The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States*

Karel Mertens and Morten O. Ravn

March 9, 2012

Abstract

This paper estimates the dynamic effects of changes in taxes in the United States. We distinguish between the effects of changes in personal and corporate income taxes using a new narrative account of federal tax liability changes in these two tax components. We develop an estimator in which narratively identified tax changes are used as proxies for structural tax shocks and apply it to quarterly post WWII US data. We find that short run output effects of tax shocks are large and that it is important to distinguish between different types of taxes when considering their impact on the labor market and the major expenditure components.

Keywords: Fiscal policy, tax changes, vector autoregressions, narrative identification, measurement error

JEL Classification: E20, E32, E62, H30

---

*We are grateful to Martin Eichenbaum, three referees, Andre Kurmann and participants at various seminars and conferences for very useful comments. We also thank Jonas Fisher and Todd Walker for sharing their data. Andres Dallal provided superb research assistance and Mertens acknowledges financial support from the Cornell Institute for the Social Sciences. *Mertens*: Department of Economics, Cornell University, km426@cornell.edu; *Ravn*: Department of Economics, University College London, m.ravn@ucl.ac.uk.

1 Introduction

This paper presents evidence on the aggregate effects of changes in tax policy in the US in the post WWII sample. Exogenous changes in taxes are identified in a vector autoregressive model by proxying latent tax shocks with narratively identified tax liability changes. We discriminate between the effects of changes in average personal income tax rates (APITRs) and the effects of changes in average corporate income tax rates (ACITRs). We find large short run effects on aggregate output of unanticipated changes in either tax rates. Cuts in personal income taxes lead to a fall in tax revenues while corporate income tax cuts on average have little impact on tax revenues. Cuts in APITRs raise employment, consumption and investment. Cuts in ACITRs boost investment, do not affect or even lower private consumption, and have no immediate effects on employment.

The key issue in estimating the impact of economic policies is identification. In the case of tax policy shocks this is particularly challenging both because of endogeneity of policy variables and because of the diversity of policy instruments. The literature has often concentrated on exogenous changes in *total* tax revenues. There is little reason to expect that the many types of taxes available to governments all have the same impact on the economy and therefore can be summarized in a single tax measure. We deviate from the literature and look instead at two broad groupings of taxes, personal income taxes and corporate income taxes. In total these two types of taxes account for more than 90 percent of total federal tax revenues and we argue that the tax categories are individually sufficiently homogeneous that one can more meaningfully estimate their impact.

Endogeneity has been addressed in alternative ways. One line of papers uses the narrative approach to identify exogenous tax changes and estimates their effects by regressing observables on the narratively identified policy shocks, e.g. Romer and Romer (2010). An attractive feature of this approach is that the narrative record summarizes the relevant features of a potentially very large information set. On the other hand, a concern with the existing literature is that the narratively identified exogenous changes in policy instruments are implicitly viewed as mapping one-to-one into the true structural shocks. In practice there is good reason to expect that narratively identified shocks suffer from measurement errors as historical records rarely are sufficiently unequivocal that calls of judgment can be avoided. Another approach adopts structural vector autoregressions (SVARs) and achieves identification by exploiting institutional features of tax and transfer systems, see e.g. Blanchard and Perotti (2002), or by introducing sign restrictions derived from economic theory, see Mountford and Uhlig (2009). This approach has the advantage that VARs provide a parsimonious characterization of the shock transmission mechanism but identification requires parameter restrictions that may be questioned.

In this paper we develop an estimation strategy that exploits the attractive features of both SVARs and the narrative method but at the same time addresses key weaknesses of the existing approaches. The estimator exploits the informational content of narrative measures of exogenous changes in taxes for identification in an SVAR framework. The key identifying assumptions that we propose are that narrative measures correlate with latent tax shocks but are orthogonal to other structural shocks. The main idea is to complement the usual VAR residual covariance restrictions with these moment conditions to achieve identification and avoid direct assumptions on structural parameters as is required in standard SVAR approaches. The resulting structural model can be estimated using a simple three step procedure and is straightforward to implement. We show that the estimator effectively extends the use of the narrative approach to cases in which the narrative shock series is measured with error and that under some additional assumptions it produces an estimate of the reliability of the narrative making it possible to judge its quality.

