We also appreciate research assistance from Andrew Okuyiga and comments from Greg Mankiw, David Romer, Jose Ursua, and participants in the Harvard macroeconomics seminar.

The global recession and financial crisis of 2008-09 have focused attention on fiscal-stimulus packages. These packages often emphasize heightened government purchases, predicated on the view (or hope) that expenditure multipliers are greater than one. The packages typically also include tax reductions, designed partly to boost disposable income and consumption (through income effects) and partly to stimulate work effort, production, and investment by lowering marginal income-tax rates (through substitution effects).

The empirical evidence on the response of real GDP and other economic aggregates to added government purchases and tax changes is thin. Particularly troubling in the existing literature is the basis for identification in isolating effects of changes in government purchases or tax revenue on real economic activity.

The present study uses long-term U.S. macroeconomic data to contribute to existing evidence along several dimensions. Spending multipliers are identified primarily from variations in defense spending, especially the changes associated with the buildups and aftermaths of wars. Tax effects are estimated mainly from changes in a newly measured time series on average marginal income-tax rates from federal and state income taxes and the social-security payroll tax. Some of the results attempt to differentiate substitution effects due to changes in tax rates from income effects due to changes in tax revenue.

Section I discusses the U.S. data on government purchases since 1914, with stress on the differing behavior of defense spending and non-defense purchases (by all levels of government). The variations up and down in defense outlays are particularly dramatic for World War II, World War I, and the Korean War. Section II describes a newly updated time series from 1912 to 2006 on average marginal income-tax rates from federal and state individual income taxes and the social-security payroll tax. Section III discusses the Romer and Romer (2008) data on "exogenous" changes in federal tax revenue since 1945. Section IV describes our empirical framework for assessing effects on real GDP from changes in government purchases, taxes, and other variables. Section V presents our empirical findings. The analysis deals with annual data ending in 2006 and starting in 1950, 1939, 1930, or 1917. Section VI summarizes the main findings and provides suggestions for additional research.

I. The U.S. History of Government Purchases: Defense and Non-defense

The graphs in Figure 1 show annual changes in per capita real defense or non-defense purchases (nominal outlays divided by the GDP deflator), expressed as ratios to the previous year's per capita real GDP.[1] The underlying data on government purchases are from the Bureau of Economic Analysis (BEA) since 1929 and, before that, from Kendrick (1961).[2] The data on defense spending apply to the federal government, whereas those for non-defense purchases pertain to all levels of government. Our analysis includes only government spending on goods and services, not transfers or interest payments. To get a long time series, we are forced to use annual data, because reliable quarterly figures are available only since 1947. The restriction to annual data has the virtue of avoiding problems related to seasonal adjustment.

The blue graph in Figure 1 makes clear the dominance of war-related variations in the defense-spending variable. For World War II, the value is 10.6% of GDP in 1941, 25.8% in

---

[1]Standard numbers for real government purchases use a government-purchases deflator that assumes zero productivity change for inputs bought by the government. We proceed instead by dividing nominal government purchases by the GDP deflator, effectively assuming that productivity advance is the same for publicly purchased inputs as it is in the private economy.

[2]The data since 1929 are the BEA's "government consumption and gross investment." This series includes an estimate of depreciation of public capital stocks (a measure of the rental income on publicly owned capital, assuming a real rate of return of zero on this capital).

1942, 17.2% in 1943, and 3.6% in 1944, followed by two negative values of large magnitude, -7.1% in 1945 and -25.8% in 1946. Thus, World War II provides an excellent opportunity to gauge the size of the government-purchases multiplier; that is, the contemporaneous effect of a change in government purchases on GDP. The favorable factors are:

- The principal changes in defense spending associated with World War II are plausibly exogenous with respect to the determination of GDP. (We neglect here a possible linkage between economic conditions and war probability.)
- These changes in defense spending are very large and include sharply positive and sharply negative values.
- Unlike the many countries that experienced major decreases in real GDP during World War II (see Barro and Ursua [2008, Table 7]), the United States did not have massive destruction of physical capital and suffered from only moderate loss of life. Hence, demand effects from defense spending should be dominant in the U.S. data.
- Because the unemployment rate in 1940 was still high, 9.4%, but then fell to a low of 1.0% in 1944, there is some information on how the size of the defense-spending multiplier depends on the amount of slack in the economy.

The U.S. time series contains two other war-related cases of large, short-term changes in defense spending. In World War I, the defense-spending variable (blue graph in Figure 1) equaled 3.5% in 1917 and 14.9% in 1918, followed by -7.9% in 1919 and -8.2% in 1920. In the Korean War, the values were 5.6% in 1951, 3.3% in 1952, and 0.5% in 1953, followed by -2.1% in 1954. As in World War II, the United States did not experience much destruction of physical capital and incurred only moderate loss of life during these wars. Moreover, the changes in defense outlays would again be mainly exogenous with respect to GDP.

In comparison to these three large wars, the post-1954 period features much more modest variations in defense spending. The largest values—1.2% in 1966 and 1.1% in 1967—occur during the early part of the Vietnam War. These values are much smaller than those for the Korean War; moreover, after 1967, the values during the Vietnam War become negligible (0.2% in 1968 and negative for 1969-71). After the end of the Vietnam conflict, the largest values of the defense-spending variable are 0.4-0.5% from 1982 to 1985 during the "Reagan defense buildup" and 0.3-0.4% in 2002-2004 during the post-2001 conflicts under George W. Bush.

The times of increased defense spending correspond well to the "defense-news" variable constructed by Ramey (2009, Table 2), who extended Ramey and Shapiro (1998) and Rotemberg and Woodford (1992, section V). The Ramey analysis uses a narrative approach, based on *Business Week* and other sources, to isolate major political events since 1939 that forecasted substantial increases in the present value of defense spending. In terms of the quantitative response of defense spending, however, the dominant events are World War II (1940-44) and the Korean War (1950). It seems unlikely that there is enough information in the variations in defense outlays after 1954 to get an accurate reading on the defense-spending multiplier.

The red graph in Figure 1 shows the movements in non-defense government purchases. Note particularly the values of 2.4% in 1934 and 2.5% in 1936, associated with the New Deal. Otherwise, the only clear pattern is a tendency for non-defense purchases to decline during major wars and rise in the aftermaths of these wars. For example, the non-defense purchases variable ranged from -1.0% to -1.2% between 1940 and 1943 and from 0.8% to 1.6% from 1946 to 1949. This broad pattern applies also to World War I, but the precise timing is puzzling. The most negative value for the non-defense purchases variable is -3.5% in 1919 (after the Armistice of 1918 but in the same year as the Versailles Treaty) and becomes positive only in the following year: 2.5% in 1920 and 2.0% in 1921.

It is hard to be optimistic about using the macroeconomic time series to isolate multipliers for non-defense government purchases. The first problem is that the variations in the non-defense variable are small compared to those in defense outlays. More importantly, the changes in non-defense purchases are likely to be endogenous with respect to GDP. That is, as with private consumption and investment, expansions of the overall economy likely induce governments, especially at the state and local level, to spend more on goods and services. As Ramey (2009, pp. 5-6) observes, outlays by state and local governments have been the dominant part of non-defense government purchases (since at least 1929). These expenditures—which relate particularly to education, public order, and transportation—are likely to respond to fluctuations in state and local revenue caused by changes in aggregate economic conditions. Whereas war and peace is a plausibly exogenous driver of defense spending, we lack similarly convincing instruments for isolating exogenous changes in non-defense purchases.

A common approach in the existing empirical literature, exemplified by Blanchard and Perotti (2002), is to include government purchases (typically, defense and non-defense combined) in a vector-auto-regression (VAR) system and then make identifying assumptions concerning exogeneity and timing. Typically, the government-purchases variable is allowed to move first, so that the contemporaneous associations with GDP and other macroeconomic aggregates are assumed to reflect causal influences from government purchases to the macro variables. This approach may be satisfactory for war-driven defense spending, but it seems problematic for other forms of government purchases.

II. Average Marginal Income-Tax Rates

Marginal income-tax rates have substitution effects that influence decisions on work versus consumption, the timing of consumption, investment, capacity utilization, and so on. Therefore, we would expect changes in these marginal tax rates to influence GDP and other macroeconomic aggregates. To gauge these effects at the macroeconomic level, we need measures of average marginal income-tax rates (AMTR)—or, possibly, other gauges of the distribution of marginal tax rates across economic agents.

Barro and Sahasakul (1983, 1986) used the Internal Revenue Service (IRS) publication *Statistics of Income, Individual Income Taxes* from various years to construct average marginal tax rates from the U.S. federal individual income tax from 1916 to 1983.[3] The Barro-Sahasakul series that we use weights each individual marginal income-tax rate by adjusted gross income or by analogous income measures available before 1944. The series takes account of non-filers, who were numerous before World War II. The 1986 study added the marginal income-tax rate from the social-security (FICA) tax on wages and self-employment income (starting in 1937 for the main social-security program and 1966 for Medicare). The analysis considered payments by employers, employees, and the self-employed and took account of the zero marginal tax rate for social security, but not Medicare, above each year's income ceiling. However, the earlier analysis and our present study do not allow for offsetting individual benefits at the margin from making social-security "contributions."

[3]The current federal individual income-tax system was implemented in 1913, following the ratification of the 16th Amendment, but the first detailed publication from the IRS applies mostly to 1916. We use IRS information from the 1916 book on tax-rate structure and numbers of returns filed in various income categories in 1914-15 to estimate average marginal income-tax rates. For 1913, we approximate based on tax-rate structure and total taxes paid.

We use the National Bureau of Economic Research (NBER) TAXSIM program, administered by Dan Feenberg, to update the Barro-Sahasakul data. TAXSIM allows for the increasing complexity of the federal individual income tax due to the alternative minimum tax (AMT), the earned-income tax credit (EITC), phase-outs of exemptions and deductions, and so on.[4] TAXSIM allows for the calculation of average marginal income-tax rates weighted in various ways—we focus on the average weighted by a concept of income that is close to labor income: wages, self-employment income, partnership income, and S-corporation income. Although this concept differs from the adjusted-gross-income measure used before (particularly by excluding most forms of capital income),[5] we find in the overlap from 1966 to 1983 that the Barro-Sahasakul and NBER TAXSIM series are highly correlated in terms of levels and changes. For the AMTR from the federal individual income tax, the correlations from 1966 to 1983 are 0.99 in levels and 0.87 in first differences. For the social-security tax, the correlations are 0.98 in levels and 0.77 in first differences. In addition, at the start of the overlap period in 1966, the levels of Barro-Sahasakul—0.217 for the federal income tax and 0.028 for social security—are not too different from those for TAXSIM—0.212 for the federal income tax and 0.022 for social security. Therefore, we are comfortable in using a merged series to cover 1912 to 2006. The merged data use the Barro-Sahasakul numbers up to 1965 (supplemented, as indicated in note 3, to include estimates for 1913-15) and the new values from 1966 on.

---

[4]The constructed average marginal income-tax rate therefore considers the impact of extra income on the EITC, which has become a major transfer program. However, the construct does not consider effects at the margin on other transfers, such as Medicaid, food stamps, and so on.

[5]The Barro-Sahasakul federal marginal tax rate does not consider the deductibility of part of state income taxes. However, since the average marginal tax rate from state income taxes up to 1965 does not exceed 0.016, this effect would be minor. In addition, the Barro-Sahasakul series treats the exclusion of employer social-security payments from taxable income as a subtraction from the social-security rate, rather than from the marginal rate on the federal income tax. However, this difference would not affect the sum of the marginal tax rates from the federal income tax and social security.

The new construct adds average marginal income-tax rates from state income taxes.[6] From 1979 to 2006, the samples of income-tax returns provided by the IRS to the NBER include state identifiers for returns with AGI under $200,000. Therefore, with approximations for allocating high-income tax returns by state, we were able to use TAXSIM to compute the AMTR from state income taxes since 1979. From 1929 to 1978, we used IncTaxCalc, a program created by Jon Bakija, to estimate marginal tax rates from state income taxes. To make these calculations, we combined the information on each state's tax code (incorporated into IncTaxCalc) with estimated numbers on the distribution of income levels by state for each year. The latter estimates used BEA data on per capita state personal income.[7] The calculations from TAXSIM and IncTaxCalc take into account that, for people who itemize deductions, an increase in state income taxes reduces federal income-tax liabilities.