Given our focus on disaggregated taxes, we construct a new narrative account of shocks to average personal and corporate tax rates for the United States. This narrative is developed from Romer and Romer's (2009a) account of changes in federal US tax liabilities which we decompose into changes in personal and corporate income tax liabilities. We use only those tax changes that Romer and Romer (2009a) classify as exogenous. Following Mertens and Ravn (2011a), we also exclude those changes with implementation lags exceeding one quarter to remove anticipation effects. The disaggregation of the Romer tax shocks poses new challenges because of the correlation between legislated changes in personal and corporate taxes, which we resolve with recursivity assumptions.

Based on this methodology we provide new estimates of the impact of tax policy shocks in the US. In our benchmark specification, we find that a one percentage point cut in the APITR raises real GDP per capita on impact by 1.4 percent and by up to 1.8 percent after three quarters. A one percentage point cut in the ACITR raises real GDP per capita on impact by 0.4 percent and by up to 0.6 percent after one year. Cuts in personal income taxes lower tax revenues while cuts in corporate taxes have no significant impact on revenues because of a very elastic response of the tax base. Translating into multipliers, the maximum personal income tax multiplier is 2.5 in the third quarter, whereas the corporate income tax multiplier is poorly defined given our finding that there is on average little impact on tax revenues from changes in corporate tax rates.

Changes in both types of taxes have important but distinct effects on other macroeconomic aggregates. A cut in the APITR raises employment, lowers the unemployment rate and increases hours worked per worker. A cut in the ACITR, on the other hand, has no immediate impact on either employment or hours per worker. Both cuts in the APITR and in the ACITR lead to increases in investment, but only cuts in personal income taxes stimulate private consumption. Cuts in corporate income taxes instead have little effect or discourage private consumption in the short run. We find no signs of any significant change in government spending or nominal interest rates following tax shocks. The differences in the size and signs of the responses to the two types of taxes illustrates the necessity of discriminating between different types of taxes.

With some additional assumptions about the nature of the measurement error, our estimation approach produces a measure of the reliability of the narrative series that may be of independent interest. This measure leads to estimates of the squared correlation between linear combinations of the narrative shocks and the true structural tax shocks. We estimate correlations between the principal components of the narrative tax shock measures and the latent tax shocks of 0.55 and 0.83. Thus, the narratives contain valuable information for identification purposes but measurement errors are nonetheless a relevant concern in practical applications.

The empirical findings support several conclusions relevant to the ongoing debate on fiscal policy. Given the currently available evidence on the multipliers associated with US government spending, see Ramey (2011b) for a recent review, our estimates indicate that the federal tax multipliers are likely to be larger than those associated with federal government purchases. If policy objectives include short run job creation and consumption stimulus, then cuts to personal income taxes are more effective than cuts to corporate profit taxes. If the objective is to raise tax revenues, increases in personal income taxes are effective, but the costs in terms of job and output losses are relatively large. Increases in corporate profit taxes are not likely to raise significant revenues.

The remainder of the paper is organized as follows. Section 2 presents the estimation procedure. In Section 3 we present the narrative series on personal income and corporate income tax changes and the benchmark estimates. This section also provides a robustness analysis. Section 4 examines the wider macroeconomic impact of tax changes. Section 5 provides some concluding remarks.

2 Estimation and Identification

This section presents our estimation procedure. The main idea of our approach is to exploit prior information in narrative accounts of policy changes to identify structural fiscal shocks in an SVAR framework. In Section 2 .1, we first describe the formal econometric framework and state the identifying assumptions on which our impulse response estimates are based. In Section 2 .2, we provide a measurement error interpretation of our framework. We make some specific assumptions about the error in measurement to elicit potential sources of bias in more conventional narrative approaches and propose measures of statistical reliability to quantify the quality of identification.