Table 1 and Figure 2 show our time series from 1912 to 2006 for the overall average marginal-income tax rate and its three components: the federal individual income tax, social-security payroll tax (FICA), and state income taxes. In 2006, the overall AMTR was 35.3%, breaking down into 21.7% for the federal individual income tax, 9.3% for the social-security levy (inclusive of employee and employer parts), and 4.3% for state income taxes.[8] For year-to-year changes, the movements in the federal individual income tax usually dominate the variations in the overall marginal rate. However, rising social-security tax rates were important

---

[6]The first state income tax was implemented by Wisconsin in 1911, followed by Mississippi in 1912. A number of other states (Oklahoma, Massachusetts, Delaware, Missouri, New York, and North Dakota) implemented an income tax soon after the federal individual income tax became effective in 1913.

[7]Before 1929, we do not have the BEA data on income by state. For this period, we estimated the average marginal tax rate from state income taxes by a linear interpolation from 0 in 1910 (prior to the implementation of the first income tax by Wisconsin in 1911) to 0.0009 in 1929. Since the average marginal tax rates from state income taxes are extremely low before 1929, this approximation would not have much effect on our results.

[8]Conceptually, our "marginal rates" correspond to the effect of an additional dollar of income on the amounts paid of the three types of taxes. The calculations consider interactions across the levies; for example, part of state income taxes is deductible on federal tax returns, and the employer part of social-security payments does not appear in the taxable income of employees.

from 1971 to 1991. Note that, unlike for government purchases, the marginal income-tax rate for each household really is an annual variable; that is, the same rate applies at the margin to income accruing at any point within a calendar year. Thus, for marginal tax-rate variables, it would not be meaningful to include variations at a quarterly frequency.[9]

Given the focus on wage and related forms of income, we view our constructed average marginal income-tax rate as applying most clearly to the labor-leisure margin, including choices of work effort over time. However, unmeasured forms of marginal tax rates (associated with corporate income taxes, sales and property taxes, means-testing for transfer programs, and so on) might move in ways correlated with the measured AMTR.

Many increases in the federal average marginal income-tax rate involve wartime, including World War II (a rise in the rate from 3.8% in 1939 to 25.7% in 1945, reflecting particularly the extension of the income tax to most households), World War I (an increase from 0.6% in 1914 to 5.4% in 1918), the Korean War (going from 17.5% in 1949 to 25.1% in 1952), and the Vietnam War (where "surcharges" contributed to the rise in the rate from 21.5% in 1967 to 25.0% in 1969). The federal AMTR tended to fall during war aftermaths, including the declines from 25.7% in 1945 to 17.5% in 1949, 5.4% in 1918 to 2.8% in 1926, and 25.1% in 1952 to 22.2% in 1954. No such reductions applied after the Vietnam War.

A period of rising federal tax rates prevailed from 1971 to 1978, with the AMTR from the individual income tax increasing from 22.7% to 28.4%. This increase reflected the shifting of households into higher rate brackets due to high inflation in the context of an un-indexed tax

[9]However, even if the tax-rate structure is set by the beginning of year t, information about a household's marginal income-tax rate for year t arrives gradually during the year—as the household learns about its income, deductions, etc.. This type of variation in perceived marginal income-tax rates within a year is clearly endogenous with respect to the determination of income (and real GDP) during the year.

system. Comparatively small tax-rate hikes include the Clinton increase from 21.7% in 1992 to 23.0% in 1994 (and 24.7% in 2000) and the rise under George H.W. Bush from 21.7% in 1990 to 21.9% in 1991. Given all the hype about Bush's violation of his pledge, "read my lips, no new taxes," it is surprising that the AMTR rose by only two-tenths of a percentage point in 1991.

Major cuts in the AMTR occurred under Reagan (25.9% in 1986 to 21.8% in 1988 and 29.4% in 1981 to 25.6% in 1983), George W. Bush (24.7% in 2000 to 21.1% in 2003), Kennedy-Johnson (24.7% in 1963 to 21.2% in 1965), and Nixon (25.0% in 1969 to 22.7% in 1971, reflecting partly the introduction of a maximum marginal rate of 60% on earned income in 1971).

Also noteworthy is the AMTR from the federal income tax during the Great Depression. The rate fell from 4.1% in 1928 to 1.7% in 1931, mainly because falling incomes within a given tax structure pushed people into lower rate brackets. Then, particularly because of attempts to balance the federal budget by raising taxes under Hoover and Roosevelt, the AMTR rose to 5.2% in 1936.

Although social-security tax rates have less high-frequency variation, they increase sharply in some periods. The AMTR from social security did not change greatly from its original value of 0.9% in 1937 until the mid 1950s but then rose to 2.2% in 1966. The most noteworthy period of rising average marginal rates is from 1971—when it was still 2.2%—until 1991, when it reached 10.8%. Subsequently, the AMTR remained reasonably stable, though it fell from 10.2% in 2004 to 9.3% in 2006 (due to expansion of incomes above the social-security ceiling).

The marginal rate from state income taxes rose from less than 1% up to 1956 to 4.1% in 1977 and has since been reasonably stable. We have concerns about the accuracy of this series, particularly before 1979, because of missing information about the distribution of incomes by state. However, the small contribution of state income taxes to the overall AMTR suggests that this measurement error would not matter a lot for our main findings. The results that we report later based on the overall AMTR turn out to be virtually unchanged if we eliminate state income taxes from the calculation of the overall marginal rate.

Our key identifying assumption is that changes in average marginal income-tax rates lagged one or more years can be satisfactorily treated as pre-determined with respect to per capita GDP. Note that the regressions used to determine per capita GDP in year t include a measure of aggregate economic conditions in year t-1 (the lagged unemployment rate, though we also considered variables based on the history of per capita real GDP). Thus, a contemporaneous association between economic conditions and changes in marginal income-tax rates need not cause problems in isolating effects from lagged tax-rate changes on GDP.

One way to evaluate our identifying assumption is from the perspective of the tax-smoothing approach to fiscal deficits (Barro [1979, 1990]; Aiyagari, Marcet, Sargent, and Seppala [2002]). A key implication of this approach is the Martingale property for tax rates: future changes in tax rates should not be predictable based on information available at date t. Redlick (2009) tests this hypothesis for the data on the overall average marginal income-tax rate shown in Table 1. He finds that the Martingale property is a good first-order approximation but that some economic variables have small, but statistically significant, predictive content for future changes in the AMTR.

If tax smoothing holds as an approximation, then the change in the tax rate for year t, τ_t-τ_{t-1}, would reflect mainly information arriving during year t about the future path of the ratio of real government expenditure, G_{t+T} (inclusive here of transfer payments), to real GDP, Y_{t+T}. Information that future government outlays would be higher in relation to GDP would cause an increase in the current tax rate. For our purposes, the key issue concerns the effects of changes in expectations about future growth rates of GDP. Under tax-smoothing, these changes would not impact the current tax rate if the shifts in expected growth rates of GDP go along with corresponding changes in expected growth rates of government spending. Thus, our identifying assumption is that any time-varying expectations about growth rates of future GDP do not translate substantially into changes in the anticipated future path of G/Y, and, therefore, do not enter significantly into the determination of tax rates.

III. The Romer-Romer Tax-Change Series

Romer and Romer (2008) use a narrative approach based on Congressional reports and other sources to assess all significant federal tax legislation from 1945 to 2007. They gauge each tax change by the size and timing of the *intended* effect on federal tax revenue. Thus, in contrast to the marginal income-tax rates discussed before, their focus is on income effects related to the federal government's tax collections. In practice, however, their tax-change variable has a high positive correlation with shifts in marginal income-tax rates; that is, a rise in their measure of intended federal receipts (expressed as a ratio to past GDP) usually goes along with an increase in average marginal income-tax rates, and vice versa.[10] Consequently, the Romer-Romer

[10]A major counter-example is the Reagan tax cut of 1986, which reduced the average marginal tax rate from the federal individual income tax by 4.2 percentage points up to 1988. Because this program was designed to be

variable or our marginal tax-rate variable used alone would pick up a combination of income and substitution effects. However, when we include the two tax measures together, we can reasonably view the Romer-Romer variable as isolating income effects,[11] with the marginal tax-rate variable isolating mainly substitution effects.[12]

Because the Romer-Romer variable is based on planned changes in federal tax revenue, assessed during the prior legislative process, the variable avoids the obvious contemporaneous endogeneity of realized tax revenue with respect to realized GDP. Thus, the major remaining concern about endogeneity involves politics; that is, Congressional legislation on taxes may involve feedback from past or prospective economic developments. To deal with this concern, Romer and Romer divide each tax bill (or sometimes parts of bills) into four bins, depending on what the narrative evidence reveals about the underlying motivation for the tax change. The four categories are (Romer and Romer [2008, abstract]): "… responding to a current or planned change in government spending, offsetting other influences on economic activity, reducing an inherited budget deficit, and attempting to increase long-run growth." They classify the first two bins as endogenous and the second two as exogenous.[13] Although we have reservations about this approach as a remedy for endogeneity, we find that the main results that use the Romer-Romer variable as an instrument for changes in federal revenue are similar whether we include

---

revenue neutral (by closing "loopholes" along with lowering rates), the Romer-Romer variable shows only minor federal tax changes in 1987 and 1988.

[11]Ricardian equivalence does not necessarily imply that these income effects would be nil. For example, a high value of the Romer-Romer tax variable might signal an increase in the ratio of expected future government spending to GDP, thereby likely implying a negative income effect.

[12]However, for a given ratio of federal revenue to GDP, an increase in the average marginal tax rate might indicate that the government had shifted toward a less efficient tax-collection system, thereby implying a negative income effect.

[13]The first bin does not actually involve endogeneity of tax changes with respect to GDP but instead reflects concern about a correlated, omitted variable—government spending—that may affect GDP. Empirically, the main cases of this type in the Romer-Romer sample associate with variations in defense purchases during and after wars, particularly the Korean War. The straightforward remedy for this omitted-variable problem is to include a measure of defense spending when attempting to explain variations in GDP. Romer and Romer (2009) include broad measures of government spending in parts of their analysis.

all four of their bins or only the two labeled as exogenous. Thus, the principal findings do not depend on their choices of which tax changes to assign to which bins or on which bins to regard as endogenous.

Romer and Romer (2008, Table 1) provide data at a quarterly frequency, but our analysis uses these data only at an annual frequency, thus conforming to our treatment for government purchases and average marginal income-tax rates. For most of our analysis, the key identifying assumption is that the Romer-Romer tax-change variable lagged one or more years can be treated as exogenous with respect to GDP. Thus, we can satisfactorily use the Romer-Romer variable as an instrument for lags of a variable constructed from changes in overall federal revenue. In the analysis of the contemporaneous relation between tax-revenue changes and GDP, we assume further that the Romer-Romer variable can be used satisfactorily as an instrument for this contemporaneous tax change.

IV. Framework for the Empirical Analysis

We estimate equations for the growth rate of per capita real GDP in the form:

$$(1) \qquad (y_t - y_{t-1})/y_{t-1} = \beta_0 + \beta_1 \cdot (g_t - g_{t-1})/y_{t-1} + \beta_2 \cdot (\tau_{t-1} - \tau_{t-2}) + \text{other variables},$$

where y_t is per capita real GDP for year t, g_t is per capita real government purchases for year t, and τ_t is a tax rate for year t. The form of equation (1) implies that the coefficient β_1 is the government-spending multiplier.[14] We are particularly interested in whether β_1 is greater than

[14]Note that the variable y_t is the per capita value of nominal GDP divided by the implicit GDP deflator, P_t (determined by the BEA from chain-weighting for 1929-2006). The variable g_t is calculated analogously as the per capita value of nominal defense spending divided by the same P_t. Therefore, the units of y and g are comparable, and β_1 reveals the effect of an extra unit of defense spending on GDP.

zero, greater than one, and larger when the economy has more slack. We gauge the last effect by adding to equation (1) an interaction between the variable $(g_t – g_{t-1})/y_{t-1}$ and the lagged unemployment rate, U_{t-1}, which is an indicator of the amount of slack in the economy.