2 .1 General Methodology

Let Y_t be an $n \times 1$ vector of observables. We assume that the dynamics of the observables are described by a system of linear simultaneous equations,

$$\mathcal{A}Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \varepsilon_t \tag{1}$$

where $\mathcal{A}$ is an $n \times n$ nonsingular matrix of coefficients, α_i, $i = 1,..,p$, are $n \times n$ coefficient matrices and ε_t is an $n \times 1$ vector of structural shocks with $E[\varepsilon_t] = 0$, $E[\varepsilon_t \varepsilon_t'] = I$, $E[\varepsilon_t \varepsilon_s'] = 0$ for $s \neq t$ where I is the identity matrix. The specification in (1) omits deterministic terms and exogenous regressors for notational brevity. An equivalent representation of the dynamics of Y_t is

$$Y_t = \sum_{i=1}^{p} \delta_i Y_{t-i} + \mathcal{B}\varepsilon_t \tag{2}$$

where $\mathcal{B} = \mathcal{A}^{-1}$, $\delta_i = \mathcal{A}^{-1}\alpha_i$.

In the SVAR literature ε_t is treated as a vector of latent variables that are estimated on the basis of the prediction errors of Y_t conditional on the information contained in the vector of lagged dependent variables $X_t = \left[Y_{t-1}', .., Y_{t-p}'\right]'$, and by imposing identifying assumptions. Let the $n \times 1$ vector u_t denote the reduced form residuals which are related to the structural shocks by

$$u_t = \mathcal{B}\varepsilon_t. \tag{3}$$

Since $E[u_t u_t'] = \mathcal{B}\mathcal{B}'$, an estimate of the covariance matrix of u_t provides $n(n+1)/2$ independent identifying restrictions. However, identification of the elements of at least one of the columns of $\mathcal{B}$ requires more identifying restrictions. The fiscal SVAR literature has accomplished this task in a variety of ways. For instance, Blanchard and Perotti (2002) exploit institutional features of the US tax system and policy reaction lags to impose coefficient restrictions on $\mathcal{B}$. Alternatively, Mountford and Uhlig (2009) impose sign restrictions on the impulse response functions implied by (2).

We propose instead to use covariance restrictions obtained from proxies for the latent shocks. Let m_t be a $k \times 1$ vector of proxy variables that are correlated with k structural shocks of interest but orthogonal to other shocks. Consider the partition $\varepsilon_t = [\varepsilon_{1t}', \varepsilon_{2t}']'$, where ε_{1t} is the $k \times 1$ vector containing the shocks of interest and the $(n-k) \times 1$ vector ε_{2t} contains all other $n-k$ shocks.[1] Without loss of generality we assume that $E[m_t] = 0$. The proxy variables can be used for identification of $\mathcal{B}$ as long as the following conditions are satisfied,

$$E[m_t \varepsilon_{1t}'] = \Phi \,, \tag{4}$$

$$E[m_t \varepsilon_{2t}'] = 0 \,, \tag{5}$$

$$E[m_t X_t'] = 0 \,, \tag{6}$$

where Φ is an unknown nonsingular $k \times k$ matrix. The first condition states that the proxy variables are correlated with the shocks of interest. The second condition requires that the proxy variables are uncorrelated with all other shocks. These two conditions are the key identifying assumptions. The third condition requires that the proxy variables are orthogonal to the history of Y_t. This assumption can be relaxed, because when a candidate proxy $\tilde{m}_t$ is correlated with the lagged dependent variables X_t then m_t can be the error from projecting $\tilde{m}_t$ on X_t. Note that we do not require that the proxies coincide exactly with the true latent shocks ε_{1t} or that they are mutually uncorrelated. As long as the proxies are orthogonal to the other shocks ε_{2t}, they contain prior information that can be exploited for identification purposes.

The assumptions on m_t translate to linear restrictions on the elements of $\mathcal{B}$ that can be used for identification. Consider the following partitioning of $\mathcal{B}$,

$$\mathcal{B} = \begin{bmatrix} \underset{n\times k}{\beta_1} & \underset{n\times(n-k)}{\beta_2} \end{bmatrix} \,, \quad \beta_1 = \begin{bmatrix} \underset{k\times k}{\beta_{11}'} & \underset{k\times(n-k)}{\beta_{21}'} \end{bmatrix}' \,, \quad \beta_2 = \begin{bmatrix} \underset{(n-k)\times k}{\beta_{12}'} & \underset{(n-k)\times(n-k)}{\beta_{22}'} \end{bmatrix}' \,,$$

[1]We assume that m_t and ε_{1t} are of the same dimension k. The case where multiple proxy variables are available, i.e. $dim(m_t) > k$, can be dealt with using factor analytic techniques.

with nonsingular β_{11} and β_{22}. Conditions (2) through (6) imply that

$$\Phi\beta_1' = \Sigma_{mu'} \ , \tag{7}$$

where henceforth we use the notation $\Sigma_{AB} \equiv E[A_t B_t]$ for any random vector or matrix A_t and B_t. The system in (7), which is of dimension $n \times k$, provides additional identifying restrictions but also depends on the k^2 unknown elements of Φ. Because we do not wish to make any assumptions on Φ other than nonsingularity, equation (7) provides really only $(n-k)k$ new identification restrictions. Partitioning $\Sigma_{mu'} = [\Sigma_{mu_1'} \ \ \Sigma_{mu_2'}]$, where $\Sigma_{mu_1'}$ is $k \times k$ and $\Sigma_{mu_2'}$ is $k \times (n-k)$ and using (7), these restrictions can be expressed as

$$\beta_{21} = (\Sigma_{mu_1'}^{-1}\Sigma_{mu_2'})'\beta_{11} \ . \tag{8}$$

Since $\Sigma_{mu_1'}^{-1}\Sigma_{mu_2'}$ is estimable, this constitutes a set of covariance restrictions of the type discussed in Hausman and Taylor (1983). In practice, estimation can proceed in three stages:

- **First Stage**: Estimate the reduced form VAR by least squares.
- **Second Stage**: Estimate $\Sigma_{mu_1'}^{-1}\Sigma_{mu_2'}$ from regressions of the VAR residuals on m_t.
- **Final Stage**: Impose the restrictions in (8) and estimate the objects of interest, if necessary in combination with further identifying assumptions.

Hausman and Taylor (1983) develop general necessary and sufficient conditions for identification in the final stage and also offer an instrumental variables interpretation.[2]

The key requirement is the availability of proxies that satisfy the conditions in equations $(4)-(6)$. We propose to use narratively identified measures of exogenous shocks to fiscal variables as proxies for the structural fiscal shocks. The use of narrative accounts has a long standing tradition in macroe-

[2] Our approach is also related to Nevo and Rosen (2010) who use weaker covariance restrictions in an instrumental variables framework to achieve partial identification, and Evans and Marshall (2009) who identify shocks in VARs with the aid of auxiliary shock measures derived from economic models.

conomics in the estimation of the effects of, for instance, fiscal and monetary policy shocks.[3] Existing applications of the narrative identification approach typically estimate the response to structural innovations by regressing the observables on (distributed lags of) the narratives or by adding them as variables in a VAR. In most of these applications, the interpretation of the results relies on implicit assumptions on Φ, the covariance between the narratives and the latent structural innovations. Our approach differs in that it does not require assumptions on Φ other than nonsingularity. Contrary to most existing narrative studies, this allows for the possibility of various types of measurement error, which is discussed next.[4]

2 .2 Measurement Problems and Reliability

A useful interpretation of the proxy variables is as imperfect measurements of (linear combinations of) latent structural shocks. Such an interpretation is natural in applications where the proxies are specified as narratively identified monetary or fiscal policy changes. Narratives of economic policy are constructed from historical sources that are used to summarize information about the size, timing, and motivation of policy interventions. But historical records sometimes contradict each other and calls of judgment are in practice impossible to avoid. Another common feature of narrative shock series is that many observations are censored to zero. These measurement problems invalidate the use of the narratives as direct observations of structural shocks and typically bias estimates in simple regressions.

Our methodology is robust to many of these measurement problems. As long as conditions (4)-(6) hold, the precise nature of the measurement error does not affect the identification of impulse responses. However, one advantage of making specific assumptions about the error in measurement is that the bias from ignoring such error can be made explicit. In our application below, this will help

[3] Examples include Romer and Romer (1989, 2010), Ramey and Shapiro (1998), Burnside, Eichenbaum and Fisher (2004), Cloyne (2011) and Ramey (2011a).