When g_t in equation (1) corresponds to defense spending, we use as an instrument the same variable interacted with a dummy for war years (1914-20, 1939-46, 1950-54, 1966-71; see the notes to Table 2). Since the main movements in defense spending are war related (Figure 1), we end up with similar results—especially in samples that cover WWII—if we include the defense-spending variable itself in the instrument list. We also consider representing g_t by non-defense purchases, but this setting leads to problems because of the lack of convincing instruments.

If τ_t in equation (1) is our average marginal income-tax rate, corresponding to the data in Table 1, we expect $\beta_2<0$ because of adverse effects on the incentive to work and, therefore, produce. As noted before, we assume that the first annual lag of the change in the tax rate, $\tau_{t-1}-\tau_{t-2}$, can be treated as pre-determined with respect to GDP. We also assess whether additional lags of tax-rate changes influence GDP. In addition, we attempt to estimate the effect on GDP from the contemporaneous tax-rate change, $\tau_t-\tau_{t-1}$, by using as an instrument for τ_t the average marginal income-tax rate that would have applied at the previous year's incomes.

If τ_t represents tax revenue collected as a share of GDP, we may get $\beta_2<0$ in equation (1) because of income effects on aggregate demand. In this context, we instrument by using the data on tax-revenue changes from Romer and Romer (2008, Table 1). Again, we consider additional lags and also attempt to estimate contemporaneous effects. We discuss later theoretical issues related to income effects in the determination of GDP.

As mentioned before, the other variables included in equation (1) include indicators of the lagged state of the business cycle. This inclusion is important because, otherwise, the fiscal variables might just pick up the dynamics of the business cycle. In the main analysis, we include the first annual lag of the unemployment rate, U_{t-1}. Given a tendency for the economy to recover from recessions, we expect a positive coefficient on U_{t-1}. We also consider the inclusion of the first lag of the dependent variable, as well as the deviation of the previous year's log of per capita real GDP from its "trend." However, these alternative variables turn out not to be statistically significant in equation (1), once U_{t-1} is included.

Many additional variables could affect GDP growth in equation (1). However, as Romer and Romer (2009) have argued, omitted variables that are orthogonal to the fiscal variables (when lagged business-cycle indicators are included) would not bias the estimated effects of the fiscal variables. The main effect that seemed important to add—particularly for samples that include the Great Depression of 1929-33—is an indicator of monetary/credit conditions. A variable that works well is the quality spread in interest rates; specifically, the gap between the yield to maturity on long-term Baa-rated corporate bonds and that on long-term U.S. government bonds. We think that this yield spread captures distortions in credit markets, and the square of the spread works in a reasonably stable way in the explanation of GDP growth in equation (1). Since the contemporaneous spread would be endogenous with respect to GDP, we instrument with the first annual lag of the spread variable. That is, given the lagged business-cycle indicators already included, we treat the lagged yield spread as pre-determined with respect to GDP. Although the inclusion of this credit variable likely improves the precision of our estimates of fiscal effects, we get broadly similar results if the credit variable is omitted.

Our discussion of instruments reflected mostly concerns about the potential simultaneity between GDP and the government spending and tax variables. However, another issue is measurement error in the right-hand-side variables, a particular concern because government purchases—which appear on the right-hand side of equation (1)—are also a component of GDP, which appears on the left-hand side. Consider a simplified version of equation (1):

$$y_t = \beta_0 + \beta_1 \cdot g_t + \text{error term}. \tag{2}$$

GDP equals government purchases plus the other parts of GDP (consumer spending, gross private domestic investment, net exports). If we label these other parts as x_t, we have:

$$y_t = g_t + x_t. \tag{3}$$

Consider estimating the equation:

$$x_t = \alpha_0 + \alpha_1 \cdot g_t + \text{error term}, \tag{4}$$

where α_1 (if negative) gauges the crowding-out effect of g_t on other parts of GDP. Measurement error in g_t tends to bias usual estimates of α_1 toward zero. However, we also have from comparing equation (2) with a combination of equations (3) and (4) that the estimate of β_1 has to coincide with 1 + estimate of α_1. Therefore, a bias in the estimate of α_1 toward zero corresponds to a bias in the estimate of β_1 toward one. Thus, if $\alpha_1<0$, spending multipliers tend to be over-estimated. To the extent that the instruments have independent measurement error, the instrumentation would help to alleviate not only simultaneity bias but also this potential bias from measurement error.

V. Empirical Results

Table 2 shows two-stage least-squares regressions with annual data of the form of equation (1). The samples all end in 2006. (Data on marginal income-tax rates for 2007 will soon be available from TAXSIM.) The starting year is 1950 (including the Korean War), 1939 (including WWII), 1930 (including the Great Depression), or 1917 (including WWI and the 1921 contraction). See the notes to Table 2 for details on the instrument lists.

A. Defense-Spending Multipliers

For all samples, the estimated defense-spending multiplier in Table 2 is significantly greater than zero, with a p-value less than 0.01.[15] For the 1950 sample, the estimated coefficient, 0.77 (s.e.=0.28), is insignificantly different from one (p-value=0.41). For samples that start in 1939 or earlier, and thereby include WWII, the estimated multiplier is much more precisely determined. These estimates are significantly greater than zero and significantly less than one with very low p-values. In the last three columns of the table, the estimated coefficient is between 0.64 and 0.67, with standard errors of 0.10 or less.[16]

As discussed before, each regression includes the lagged unemployment rate, U_{t-1}, to pick up business-cycle dynamics. The estimated coefficients on U_{t-1} in Table 2 are significantly positive for each sample, indicating a tendency for the economy to recover by growing faster when the lagged unemployment rate is higher. We also tried business-cycle variables measured by the lag of the dependent variable or the lag of the deviation of the log of per capita GDP from its trend (gauged by a one-sided Hodrick-Prescott filter). In all cases, the estimated coefficients

[15]See Barro (1984, pp. 312-315) for an earlier analysis of the effects of wartime spending on output. Hall (2009, Table 1) also presents estimates of defense-spending multipliers associated with wars.
[16]If we add an additional lag of the defense-purchases variable, the estimated coefficient differs insignificantly from zero for all samples, and the estimated coefficient and standard error for the contemporaneous variable change little.

of these alternative variables were not statistically significantly different from zero, whereas the estimated coefficient on the lagged unemployment rate remained significantly positive.

The interaction term, (Δg-defense)* U_{t-1} in Table 2, indicates how the defense-spending multiplier depends on the amount of slack in the economy, gauged by the lagged unemployment rate. The variable U_{t-1} in this interaction term is measured as a deviation from the median unemployment rate of 0.0557 (calculated from 1914 to 2006). Therefore, the coefficient on the variable Δg-defense reveals the multiplier when the lagged unemployment rate is at the sample median, and the interaction term indicates how the defense-spending multiplier varies as U_{t-1} deviates from its median. For the 1950 sample, there is insufficient variation in Δg-defense to estimate the interaction term with any precision.[17] For samples that start in 1939 or earlier, and thereby include WWII, it is possible to get meaningful estimates for the interaction term. For the 1939 sample in column 3 of the table, the estimated coefficient is 5.1 (s.e.=2.2), which is significantly different from zero with a p-value less than 0.05. Since the estimated coefficient on Δg-defense is 0.67 (0.07), the point estimates imply a defense-spending multiplier of 0.67 at the median unemployment rate of 5.6%. But the estimated multiplier rises by around 0.10 for each 2 percentage points by which the unemployment rate exceeds 5.6%. Hence, the estimated multiplier reaches 1.0 when the unemployment rate gets to about 12%. Conversely, the estimated multiplier is less than 0.67 when the unemployment rate is below its median.

As already noted, the wartime experiences include substantially positive and negative values for Δg-defense. The estimates shown in Table 2 assume that the effects on GDP are the same for increases and decreases in spending, notably, for war buildups and demobilizations.

[17] If the interaction term is added, along with the variable $\Delta[\tau^*(\text{g-def/y})](-1)$, to the 1950 equation in Table 2, the estimated coefficients are -23.6 (s.e.=25.3) on the interaction term and 0.58 (0.37) on Δg-defense.

Tests of this hypothesis are accepted at high p-values for the various samples considered in the table. For the 1939 sample (corresponding to column 3 of Table 2), the estimated coefficients are 0.69 (s.e.=0.08) for positive values of Δg-defense and 0.54 (0.17) for negative values, with a p-value of 0.44 applying to a test of equal coefficients. For the 1930 and 1917 samples, the estimated coefficients for positive and negative values of Δg-defense are even closer, and the p-values exceed 0.9. For the 1950 sample (corresponding to column 1 of Table 2), the estimated coefficients are 0.80 (0.34) for positive values of Δg-defense and 1.16 (0.83) for negative values, with a p-value of 0.71 applying to the hypothesis of equal coefficients.[18] Thus, the evidence is consistent with the condition that spending multipliers are the same for increases and decreases in defense spending.

The two-stage least-squares estimates of the defense-spending multipliers in Table 2 rely only on the variations in defense spending associated with major wars (including a year of war aftermath in each case[19]). Hence, the estimation excludes more moderate—possibly also exogenous—variations in defense spending, such as those associated with the defense buildups under Reagan and George H.W. Bush. If we can think of all of the variations in defense spending as exogenous, we can improve on the estimates by allowing for all of the sample variations in the variable Δg-defense. However, because the wartime observations already capture the principal fluctuations in defense spending, the results on defense-spending multipliers change little if we modify the instrument list to include Δg-defense. For the 1950 sample, the estimated coefficient on Δg-defense becomes 0.65 (s.e.=0.26), somewhat below the

[18]For the samples starting 1939 or earlier, the equations include the interaction term between Δg-defense and the lagged unemployment rate, with the coefficient constrained to be the same for positive and negative values of Δg-defense. If these coefficients are allowed to differ for positive and negative values, we still accept with high p-values the hypothesis of equal coefficients (for Δg-defense and for Δg-defense interacted with the lagged unemployment rate).

[19]We treated WWI as ending in 1919 and thereby included 1920 as the year of war aftermath. However, the results change little if we instead treat the war as ending in 1918, so that 1919 is the year of war aftermath.

value in Table 2, column 1, and closer to the estimates for the longer samples. For the samples that start in 1939 or earlier, the change in the instrument list has a negligible impact. For example, for the 1939 sample in column 3, the estimated coefficient on Δg-defense becomes 0.65 (s.e.=0.07) and that on the interaction with U_{t-1} becomes 4.8 (2.2).

For samples that start after the Korean War, such as 1954-2006, the estimated defense-spending multiplier differs insignificantly from zero. For example, if the instrument list excludes Δg-defense but includes this variable interacted with war years, as in Table 2, the estimated multiplier is 0.94 (s.e.=0.57). The estimated coefficients of the other variables are close to those for the 1950-2006 sample in Table 2, column 1. If we include Δg-defense on the instrument list, the estimated multiplier for the 1954-2006 sample becomes 0.53 (s.e.=0.51), and the other estimated coefficients are close to those in Table 2, column 1. The main conclusion is that, with the Korean War excluded, there is insufficient variation in defense spending to get an accurate reading on the spending multiplier.

B. Effects of Marginal Income-Tax Rates

The equations in Table 2 include the lagged change in the average marginal income-tax rate, Δτ(-1). For the sample that starts in 1950 in column 1, the estimated coefficient is -0.58 (s.e.=0.21), which is significantly negative with a p-value less than 0.01. Thus, the estimate is that a cut in the AMTR by 1 percentage point raises next year's per capita GDP by around 0.6%.

We can compare our estimated effect of tax-rate changes on GDP to microeconomic estimates of labor-supply elasticities, as summarized by Chetty (2009, Table 2). His results apply to elasticities of taxable income with respect to 1-τ, where τ is the marginal income-tax rate. For 13 studies (excluding the two based on macroeconomic data), the mean of the estimated elasticities, η, is 0.39. The implied effect of a change in τ on the log of taxable income entails multiplying η by $-1/(1-\tau)$. Our AMTR has a mean of 0.33 from 1950 to 2006. If we evaluate the microeconomic estimates at the point where $\tau=0.33$, the effect from a change in τ on the log of taxable income equals $-\eta/(1-\tau) = -0.39\cdot(1.49) = -0.58$. If GDP moves in the same proportion as taxable income, this number should correspond to the estimated coefficient on the variable $\Delta\tau(-1)$ in Table 2. Since that point estimate happens also to be -0.58, there does turn out to be a close correspondence. That is, our macroeconomic estimate of the response of GDP to a change in the AMTR accords with typical microeconomic estimates of labor-supply elasticities.