[4] Moreover, our approach offers a more parsimoniously parametrized alternative for narrative measures with relatively few nonzero observations (which is the norm in the literature). In addition, the estimator that we propose identifies not only impulse response functions, but also the entire realized shock sequence in the sample of observations for Y_t and thus permits for instance forecast error variance decompositions.

explain differences with traditional approaches. A second advantage of imposing more structure on the nature of the measurement error is that it allows the use of the statistical reliability of m_t as a diagnostic tool.

Consider an augmented system consisting of the SVAR in (2) and the following system of measurement equations,

$$m_t = D_t\left(\Gamma\varepsilon_{1t} + \upsilon_t\right)\ , \tag{9}$$

where Γ is a $k\times k$ nonsingular matrix, υ_t is a $k\times 1$ vector of measurement errors with $E[\upsilon_t]=0$, $E[\upsilon_t\varepsilon_{1t}']=0$, $E[\upsilon_t\upsilon_t']=\Sigma_{\upsilon\upsilon'}$ and $E[\upsilon_t\upsilon_s']=0$ for $s\neq t$. D_t is a $k\times k$ diagonal matrix containing random (0,1)-indicators tracking zero observations. We assume that the diagonal elements of D_t are perfectly correlated, i.e. when $k>1$ the proxy variables are identically censored. We also assume that $E[D_t\upsilon_t\varepsilon_{1t}']=0$, but we do not require that the censoring process D_t is independent of ε_{1t}. For instance, it is possible that larger realizations (in absolute value) of ε_{1t} are more likely to be measured. Note that (9) allows for both an additive correlated noise υ_t and an arbitrary scale for m_t.[5]

Combining (9) with the SVAR in (2) results in a system of structural equations with latent variables, as discussed in Bollen (1989). Rewrite the model as:

$$Y_t = \theta' X_t^* + w_t\ , \tag{10}$$

where $X_t^* = [Y_{t-1}',...,Y_{t-p}'\ ,\varepsilon_{1t}']'$, $\theta=[\delta',\beta_1]'$, $\delta=[\delta_1,..,\delta_p]'$ and $w_t=\beta_2\varepsilon_{2t}$. X_t^* is not fully observable because it contains ε_{1t}. The enlarged system is a measurement error model of the form

$$Y_t = \gamma'\bar{X}_t + z_t\ , \tag{11}$$

$$\bar{X}_t = \Omega X_t^* + \Upsilon_t\ , \tag{12}$$

[5] Other measurement error specifications may account for the possible discrete nature of m_t, for multiplicative errors, heteroskedastic errors, etc.

where $\bar{X}_t = [Y'_{t-1}, \ldots, Y'_{t-p}, m'_t]'$ and

$$\theta = \Omega'\gamma \ , \ \ w_t = z_t + \gamma'\Upsilon_t \ , \ \ \Omega = \begin{bmatrix} I & 0 \\ 0 & \Gamma \end{bmatrix} \ , \ \Upsilon_t = \begin{bmatrix} 0 \\ D_t\upsilon_t + (D_t - I_k)\Gamma\varepsilon_{1t} \end{bmatrix} .$$

Note that because of censoring, $E[X_t^* \Upsilon_t'] \neq 0$ and the measurement error Υ_t is not classical. From $\Sigma_{\bar{X}w'} = 0$, we obtain

$$\theta = \Omega' \Lambda_{\bar{X}}^{-1} \Sigma_{\bar{X}\bar{X}'}^{-1} \Sigma_{\bar{X}Y} \ , \tag{13}$$

where $\Lambda_{\bar{X}}$ is the reliability matrix of (the uncensored realizations) of $\bar{X}_t$, given by

$$\Lambda_{\bar{X}} = \begin{bmatrix} I & 0 \\ 0 & \Sigma_{mm'}^{-1}\Phi\Gamma' \end{bmatrix} . \tag{14}$$

Most existing narrative studies estimate a version of (11), often also including lags of m_t. But unless there is no measurement error, the resulting naive estimator $\Sigma_{\bar{X}\bar{X}'}^{-1}\Sigma_{\bar{X}Y}$ is generally biased because of scaling ($\Omega' \neq I$), and measurement error ($\Lambda_{\bar{X}}^{-1} \neq I$).[6] The elements of θ reduce to

$$\delta = \Sigma_{XX'}^{-1}\Sigma_{XY'} \ , \ \ \beta_1' = \Phi^{-1}\Sigma_{mY'} \ ,$$

and since $\Sigma_{mY'} = \Sigma_{mu'}$, the three stage procedure described in the previous section is equivalent to estimating a measurement error model in which Y_t has perfect reliability and m_t is measured with error.

Under the additional assumption of independent random censoring, we show in appendix A that it is possible to identify the reliability matrix (14). In that case, $\Phi = E[D_t]\Gamma$ and the $k \times k$ reliability matrix of m_t is given by

$$\Lambda = \Sigma_{mm'}^{-1} E[D_t]\Gamma\Gamma' \ , \tag{15}$$

[6]When $k > 1$ and the proxy variables are not identically censored and if in addition the off-diagonal elements of Γ are nonzero, then (14) needs to be further decomposed into a reliability matrix and yet another bias term that is due to omitted variables.

which is a generalization of the reliability ratio of a scalar measurement. When $k = 1$, Λ is the fraction of the variance in the uncensored measurements that is explained by the variance of the latent variable or equivalently the squared correlation between the narrative measure and the true structural shock of interest. Since $0 \leq \Lambda \leq 1$, measurement error bias manifests itself in this case as shrinkage towards zero. When $k > 1$, the bias can go in either direction. The eigenvalues of Λ can be interpreted as the scalar reliabilities of the principal components of the uncensored observations in m_t. The smallest eigenvalue of Λ corresponds to the smallest scalar reliability of any linear combination of m_t, see Gleser (1992). An estimate of Λ can be used for testing the hypothesis that some linear combination of m_t has scalar reliability zero. It provides a metric for evaluating how closely the proxy variables are related to the true shocks, and therefore for the estimability of the structural parameters and the quality of identification. SVAR shocks are sometimes criticized for being at odds with historical events or descriptive records, see for instance Rudebusch (1998). The reliability of proxies constructed from the historical record of policy changes quantifies the extent to which this criticism applies.

3 Do Tax Cuts Stimulate Economic Activity?

In this section we apply our methodology to the estimation of the impact of exogenous tax shocks on economic activity in the United States over the postwar period. Here we concentrate mainly on the effects on output. The subsequent section provides evidence for a broader set of macroeconomic aggregates.

The empirical analysis in this paper differs from existing estimates of the effects of unexpected changes in tax policy in three ways. First, we apply the new SVAR estimator presented above. Our candidates for the proxies that are used for identification purposes are narrative data for legislated federal tax changes. Second, we take several steps to ensure that our estimates are not affected by anticipation effects. Third, while much of the macro literature has estimated the impact of changes in the average total tax rate (or in total tax revenues), we investigate the impact of changes in more disaggregated average tax rates. Ideally, one would like to examine the effects of changes in very narrowly defined tax instruments. However, there are practical limits to the level of disaggregation determined by data availability. We concentrate on changes in two tax categories, personal income and corporate income taxes. In our sample, personal income tax revenues (we include contributions to social insurance in our definition of personal income taxes) have accounted for on average 74.2 percent of total federal tax revenues while corporate income taxes have accounted for 16.4 percent. Thus, the two components comprise the bulk of total federal tax revenue generation.

The macroeconomic literature instead often distinguishes between labor and capital income taxes, see e.g. Mendoza, Razin and Tesar (1994), Jones (2002) or Burnside, Eichenbaum and Fisher (2004), which is appealing in terms of economic modeling. However, the division into personal and corporate income taxes corresponds more closely to the actual policy instruments and observed changes in federal tax liabilities can much more easily be assigned to one of these tax categories. The next subsection describes the proxies for each of the two types of tax shocks.

3 .1 A Tax Narrative for Personal and Corporate Income Taxes

We produce a narrative account of legislated federal personal and corporate income tax liability changes in the US for a sample of quarterly observations covering the period 1950Q1 - 2006Q4. The narrative extends Romer and Romer's (2009a) analysis by decomposing the total tax liabilities changes recorded by Romer and Romer (2009a) into the following subcomponents: corporate income tax liabilities (CI), individual income tax liabilities (II), employment taxes (EM) and a residual category with other revenue changing tax measures (OT). We discard the latter group because it is very heterogeneous.[7] The decomposition is based on the same sources as Romer and Romer (2009a) supplemented with additional information from sources such as congressional records, the Economic Report of the President, CBO reports, etc. whenever required. In an online appendix, we describe the construction of the data and the historical sources in detail.