The estimated coefficient of -0.58 on $\Delta\tau(-1)$ in Table 2, column 1, does not correspond to a usual tax multiplier for GDP. Our results connect the change in GDP to a shift in the average marginal income-tax rate, not to variations in tax revenue, per se. As an example, for a revenue-neutral change in the tax-rate structure (corresponding to the plan for the 1986 tax reform), the conventional tax multiplier would be minus infinity. However, the typical pattern is that increases in the ratio of tax revenue to GDP accompany increases in the AMTR, and vice versa. We can, therefore, compute a tax multiplier that gives the ratio of the change in GDP to the change in tax revenue when we consider the typical relation of tax revenue to the AMTR.

Let T be the average tax rate, gauged by the ratio of federal revenue to GDP. Then real federal revenue is T·GDP. The change in this revenue, when expressed as a ratio to GDP, is:

$$\Delta(\text{revenue})/GDP = T\cdot\Delta GDP/GDP + \Delta T. \tag{5}$$

The estimates in Table 2, column 1, suggest $\Delta GDP/GDP = -0.58\cdot\Delta\tau$, where τ is the average marginal income-tax rate (applying here to federal taxes).

We now have to connect the change in the average tax rate, ΔT, to the change in the AMTR, $\Delta\tau$. From 1950 to 2006, the average of T (nominal federal revenue divided by nominal GDP) is 0.182. The average for τ (based only on the federal individual income tax plus social security) is 0.297. We take as a typical relation that an increase in τ by one percentage point is associated with an increase in T by 0.613 of a percentage point (the ratio of 0.182 to 0.297). If we substitute this result and the previous one for $\Delta GDP/GDP$ into equation (2), we get

$$\Delta(\text{revenue})/GDP = (-0.58\cdot T + 0.613)\cdot\Delta\tau. \tag{6}$$

If we evaluate equation (6) at the sample average for T of 0.182, we get

$$\Delta(\text{revenue})/GDP = 0.51\cdot\Delta\tau. \tag{7}$$

Finally, if we use equation (7), we get that the "tax multiplier" is

$$\begin{aligned} \Delta GDP/\Delta(\text{revenue}) &= [\Delta GDP/GDP]/[\Delta(\text{revenue})/GDP] \\ &= -0.58\cdot\Delta\tau/0.51\cdot\Delta\tau = -1.14. \end{aligned} \tag{8}$$

Thus, the results correspond to a conventional tax multiplier of around -1.1.

Samples that start earlier than 1950 show less of an impact from $\Delta\tau(-1)$ on GDP growth; for the sample that starts in 1939, in Table 2, column 2, the estimated coefficient is 0.10 (s.e.=0.16). One problem is that the dramatic rise in the AMTR from 0.069 in 1940 to 0.263 in 1944 does not match up well with the strong GDP growth during WWII (even when one factors in the positive effect from the rise in defense purchases from 1941 to 1944).

Aside from the dramatic rise in defense spending and tax rates, World War II featured an expansion of direct governmental control over the allocation of private resources, including the large-scale military draft and mandates on production. Similar forces arose during World War I. For present purposes, the key point is that the sensitivity of GDP to the marginal income-tax rate is likely to be weaker the larger the fraction of GDP that is allocated by governmental directive. To gauge this effect empirically, we use the ratio of defense purchases to GDP to proxy for the extent of command and control. An interaction term between the defense-purchases ratio and the marginal tax rate, τ_t, then picks up the attenuation of the impact of tax rates on GDP. The regressions shown in columns 3-5 of Table 2 include the lagged change in this interaction term, along with the lagged change in the AMTR, $\Delta\tau(-1)$. For example, for the sample that starts in 1939 in column 3, the estimated coefficient of the interaction term is significantly positive, whereas that for $\Delta\tau(-1)$ is negative (though not statistically significant on its own). The two tax-rate variables are jointly significant with a p-value of 0.03. The point estimates imply that the net effect of a lagged change in the AMTR on GDP growth is negative in most situations but becomes nil under extreme wartime conditions (when the defense purchases ratio reaches 0.36, compared to the peak of 0.44 in 1944).

C. The Yield Spread

Table 2 shows that the estimated coefficient on the yield-spread variable is significantly negative for each sample. Hence, a larger gap between Baa and U.S. Treasury bond yields predicts lower GDP growth. The magnitude of the coefficient is similar across samples, except when the sample goes back to include the Great Depression. The inclusion of these years raises the magnitude of the estimated coefficient (to fit the extremely low growth rates of 1930-33). For example, for the 1930 sample, if one allows for two separate coefficients on the yield-spread variable, the estimated coefficients are -111.3 (s.e.=14.9) for 1930-38 and -48.0 (28.8) for 1939-2006. (This regression includes separate intercepts up to 1938 and after 1938.) The two estimated coefficients on the yield-spread variable differ significantly with a p-value of 0.051. Similar results apply if the sample starts in 1917.

For our purposes, an important result is that the estimated coefficients on the defense-spending and tax-rate variables do not change a lot if the equation excludes the yield-spread variable. For example, for the 1939 sample (corresponding to column 3 of Table 2), the estimated coefficients become 0.71 (s.e.=0.08) on Δg-defense, 0.47 (0.14) on U_{t-1}, 6.1 (2.4) on (Δg-defense)* U_{t-1}, -0.31 (0.21) on Δτ(-1), and 0.72 (0.31) on Δ[τ*(g-def/y)](-1). Thus, the main change is for the variable U_{t-1}, which picks up business-cycle dynamics. For the 1930 and 1917 samples, the main changes are again in the estimated coefficients on U_{t-1}. For the 1950 sample (corresponding to column 1 of Table 2), the deletion of the yield-spread variable raises the magnitudes of the estimated fiscal effects. The estimated coefficients become 1.01 (s.e.=0.30) on Δg-defense, 0.40 (0.19) on U_{t-1}, and -0.69 (0.23) on Δτ(-1).

Since we think that holding fixed a measure of credit conditions sharpens the estimates of the effects of the fiscal variables, we focus on the results presented in Table 2. However, the robustness of the main results to deletion of the interest-rate spread variable heightens our confidence in the findings with regard to fiscal effects.

D. More Results on Government Purchases

The results in Table 2 seem to provide reliable estimates of the multiplier for defense spending, particularly for long samples that include WWII. This multiplier is estimated to be around 0.6-0.7 at the median unemployment rate. However, to evaluate typical fiscal-stimulus packages, we are more interested in multipliers associated with non-defense purchases. The problem, already mentioned, is that this multiplier is hard to estimate because observed movements in non-defense purchases are likely to be endogenous with respect to GDP. Given this problem, it is important to know whether the defense-spending multiplier provides an upper or lower bound for the non-defense multiplier.

Consider, from a theoretical standpoint, how the multiplier for non-defense purchases relates to that for defense spending. One point is that the movements in defense spending, driven to a considerable extent by war and peace, are likely to be more temporary than those in non-defense purchases, a property stressed by Barro (1981). However, in a baseline model worked out by Barro and King (1984), the extent to which a change in government purchases is temporary has no impact on the size of the multiplier. The key features of this model are the existence of a representative agent with time-separable preferences over consumption and leisure, the absence of durable goods, lump-sum taxation, and "market clearing." Parts of the analysis rely also on an assumption that consumption and leisure are both normal goods.

Consider now some realistic deviations from the Barro-King setting. With distorting taxes, the government's incentive to tax-smooth implies that tax rates tend to rise less when a given size change in government purchases is more temporary.[20] Since higher tax rates discourage economic activity, a temporary change in purchases, as in wartime, tends to have a larger effect on GDP than an equal-size permanent change. However, in the empirical analysis, we already held constant measures of marginal income-tax rates. Within a Barro-King setting, but with distorting taxes, the spending multiplier estimated for given tax rates should not depend on whether the change in spending is temporary or permanent.

[20]In World War II, defense spending was temporarily high, implying, on tax-smoothing grounds, that much of the added spending was deficit financed. However, the war likely also substantially raised the anticipated long-run ratio of government spending to GDP. This change motivated a rise in tax rates—the average marginal tax rate from the federal individual income tax increased from 3.8% in 1939 to 25.7% in 1945 (see Table 1).

If we allow for durable goods and, hence, investment, the crowding out of investment tends to be larger when the increase in government purchases is more temporary. The decline in investment means that consumption and leisure fall by less than otherwise and, hence, that work effort rises less by less than otherwise. If the change in production depends in the short run only on the change in labor input, the multiplier is smaller when the increase in purchases is more temporary (as in wartime). However, this conclusion need not hold if we allow for variable capital utilization, which tends to expand more when the increase in purchases is more temporary.

As already mentioned, wars typically feature command-and-control techniques, including rationing private expenditure on goods and services, drafting people to work in the military, and forcing companies to produce tanks rather than cars (all without reliance on explicit prices). Rationing tends to hold down private demand for goods and services, thereby making the spending multiplier smaller than otherwise. However, mandated increases of production and labor input tend to raise the aggregate supply of goods and services, thereby increasing the spending multiplier. An offsetting force is that government-mandated output may be under-valued in the computation of GDP—if tanks carry unrealistically low "prices" and if draftee wages (including provision of food, housing, etc.) fall short of private-sector wages. A final consideration is that, during a popular war such as WWII, patriotism likely shifts labor supply outward. This boost to the aggregate supply of goods and services tends to make the wartime multiplier comparatively large.

Our conjecture is that, because of command-and-control and patriotism considerations, the defense-spending multiplier associated with war and peace would exceed the non-defense multiplier. In this case, the defense-spending multiplier—for which we have good estimates—

would provide an upper bound for the non-defense multiplier. However, since the comparison between the two multipliers is generally ambiguous on theoretical grounds, it would obviously be desirable to have direct, reliable estimates of the non-defense multiplier.

Table 3 shows regression results when a non-defense purchases variable—constructed analogously to the defense-spending variable—is added to the regressions shown in Table 2. Crucially, the instrument list includes the contemporaneous value of the non-defense purchases variable. The estimated multiplier is large and significantly different from zero for the 1950 sample—the estimated coefficient is 1.93 (s.e.=0.90). However, the estimated coefficients are insignificantly different from zero for the other samples, which start in 1939, 1930, or 1917. A reasonable interpretation is that the estimated coefficients on the non-defense purchases variable pick up primarily reverse causation from GDP to government spending but that the nature of this reverse linkage varies over time.

To illustrate the potential for spuriously high estimated multipliers due to endogeneity, in Table 3, columns 5 and 6, we replaced the non-defense purchases variable by analogously constructed variables based on sales of two large U.S. corporations with long histories—General Motors and General Electric. In column 5, the estimated "multiplier" based on GM sales for a 1950 sample is 3.4 (s.e.=0.8). That is, this "multiplier" exceeds three. For GE sales—which are less volatile than GM's but also more correlated with GDP—the result is even more extreme, 17.9 (4.6). Moreover, unlike for non-defense purchases, the estimated GM and GE coefficients are similar over longer samples. Clearly, the GM and GE estimates reflect mainly reverse causation from GDP to sales of individual companies. We think that a similar perspective applies to the apparent multiplier above one estimated for non-defense purchases for the 1950 sample in column 1.

The problem is that, lacking good instruments, we cannot estimate multipliers satisfactorily for non-defense purchases.[21] The vector-autoregression (VAR) literature makes identifying assumptions based typically on changes in government purchases being pre-determined within a quarter (see, for example, Blanchard and Perotti [2002]). This procedure, which has been criticized by Ramey (2009), corresponds in annual data to the kinds of estimates for non-defense purchases shown in columns 1-4 of Table 3. We think these results are not meaningful. Probably a more productive avenue is a search for satisfactory instruments for changes in non-defense government purchases for the United States or other countries. Political variables related to spending programs may provide satisfactory instruments. Applications of this approach to cross-sectional allocations of U.S. New Deal spending include Wright (1974) and Fishback, Horrace, and Kantor (2005). An application to Swedish municipalities is Johansson (2003).

E. Effects on Components of GDP

Table 4 returns to the setting of Table 2, but with the dependent variables corresponding to changes in components of GDP, rather than overall GDP. For example, for personal consumer expenditure, the dependent variable is the difference between this year's per capita real expenditure (nominal spending divided by the GDP deflator) and the previous year's per capita real expenditure, all divided by the prior year's per capita real GDP. The same approach applies in Table 4 to gross private domestic investment, I, non-defense government purchases, and net exports. Note that this method relates spending on the various parts of GDP to defense spending and the other right-hand-side variables considered in Table 2 but does not allow for effects from

[21]Lagged changes in the non-defense purchases variable provide only weak instruments and tend to produce estimated coefficients with high standard errors. With the first lag used as an instrument for the 1950 sample, the estimated coefficient of the non-defense-purchases variable is 5.6 (s.e.=3.1). With the second lag used as an instrument, the result is 1.9 (3.1).

changing relative prices, for example, for C versus I. This spending approach is valid even though the GDP deflator comes from chain-weighting—since all nominal magnitudes are divided by the same price index (the GDP deflator) in each year. Moreover, in this approach, the effects found for overall GDP in Table 2 correspond to the sum of the effects for the components of GDP considered in Table 4. For example, the defense-spending multiplier estimated in Table 2 equals one plus the sum of the estimated effects on the four components of GDP in Table 4. For the other right-hand-side variables, the estimated effect in Table 2 equals the sum of the estimated effects in Table 4.

The data for the components of GDP in Table 4 come from the BEA information available annually since 1929. Therefore, the samples considered—1950-2006, 1939-2006, and 1930-2006—do not go back before 1930. (The 1917-2006 sample in Table 2 used non-BEA data before 1929 for GDP and government purchases.)

Recall that the defense-spending multipliers estimated for GDP in Table 2 were 0.6-0.7. Correspondingly, the sum of the magnitudes of the crowding-out coefficients for the four parts of GDP considered in Table 4 is 0.3-0.4. For example, for the 1930-2006 sample, the estimated coefficient for C is -0.02 (s.e.=0.05)—not statistically significantly different from zero—that for I is -0.20 (0.06), that for non-defense government purchases is -0.06 (0.02), and that for net exports is -0.06 (0.02). Thus, the clearest crowding out from defense spending applies to private investment. However, crowding out is also significant for non-defense government purchases and net exports but not for private consumer expenditure.

We can go further to consider a finer breakdown among the components of GDP (not shown in Table 4). For the 1930-2006 sample, the coefficient of -0.20 (s.e.=0.06) for gross private domestic investment divides up as -0.09 (0.03) for non-residential fixed investment, -0.05 (0.02) for residential investment, and -0.05 (0.03) for inventory change. For consumer expenditure, the coefficient of -0.02 (0.05) breaks down into -0.06 (0.02) for consumer durables and 0.04 (0.04) for consumer non-durables and services—thus, the crowding-out for purchases of consumer durables is comparable to that for business investment. For non-defense government purchases, the coefficient of -0.06 (0.02) breaks down as -0.04 (0.01) for federal non-defense purchases and -0.03 (0.01) for state-and-local purchases. Finally, the coefficient of -0.06 (0.02) for net exports divides up as -0.05 (0.02) for exports and 0.00 (0.02) for imports.

The negative effect from the average marginal income-tax rate on overall GDP shows up most clearly in Table 2 for the 1950-2006 sample, with an estimated coefficient of -0.58 (s.e.=0.21). Table 4 shows that this response divides up roughly equally between consumer expenditure, C (-0.29 [0.11]) and gross private domestic investment, I (-0.35 [0.15]). For the finer breakdown (not shown in the table), the response of C divides up nearly equally between durables (-0.12 [.06]) and non-durables and services (-0.16 [0.07]). For I, the effects break down as 0.02 (0.07) for non-residential fixed investment, -0.14 (0.07) for residential investment, and -0.24 (0.08) for inventory change. The absence of an effect on non-residential fixed investment may arise because our measured AMTR applies mostly to wage income.

For the interest-rate spread variable, Table 2 shows negative effects on GDP for all samples, with a larger magnitude of response in the 1930 sample. In Table 4, the main negative effect in the 1950 and 1939 samples is on gross private domestic investment, I. The change in the 1930 sample is a sharply negative effect on private consumer spending, C; this response did not appear in the samples that exclude the Great Depression.

F. Results for Total Government Purchases

Since an expansion of defense spending tends to crowd out non-defense purchases, the rise in overall government purchases typically falls short of the increase in defense spending. Therefore, a multiplier calculated from defense spending alone tends to understate the multiplier computed for overall government purchases. If we assume that the non-defense and defense multipliers are equal,[22] we can estimate the multiplier for overall purchases by replacing the defense-related variables in Table 2 with corresponding measures based on total government purchases. In these revised equations, the instrument list still includes the defense-spending variables interacted with war years, as in Table 2.

Results from equations based on total government purchases are in Table 5. As expected, the estimated multipliers are larger than those in Table 2. For example, for the 1950 sample, the estimated value in Table 5, column 1, is 0.82 (s.e.=0.29), compared to 0.77 (0.28) in Table 2, column 1. For the 1939 sample, the estimated multiplier (evaluated at the median unemployment rate) in Table 5, column 2, is 0.71 (0.08), compared to 0.67 (0.07) in Table 2, column 3. The interaction term with the lagged unemployment rate is also higher in Table 5, column 2 (6.2 [2.4]) than the corresponding value in Table 2, column 3 (5.1 [2.2]). Results for the 1930 and 1917 samples are similar to those from the 1939 sample. Thus, the bottom line is that the estimated multiplier for total government purchases is higher by around 0.05 than the values computed for defense spending alone. However, this conclusion depends on the untested assumption that defense and non-defense spending multipliers are the same.

[22] We cannot test this proposition without satisfactory instruments related to non-defense purchases.

G. More Results on Taxes

Thus far, the findings on taxes involve changes in average marginal income-tax rates. GDP can respond to tax-rate changes due to substitution effects involving consumption versus leisure and work today versus work tomorrow. However, tax changes may also matter through income effects, the channel stressed by Romer and Romer (2009). These effects would involve changes in tax revenue, rather than tax rates, per se.

Empirically, movements in the average marginal income-tax rate are substantially positively correlated with changes in tax revenue. From 1950 to 2006, the correlation is 0.63 between the change in the ratio of federal revenue to GDP and the change in the federal part of our AMTR. The correlation is 0.76 between the change in the federal AMTR and the variable that Romer and Romer (2008) constructed to gauge incremental federal taxes—this variable is detailed in the notes to Table 6. Given these correlations, the AMTR used in Table 2 could be picking up a combination of substitution and income effects in the determination of GDP. In this section, we try to sort out these effects.

Before turning to the empirical results, consider what the theoretical analysis in section D, above, implies about income effects from changes in tax revenue. We can think of a pure positive income effect as reflecting information that a lump-sum bonanza will be coming in the future. One possible bonanza is a reduction in the future ratio of government purchases to GDP (a good thing if the services from government purchases are not highly valued)—and this change may be signaled currently by a decrease in the ratio of tax revenue to GDP. In the Barro and King (1984) setup without investment, however, a pure income effect has no impact on current GDP. If we allow for durable goods, we get that current investment falls, while current consumption and leisure rise (thereby absorbing part of the future bonanza). The rise in leisure implies, if labor is the only variable input into production, that current GDP falls. Thus, the prediction is that the income effect from a change in tax revenue would, if anything, impact negatively on current GDP.

Table 6 presents further results on tax variables for the 1950-2006 sample. Column 1 is essentially the same as column 1 of Table 2, with a minor difference in the instrument list. The estimated coefficient on the first lag of the change in the average marginal income-tax rate is -0.57 (s.e.=0.21). Column 2 of Table 6 adds the second lag of the tax-rate variable. The estimated coefficient is negative, -0.27 (0.19), but not statistically significant, and the first lag remains significantly negative, -0.45 (0.22).[23] The p-value for the joint significance of the two lags is 0.007.

The problem with adding the contemporaneous value of $\Delta\tau$ as a regressor is endogeneity. For a given tax structure, a rise in GDP tends to raise τ because households are pushed into higher marginal-rate brackets. In addition, the government might respond within the year to higher GDP by changing the tax law applicable to income for the current calendar year. To take account of the first channel, we calculated what each household's marginal income-tax rate would have been based on the prior year's (nominal) income when applied to the current year's tax law. We then averaged the marginal rates based on the previous year's incomes. We were able to construct this variable, using the NBER's TAXSIM program, for the federal individual income tax and social security from 1967 to 2006. We then formed an instrument for the current

[23]Our focus is on the overall marginal income-tax rate; that is, we implicitly have the same coefficients for changes in federal and state income-tax rates as for changes in social-security tax rates. If we separate the two income-tax rates from the social-security rate, we surprisingly get larger size coefficients for social security. The hypothesis of equal magnitude coefficients for each of the two lags is rejected with a p-value of 0.02. We have no good economic explanation for this result. However, a key part of the data pattern is that the increases in the AMTR from social security starting in the early 1970s fit well with the recessions of the mid 1970s and the early 1980s.

value of $\Delta\tau$ by taking the value of τ_t constructed as just described for the federal income tax and social security and subtracting the actual values of average marginal rates from the federal income tax and social security for the previous year. (This instrument assumes a value of zero for the change in the AMTR from state income taxes.) For the period 1950-66, the instrument takes on the constant value -0.0005, which is the median change from 1950 to 2006 (or, it turns out, also from 1917 to 2006).

If we add the contemporaneous change in τ to column 2 of Table 6, while adding the variable just described to the instrument list, we get that the estimated coefficients on the $\Delta\tau$ variables are 0.26 (s.e.=0.30) for the current year, -0.56 (0.25) for the first lag, and -0.22 (0.19) for the second lag.[24] Thus, the contemporaneous effect differs insignificantly from zero. Moreover, the positive point estimate likely still reflects endogeneity, in the sense of the government tending to raise tax rates when the economy is doing well, and vice versa.

Columns 3 and 4 of Table 6 show results when tax changes are measured by variations in overall federal revenue. Analogous to the construction of the government-purchases variables, we use the change in per capita real federal revenue (nominal revenue divided by the GDP deflator), expressed as a ratio to the previous year's per capita real GDP. As noted before, Romer and Romer (2008) attempt to isolate the exogenous part of the increment to federal revenue. To make use of their data, we exclude the federal-revenue variables from the instrument list but include the Romer-Romer tax-increment measures on this list. (See the notes to Table 6.) This approach should be superior to our previous assumption that lags of the federal-revenue variable are pre-determined with respect to GDP. In column 3, the estimated

[24]If we include the actual value of the current change in τ on the instrument list, the estimated coefficient on the current value of $\Delta\tau$ becomes 0.35 (s.e.=0.23), larger than before, as expected, but still insignificantly different from zero.

coefficient of the first lag of the revenue variable, -0.67 (s.e.=0.30), is significantly negative. In column 4, the second-lag coefficient is negative, -0.56 (0.34), but not significantly different from zero, and the first lag is still significant, -0.65 (0.31). The two lags are jointly significant with a p-value of 0.025.

Given the form of the tax-revenue variable, the estimated coefficients in Table 6, columns 3 and 4, can be interpreted as tax multipliers for GDP. Hence, the point estimate of the multiplier is -0.7 with a one-year lag and -1.2 when cumulated over the two lags.

This time pattern of results for federal tax revenue is broadly consistent with the findings of Romer and Romer (2009, Figure 9, panel c), based on a quarterly VAR system with one of their four categories of tax changes. That is, their main estimated negative response of GDP to a tax hike occurs with a lag of 3 to 10 quarters. The magnitudes of their estimated tax multipliers tend to exceed those that we found. However, one reason for this difference is that they enter their exogenous tax-change variable directly into an equation for determining GDP, rather than using this variable as an instrument for the change in overall federal revenue.[25]

The results change only moderately if we instrument the changes in the federal-revenue variable with the changes that Romer and Romer (2008) designate as exogenous. In this case, the estimated coefficient on the first lag of the federal-revenue variable (corresponding to Table 6, column 3) is -0.63 (s.e.=0.29). With the second lag included (corresponding to

[25]Imagine that, instead of including the change in per capita real federal revenue as a regressor, we had used one-half of this change. In that case, all results would be unchanged except that the coefficient—the tax multiplier—would be twice as large in magnitude. The inclusion of the Romer-Romer variable directly into an equation for GDP is analogous, except that the correlation between this variable and the federal-revenue variable is less than one.

column 4), the estimated coefficients become -0.65 (0.32) on the first lag and -0.81 (0.40) on the second, with a p-value for joint significance of 0.018.[26]

One virtue of the Romer-Romer variable is that it may allow for satisfactory estimation of the contemporaneous effect of a tax-revenue change on GDP. If we add the contemporaneous federal-revenue variable to the equation in column 4 of Table 6, while adding the contemporaneous Romer-Romer variable to the instrument list, the coefficient estimates are 0.81 (s.e.=0.37) on the current value, -0.58 (0.23) on the first lag, and -0.22 (0.31) on the second lag. The positive coefficient on the current variable suggests reverse causation; that is, the Romer-Romer variable that includes all of their bins does not seem to isolate the exogenous part of tax-revenue changes. If we include on the instrument list only the tax changes from the two bins that they designate as exogenous, the results become 0.24 (0.65), -0.61 (0.29), and -0.67 (0.48). Thus, in this case, the contemporaneous effect differs insignificantly from zero. In any event, the results provide no evidence for an important negative, contemporaneous effect of federal-revenue changes on GDP.

To try to distinguish substitution from income effects of tax changes, columns 5 and 6 of Table 6 include the marginal tax-rate and federal-revenue variables at the same time. Unfortunately, it is hard to separate the effects from these two aspects of tax changes. In column 6, with two lags of each tax variable included, the four tax variables are jointly significant with a p-value of 0.023. However, the p-value for joint significance of the two tax-rate variables is 0.29, whereas that for the two federal-revenue variables is 0.52. Thus, we would accept the hypothesis that, given the behavior of marginal income-tax rates, changes in federal

[26]The results with one lag of the tax-change variable, corresponding to Table 6, column 3, also do not change greatly if we include this variable itself on the instrument list. In that case, the coefficient is -0.60 s.e..=0.25. However, the results differ more when two lags are included; in this case, the coefficients are -0.60 (0.25) on the first lag and -0.10 (0.23) on the second, with a p-value for joint significance of 0.058.

revenue have no effect on GDP, so that income effects are nil. However, we would also accept the hypothesis that, given the behavior of federal revenue, changes in average marginal income-tax rates have no effect on GDP, so that substitution effects are nil. There is little power here in distinguishing substitution effects of tax changes from income effects.

We observed before, in Table 2, that the effects from changes in marginal income-tax rates look weaker in samples that start before 1950. The Romer-Romer data begin in 1945, so the earliest starting date that we can consider while allowing for two lags of their variable is 1947. In practice, beginning even one year earlier than 1950 eliminates the joint significance of the federal-revenue variables shown in Table 6, column 4 (where the p-value for joint significance of the two lags was 0.025). For example, for a sample that starts in 1948, the estimated coefficients on the federal-revenue variables are -0.35 (s.e.=0.31) on the first lag and -0.40 (0.32) on the second, with a p-value for joint significance of 0.15. The key problem is that the tax cut in 1948 does not match up well with the 1949 recession. Note that Romer and Romer (2009) use a sample that starts in 1950.

The marginal tax-rate variable holds up somewhat better with an earlier starting date. For example, for a sample that begins in 1948, the estimated coefficients (corresponding to Table 6, column 2) are -0.12 (s.e.=0.20) on the first lag and -0.40 (0.19) on the second, with a p-value for joint significance of 0.052.[27]

[27]If the marginal tax-rate and federal-revenue variables are entered together for a sample starting in 1948, the estimated coefficients for the tax-rate variable are 0.01 (s.e.=0.25) for the first lag and -0.66 (0.35) for the second, and those for the federal-revenue variable are -0.38 (0.36) for the first lag and 0.50 (0.57) for the second. P-values for joint significance are 0.16 for the tax-rate variables, 0.40 for the federal-revenue variables, and 0.10 for all tax variables.

VI. Concluding Observations

For samples that include World War II, the estimated multiplier connecting changes in defense spending to contemporaneous responses of GDP is reliably determined to be in the range of 0.6-0.7 when evaluated at the median unemployment rate. There is some evidence that this multiplier rises with the amount of economic slack, with the estimated multiplier reaching one when the unemployment rate is around 12%. In contrast, we lack reliable estimates of the multiplier for non-defense purchases, because the lack of good instruments makes it infeasible to isolate the direction of causation between these purchases and GDP. Since the defense-spending multiplier is significantly less than one in long samples, a rise in defense spending is estimated to crowd out significantly other components of GDP. The main crowding out shows up for gross private domestic investment but significant negative effects show up also for non-defense government purchases and net exports. The estimated effect on private consumer expenditure differs insignificantly from zero.

The sample that starts in 1950 indicates significantly negative effects from changes in average marginal income-tax rates on GDP. When interpreted as a tax multiplier (using the historical association between changes in tax revenue and changes in average marginal income-tax rates), we got a value around -1.1. However, these tax-rate effects are less reliably estimated in long samples, for example, those that include World War II, the Great Depression, and World War I. We used data from Romer and Romer (2008) to try to distinguish substitution effects of tax-rate changes from income effects of changes in tax revenue. However, we were unable to get a clear statistical distinction between these two channels.

The defense-spending multiplier of 0.6-0.7 applies in Table 2 while holding fixed tax rates and, implicitly, tax revenue. Thus, this multiplier accords most clearly with temporary variations in spending that are deficit financed. If, instead, higher spending goes along with correspondingly higher tax revenue, we also have to factor in the negative tax multiplier, which we estimated to be around -1.1. Thus, the full effect from greater defense spending and correspondingly higher tax revenue is a multiplier of about -0.5; that is, the balanced-budget multiplier is negative.

We are in the process of applying the methodology to long-term macroeconomic data for other countries. However, the U.S. evidence for isolating defense-spending multipliers works well because the main wars involved little destruction of domestic capital stock and only moderate loss of American life. The massive destruction during the world wars for most other OECD countries would preclude a similar analysis; that is, adverse supply shocks would confound the demand effects from greater government spending.

Promising cases that seem analogous to the U.S. experience are Australia, Canada, and New Zealand. These cases are particularly interesting because the entry dates into the world wars, 1914 and 1939, are earlier than the U.S. dates. In particular, the earlier entry into WWII means that the dramatic increases in defense spending took place in the context of unemployment rates than were higher than those in the United States. For example, for Canada, the dramatic rise in defense spending in 1940 matches up with an unemployment rate of over 14% in 1939. Therefore, the three countries should provide clearer evidence about how the defense-spending multiplier interacts with the amount of slack in the economy. However, further research will be required to construct time series on average marginal income-tax rates—the variable used in the present study for the United States.

References

Aiyagari, R., A. Marcet, T. Sargent, and J. Seppala (2002). "Optimal Taxation without State-Contingent Debt," *Journal of Political Economy*, 110, December, 1220-1254.

Barro, R.J. (1979). "On the Determination of the Public Debt, "*Journal of Political Economy*, 87, October, 940-971.

Barro, R.J. (1981). "Output Effects of Government Purchases," *Journal of Political Economy*, 89, December, 1086-1121.

Barro, R.J. (1984). *Macroeconomics*, New York, Wiley.

Barro, R.J. (1990). "On the Predictability of Tax-Rate Changes," in R.J. Barro. ed., *Macroeconomic Policy*, Cambridge MA, Harvard University Press, 268-297.

Barro, R.J. and R.G. King (1984). "Time-Separable Preferences and Intertemporal-Substitution Models of Business Cycles," *Quarterly Journal of Economics*, 99, November, 817-839.

Barro, R.J. and C. Sahasakul (1983). "Measuring the Average Marginal Tax Rate from the Individual Income Tax," *Journal of Business*, 56, October, 419-452.

Barro, R.J. and C. Sahasakul (1986). "Average Marginal Tax Rates from Social Security and the Individual Income Tax," *Journal of Business*, 59, October, 555-566.

Barro, R.J. and J.F. Ursua (2008). "Consumption Disasters since 1870," *Brookings Papers on Economic Activity*, spring, 255–335.

Blanchard, O. and R. Perotti (2002). "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics*, 117, November, 1329-1368.

Chetty, R. (2009). "Bounds on Elasticities with Optimization Frictions: A Reconciliation of Micro and Macro Labor Supply Elasticities," unpublished, Harvard University, October.

Fishback, P.V., W.C. Horrace, and S. Kantor (2005). "Did New Deal Grant Programs Stimulate Local Economies? A Study of Federal Grants and Retail Sales During the Great Depression," *Journal of Economic History*, 65, March, 36-71.

Hall, R.E. (2009). "By How Much Does GDP Rise if the Government Buys More Output?," unpublished, Stanford University, November, forthcoming in *Brookings Papers on Economic Activity*.

Johansson, E. (2003). "Intergovernmental Grants as a Tactical Instrument: Empirical Evidence from Swedish Municipalities," *Journal of Public Economics*, 87, May, 883-915.

Kendrick, J.W. (1961). *Productivity Trends in the United States*, Princeton NJ, Princeton University Press.

Ramey, V.A. (2009). "Identifying Government Spending Shocks: It's All in the Timing," unpublished, University of California San Diego, October.

Ramey, V.A. and M. Shapiro (1998). "Costly Capital Reallocation and the Effects of Government Spending," *Carnegie-Rochester Conference Series on Public Policy*, 48, June, 145-194.

Redlick, C.J. (2009). *Average Marginal Tax Rates in the United States: A New Empirical Study of their Predictability and Macroeconomic Effects, 1913-2006*, unpublished undergraduate thesis, Harvard University.

Romer, C.D. and D.H. Romer (2008). "A Narrative Analysis of Postwar Tax Changes," unpublished, University of California Berkeley, November.

Romer, C.D. and D.H. Romer (2009). "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," unpublished, University of California Berkeley, April, forthcoming in *American Economic Review*.

Rotemberg, J. and M. Woodford (1992). "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity," *Journal of Political Economy*, 100, December, 1153-1297.

Wright, G. (1974). "The Political Economy of New Deal Spending: An Econometric Analysis," *The Review of Economics and Statistics*, 56, February, 30-38.

Table 1 Data on Average Marginal Income-Tax Rates				
Year	**Overall Marginal Tax Rate**	**Federal Individual Income Tax**	**Social-Security Payroll Tax**	**State Income Taxes**
1912	0.000	0.000	0.000	[0.0001]
1913	0.003	0.003	0.000	[0.0001]
1914	0.005	0.005	0.000	[0.0002]
1915	0.007	0.007	0.000	[0.0002]
1916	0.013	0.013	0.000	[0.0003]
1917	0.037	0.037	0.000	[0.0003]
1918	0.054	0.054	0.000	[0.0004]
1919	0.052	0.052	0.000	[0.0004]
1920	0.046	0.046	0.000	[0.0005]
1921	0.043	0.042	0.000	[0.0005]
1922	0.047	0.046	0.000	[0.0006]
1923	0.034	0.033	0.000	[0.0006]
1924	0.036	0.035	0.000	[0.0007]
1925	0.031	0.030	0.000	[0.0007]
1926	0.029	0.028	0.000	[0.0007]
1927	0.033	0.032	0.000	[0.0008]
1928	0.042	0.041	0.000	[0.0008]
1929	0.036	0.035	0.000	0.0009
1930	0.024	0.023	0.000	0.0007
1931	0.018	0.017	0.000	0.0006
1932	0.030	0.029	0.000	0.0006
1933	0.032	0.031	0.000	0.0015
1934	0.036	0.034	0.000	0.0018
1935	0.041	0.038	0.000	0.0028
1936	0.055	0.052	0.000	0.0032
1937	0.058	0.046	0.009	0.0035
1938	0.046	0.034	0.009	0.0032
1939	0.051	0.038	0.009	0.0036
1940	0.069	0.056	0.009	0.0038
1941	0.126	0.113	0.009	0.0038
1942	0.205	0.192	0.008	0.0047
1943	0.221	0.209	0.007	0.0048
1944	0.263	0.252	0.006	0.0052
1945	0.268	0.257	0.006	0.0047
1946	0.238	0.226	0.007	0.0052
1947	0.238	0.226	0.006	0.0056
1948	0.193	0.180	0.006	0.0072
1949	0.187	0.175	0.005	0.0072
1950	0.211	0.196	0.007	0.0079

Table 1, continued				
Year	**Overall Marginal Tax Rate**	**Federal Individual Income Tax**	**Social-Security Payroll Tax**	**State Income Taxes**
1951	0.248	0.231	0.009	0.0085
1952	0.268	0.251	0.008	0.0086
1953	0.266	0.249	0.008	0.0086
1954	0.241	0.222	0.010	0.0087
1955	0.250	0.228	0.012	0.0098
1956	0.254	0.232	0.012	0.0101
1957	0.255	0.232	0.013	0.0104
1958	0.253	0.229	0.013	0.0114
1959	0.265	0.236	0.016	0.0130
1960	0.265	0.234	0.018	0.0129
1961	0.270	0.240	0.017	0.0132
1962	0.275	0.244	0.017	0.0142
1963	0.280	0.247	0.018	0.0146
1964	0.253	0.221	0.017	0.0155
1965	0.244	0.212	0.016	0.0164
1966	0.251	0.212	0.022	0.0173
1967	0.256	0.215	0.021	0.0202
1968	0.286	0.238	0.026	0.0229
1969	0.298	0.250	0.024	0.0245
1970	0.286	0.237	0.022	0.0270
1971	0.278	0.227	0.022	0.0291
1972	0.289	0.231	0.025	0.0332
1973	0.305	0.239	0.034	0.0327
1974	0.325	0.247	0.042	0.0354
1975	0.333	0.254	0.043	0.0370
1976	0.340	0.257	0.043	0.0391
1977	0.361	0.277	0.043	0.0410
1978	0.369	0.284	0.043	0.0421
1979	0.384	0.273	0.068	0.0420
1980	0.400	0.286	0.072	0.0412
1981	0.418	0.294	0.084	0.0403
1982	0.404	0.275	0.087	0.0414
1983	0.391	0.256	0.091	0.0450
1984	0.393	0.254	0.095	0.0446
1985	0.399	0.260	0.095	0.0442

Table 1, continued				
Year	**Overall Marginal Tax Rate**	**Federal Individual Income Tax**	**Social-Security Payroll Tax**	**State Income Taxes**
1986	0.401	0.259	0.097	0.0447
1987	0.375	0.237	0.096	0.0422
1988	0.356	0.218	0.097	0.0418
1989	0.360	0.218	0.100	0.0421
1990	0.362	0.217	0.102	0.0421
1991	0.371	0.219	0.108	0.0438
1992	0.369	0.217	0.108	0.0448
1993	0.379	0.224	0.110	0.0446
1994	0.385	0.230	0.111	0.0446
1995	0.386	0.232	0.109	0.0445
1996	0.385	0.235	0.107	0.0441
1997	0.386	0.237	0.105	0.0440
1998	0.387	0.239	0.104	0.0440
1999	0.390	0.243	0.103	0.0442
2000	0.392	0.247	0.101	0.0442
2001	0.385	0.238	0.103	0.0440
2002	0.380	0.231	0.105	0.0436
2003	0.359	0.211	0.104	0.0441
2004	0.358	0.213	0.102	0.0433
2005	0.351	0.216	0.092	0.0433
2006	0.353	0.217	0.093	0.0432

Note: See the text on the construction of average (income-weighted) marginal tax rates for the federal individual income tax, social-security payroll tax, and state income taxes. Values shown in brackets for state income taxes for 1912-28 are interpolations. The total is the sum of the three pieces. The construction of these data are detailed in an appendix posted at http://www.economics.harvard.edu/faculty/barro/data_sets_barro.

Table 2 Equations for GDP Growth, Various Samples					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Starting date	**1950**	**1939**	**1939**	**1930**	**1917**
Δg-defense	0.77** (0.28)	0.59** (0.06)	0.67** (0.07)	0.66** (0.09)	0.64** (0.10)
U(-1)	0.52** (0.17)	0.67** (0.15)	0.62** (0.15)	0.62** (0.10)	0.48** (0.11)
(Δg-defense)*U(-1)	--	--	5.1* (2.2)	4.6 (2.9)	4.6 (3.3)
Δτ(-1)	-0.58** (0.21)	0.10 (0.16)	-0.26 (0.19)	-0.39 (0.25)	-0.33 (0.28)
Δ[τ*(g-def/y)](-1)	--	--	0.73** (0.28)	0.84* (0.38)	0.80 (0.42)
Interest-rate spread, square	-48.3* (20.5)	-56.4* (23.5)	-48.1* (22.9)	-104.1** (13.2)	-76.1** (12.4)
p-value for τ variables	0.007	0.53	0.030	0.083	0.16
R^2	0.46	0.77	0.80	0.73	0.65
σ	0.018	0.021	0.020	0.027	0.031

*Significant at 0.05 level. **Significant at 0.01 level.

Notes to Table 2

Data are annual from the starting year shown through 2006. The dependent variable is the change from the previous year in per capita real GDP divided by the previous year's per capita real GDP. Δg-defense is the change from the previous year in per capita real defense spending (nominal spending divided by the GDP deflator) divided by the previous year's per capita real GDP. U(-1) is the lagged unemployment rate. (Δg-defense)*U(-1) is an interaction between the first two independent variables, each expressed as a deviation from its respective median over 1914-2006: 0.0002 for Δg-defense and 0.0557 for U(-1). $\Delta\tau$ is the change from the previous year in the average marginal income-tax rate from federal and state income taxes and social security, as shown in Table 1. The lagged value of this variable appears in the equations. $\Delta[\tau*(\text{g-def/y})](-1)$ is the lagged value of the change in the interaction between the average marginal income-tax rate and the ratio of defense spending to GDP. The interest-rate spread is the difference between the yield on long-term Baa corporate bonds and that on long-term U.S. Treasury bonds. Before 1919, the spread is estimated from data on long-term Aaa corporate bonds. The square of the spread appears in the equations. Data on yields are from Moody's, as reported on the website of the Federal Reserve Bank of St. Louis.

Estimation is by two-stage least-squares, using as instruments all of the independent variables in this table, except for the square of the interest-rate spread, which is replaced by its lagged value, and the variables involving the changes in defense spending. These variables are replaced by interactions with a dummy variable for war years (1914-20, 1939-46, 1950-54, 1966-71). For WWI, the period starts with the initiation of war in Europe, 1914, goes to the signing of the Versailles Treaty in 1919, then includes one year of war aftermath. Similarly, WWII covers 1939-46. The other years designated as "war" are those with a casualty rate of at least 0.01 per thousand population or for the year following a casualty rate at least this high. The instrument list also contains the first lag of the dependent variable. The p-value is for a test that the coefficients on one or two variables involving the tax rate are all equal to zero. σ is the standard-error-of-estimate.

Table 3 More Results on Government Purchases						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Starting date	**1950**	**1939**	**1930**	**1917**	**1950**	**1950**
Δg-defense	0.89** (0.28)	0.65** (0.08)	0.65** (0.10)	0.64** (0.10)	0.75** (0.24)	0.59* (0.26)
U(-1)	0.64** (0.18)	0.61** (0.15)	0.62** (0.11)	0.48** (0.11)	0.31* (0.16)	0.56** (0.15)
(Δg-defense)*U(-1)	--	4.7* (2.4)	4.5 (3.1)	4.7 (3.3)	--	--
Δτ(-1)	-0.54** (0.20)	-0.29 (0.20)	-0.40 (0.26)	-0.33 (0.28)	-0.36* (0.18)	-0.43* (0.19)
Δ[τ*(g-def/y)](-1)	--	0.72** (0.28)	0.84* (0.38)	0.80 (0.42)	--	--
Interest-rate spread, square	-41.4* (20.2)	-49.8* (23.0)	-104.7** (13.6)	-75.8** (12.4)	-40.5* (18.4)	-24.0 (20.1)
Δg-non-defense	1.93* (0.90)	-0.34 (0.73)	-0.13 (0.63)	0.04 (0.48)	--	--
Δ(GM sales)	--	--	--	--	3.44** (0.82)	--
Δ(GE sales)	--	--	--	--	--	17.86** (4.59)
R^2	0.49	0.80	0.73	0.65	0.51	0.56
σ	0.017	0.020	0.028	0.031	0.017	0.016

*Significant at 0.05 level. **Significant at 0.01 level.

Note: See the notes to Table 2. The first four columns include the variable Δg-non-defense, which is the change from the previous year in per capita real non-defense government purchases (nominal purchases divided by the GDP deflator) for all levels of government, divided by the previous year's per capita real GDP. The variable Δg-non-defense is included in the instrument lists for the first four columns. Δ(GM sales) is the change from the previous year in per capita real net sales of General Motors Corporation, expressed as a ratio to the previous year's per capita real GDP. Real net sales are nominal sales divided by the GDP deflator. This variable is included in the instrument list for column 5. Δ(GE sales), in column 6, is treated analogously but using net sales of General Electric Corporation. The GM and GE data come from annual reports of the two companies.

Table 4 Effects on Components of GDP				
Sample: 1950-2006				
Dependent variable	**ΔC**	**ΔI**	**Δ(G-non def)**	**Δ(X-M)**
Δg-defense	0.00 (0.15)	-0.13 (0.21)	-0.06 (0.04)	-0.04 (0.09)
U(-1)	0.22* (0.09)	0.43** (0.12)	-0.06* (0.03)	-0.08 (0.05)
Δτ(-1)	-0.29** (0.11)	-0.35* (0.15)	-0.02 (0.03)	0.07 (0.07)
Interest-rate spread, square	-10.1 (11.2)	-31.6* (14.9)	-3.6 (3.2)	-2.1 (6.5)
R^2, σ	0.29, 0.010	0.44, 0.013	0.15, 0.003	0.08, 0.006
σ	0.010	0.013	0.003	0.006
Sample: 1939-2006				
Δg-defense	-0.01 (0.04)	-0.22** (0.06)	-0.05** (0.01)	-0.05* (0.02)
U(-1)	0.20** (0.07)	0.48** (0.12)	-0.04 (0.03)	-0.02 (0.05)
(Δg-defense)*U(-1)	2.7* (1.1)	2.4 (1.7)	-1.2** (0.4)	1.2 (0.7)
Δτ(-1)	-0.19* (0.10)	-0.08 (0.15)	-0.09** (0.03)	0.09 (0.06)
Δ[τ*(g-def/y)](-1)	0.58** (0.14)	0.29 (0.22)	-0.03 (0.05)	-0.11 (0.09)
Interest-rate spread, square	-6.3 (11.5)	-30.4 (18.0)	-5.0 (4.2)	-5.5 (7.3)
R^2	0.35	0.54	0.51	0.35
σ	0.010	0.016	0.004	0.006
Sample: 1930-2006				
Δg-defense	-0.02 (0.05)	-0.20** (0.06)	-0.06** (0.02)	-0.06** (0.02)
U(-1)	0.25** (0.06)	0.32** (0.07)	0.03 (0.02)	0.00 (0.02)
(Δg-defense)*U(-1)	2.1 (1.7)	2.7 (1.9)	-1.3* (0.6)	1.2 (0.7)
Δτ(-1)	-0.25 (0.14)	-0.15 (0.16)	-0.09 (0.05)	0.10 (0.06)
Δ[τ*(g-def/y)](-1)	0.66** (0.22)	0.29 (0.25)	0.00 (0.07)	-0.11 (0.09)
Interest-rate spread, square	-56.1** (7.6)	-42.9** (8.6)	-4.2 (2.6)	-0.6 (3.0)
R^2, σ	0.46, 0.016	0.54, 0.018	0.30, 0.005	0.34, 0.006
σ	0.016	0.018	0.005	0.006

*Significant at 0.05 level. **Significant at 0.01 level.

Notes to Table 4

These results correspond to Table 2, except for the specifications of the dependent variables, which are now based on components of GDP. For personal consumer expenditure, C, the dependent variable equals the change in per capita real expenditure (nominal expenditure divided by the GDP deflator), expressed as a ratio to the previous year's per capita real GDP. The same approach applies to gross private domestic investment, I, non-defense government purchases (by all levels of government), and net exports. Data since 1929 are from the Bureau of Economic Analysis.

Table 5 Results Using Total Government Purchases				
	(1)	**(3)**	**(4)**	**(5)**
Starting date	**1950**	**1939**	**1930**	**1917**
Δg	0.82** (0.29)	0.71** (0.08)	0.73** (0.10)	0.68** (0.10)
U(-1)	0.57** (0.16)	0.66** (0.15)	0.55** (0.11)	0.44** (0.11)
Δg*U(-1)	--	6.2** (2.4)	6.2 (3.2)	5.6 (3.6)
$\Delta\tau$(-1)	-0.56** (0.20)	-0.20 (0.19)	-0.33 (0.25)	-0.26 (0.27)
Δ[τ*(g-def/y)](-1)	--	0.74** (0.28)	0.81* (0.37)	0.81* (0.42)
Interest-rate spread, square	-45.3* (20.4)	-44.6 (23.6)	-97.8** (13.3)	-71.8** (12.6)
p-value for τ variables	0.006	0.028	0.092	0.15
R^2	0.48	0.79	0.74	0.65
σ	0.017	0.020	0.027	0.031

*Significant at 0.05 level. **Significant at 0.01 level.

Note: These results correspond to those in Table 2, except that the variables Δg and Δg*U(-1) now refer to total government purchases, rather than defense purchases. The instruments are the same as those used in Table 2.

Table 6 More Results on Tax Variables, 1950-2006						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Δg-defense	0.75** (0.28)	0.70* (0.28)	0.79** (0.31)	0.75* (0.32)	0.86** (0.30)	0.82** (0.30)
U(-1)	0.53** (0.17)	0.51** (0.17)	0.49** (0.18)	0.33 (0.22)	0.48** (0.18)	0.50* (0.24)
Δτ(-1)	-0.57** (0.21)	-0.45* (0.22)	--	--	-0.43 (0.25)	-0.31 (0.26)
Δτ(-2)	--	-0.27 (0.19)	--	--	--	-0.37 (0.36)
[Δ(federal taxes)/Y(-1)](-1)	--	--	-0.67* (0.30)	-0.65* (0.31)	-0.34 (0.34)	-0.35 (0.35)
[Δ(federal taxes)/Y(-1)](-2)	--		--	-0.56 (0.34)	--	0.19 (0.67)
Interest-rate spread, square	-53.6** (20.5)	-45.6* (21.0)	-73.2** (20.8)	-64.9** (22.3)	-59.9** (21.4)	-50.8* (22.2)
p-value for τ	0.008	0.007	--	--	0.082	0.29
p-value for federal taxes	--	--	0.028	0.025	0.33	0.52
p-value for all tax variables	0.008	0.007	0.028	0.025	0.013	0.023
R^2	0.46	0.48	0.42	0.39	0.47	0.51
σ	0.018	0.017	0.018	0.019	0.018	0.017

*Significant at 0.05 level. **Significant at 0.01 level.

Note: See the notes to Table 2. Data are annual from 1950 through 2006. The equations in columns 1, 2, 5, and 6 include one or two lags of the change in the average marginal income-tax rate, Δτ. The variable Δ(federal taxes)/Y(-1) is the change in per capita real federal revenue (total nominal receipts from Bureau of Economic Analysis divided by the GDP deflator), expressed as a ratio to the previous year's per capita real GDP. The equations in columns 3-6 include one or two lags of this variable. The instrument list for all equations consists of Δg-defense, U(-1), Δτ(-1), Δτ(-2), the first lags of the dependent variable and the square of the interest-rate spread, and two lags of a federal tax variable constructed from information given in Romer and Romer (2008, Table 1, columns 1-4). Their variable measures legislated, projected incremental nominal federal tax revenue for year t. We use their series that neglects retroactive changes. Their data indicate the quarter in which the incremental revenue flow (expressed at an annual rate) is supposed to start. We aggregated their quarterly figures for each year to get annual numbers and then expressed these values as ratios to the previous year's nominal GDP.

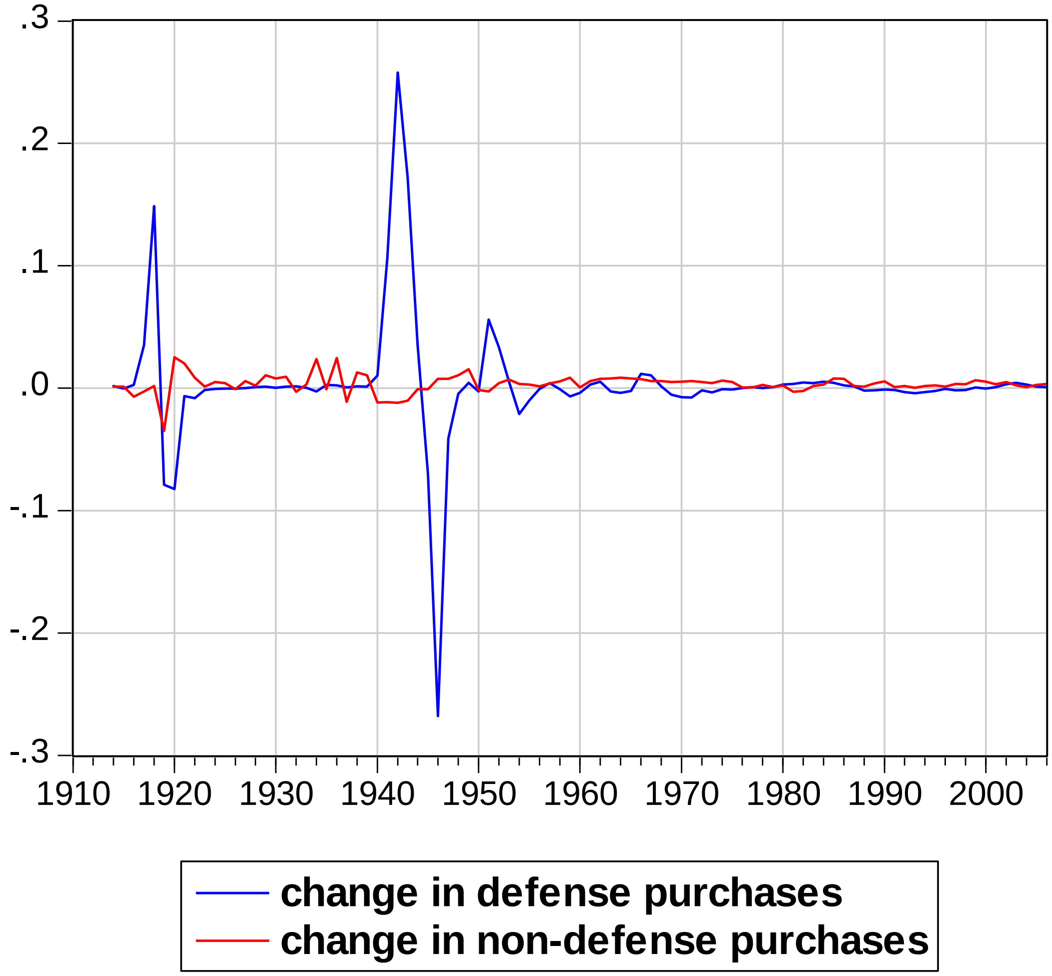

Figure 1

Changes in Defense and Non-Defense Government Purchases, 1914-2006

(expressed as ratios to the previous year's GDP)

Note: The figure shows the change in per capita real government purchases (nominal purchases divided by the GDP deflator), expressed as a ratio to the prior year's per capita real GDP. The blue graph is for defense purchases, and the red graph is for non-defense purchases by all levels of government. The data on government purchases since 1929 are from Bureau of Economic Analysis and, before that, from Kendrick (1961). The GDP data are described at http://www.economics.harvard.edu/faculty/barro/data_sets_barro.

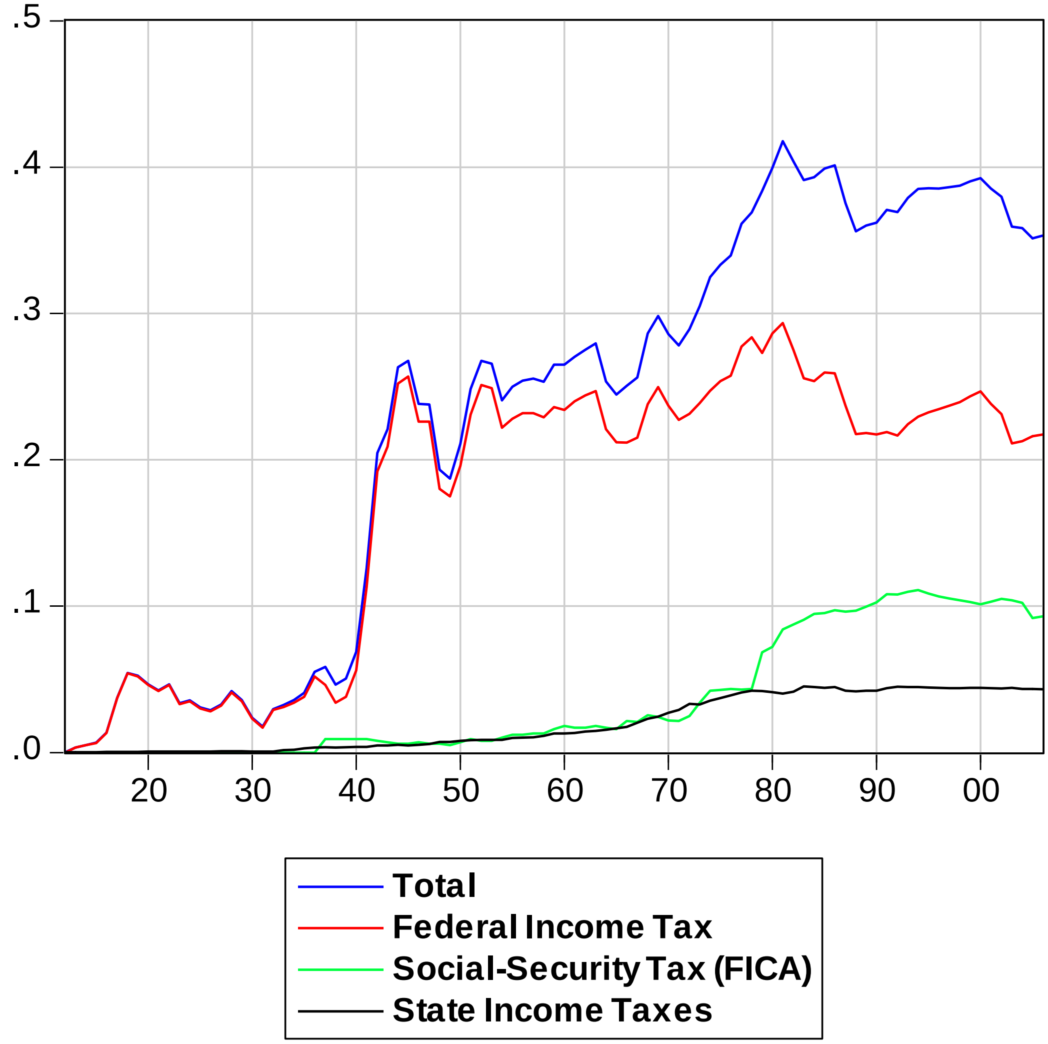

Figure 2

Average Marginal Income-Tax Rates, 1912-2006

Note: The red graph is for the federal individual income tax, the green graph is for the social-security payroll tax (FICA), and the black graph is for state income taxes. The blue graph is the total average marginal income-tax rate. The data are from Table 1.