[7] II and EM tax changes include adjustments to marginal rates and various deductions and tax credits. CI tax changes include a few adjustments to marginal rates and otherwise mainly changes in depreciation allowances and investment tax credits. The other tax changes mostly include excise taxes, often targeted to specific industries (transportation) or goods (gasoline, automobiles, sporting goods,...), and gift and estate taxes. See the online data appendix for details.

To comply with condition (5), which requires that the proxies are orthogonal to all non-tax structural shocks, we retain only those changes in tax liabilities that were unrelated the current state of the economy. To this end, we adopt Romer and Romer's (2009a) selection of exogenous changes in tax liabilities, which is based on a classification of the motivation for the legislative action either as ideological or as arising from inherited deficit concerns. Another important issue is that many changes in the tax code are legislated well in advance of their scheduled implementation. In Mertens and Ravn (2011a) we distinguish between unanticipated and anticipated exogenous tax changes on the basis of the implementation lag (the difference between the dates at which the tax change becomes law and goes into effect). Around half of the exogenous changes in tax liabilities were legislated at least 90 days before their implementation and in Mertens and Ravn (2011a) we show that there is evidence for macroeconomic effects of legislated tax shocks prior to their implementation. These findings mean that condition (6) may fail to hold for the subset of pre-announced tax changes. For that reason, we retain only those exogenous tax changes for which the legislation and implementation date are less than one quarter apart.

Romer and Romer (2009a) describe almost 50 legislative changes in the tax code over the sample period, many containing multiple changes in tax liabilities implemented at different points in time. Our narrative measures are a much smaller subset because we eliminate all endogenous and/or pre-announced tax changes. Our tax narrative contains 13 observations of individual income tax liability changes, 2 observations for employment tax liability changes and 16 observations for corporate income tax liability changes deriving from 21 separate legislative changes to the federal tax code. The vast majority of these changes were legislated as permanent changes to the tax code. Because there are too few observations for a separate employment tax category, we merge the EM and II taxes into a personal income (PI) tax category. All our results are very similar if we instead leave out the employment taxes.

We convert the tax liability changes into the corresponding average tax rate changes as follows,

$$\Delta T_t^{CI,narr} = \frac{\text{CI tax liability change}_t}{\text{Corporate Taxable Income}_{t-1}} \ ,$$

$$\Delta T_t^{PI,narr} = \frac{\text{II tax liability change}_t + \text{EM tax liability change}_t}{\text{Personal Taxable Income}_{t-1}} \ .$$

We scale the tax liability changes by previous quarter taxable income, but our results are nearly identical if we instead scale by the contemporaneous or previous year taxable income. The resulting narrative measures are depicted in Figure 1 together with the corresponding average tax rates computed from the national income and product (NIPA) accounts. The average personal income tax rate (APITR) is defined as the sum of federal personal current taxes and contributions to government social insurance divided by personal income less transfers plus contributions for government social insurance. The average corporate income tax rate (ACITR) is constructed as federal taxes on corporate income excluding Federal Reserve banks as a ratio of corporate profits. Appendix B provides further details on the data.

The two average tax rates vary considerably over time. The average rates are very broadly defined and are affected by legislative adjustments to tax rates, tax brackets as well as changes in tax expenditures. The average tax rates also display endogenous movements unrelated to legislative actions that occur for a variety of reasons, such as cyclical fluctuations in the administrative definition of taxable income versus NIPA income, tax progressivity and changes in the distribution of income, cyclical variations in tax compliance and evasion, etc. Even though total federal revenues as a share of GDP have remained fairly stationary around 18 percent, the APITR and ACITR measures both display trends over the sample. Figure 1 shows that the APITR has slowly risen from around 10 percent at the beginning of the sample to approximately 18 percent at the end of 2006. The two most significant exogenous changes in personal income taxes relate to the Revenue Act of 1964, which reduced marginal tax rates on individual income, and to the Jobs and Growth Tax Relief Reconciliation Act of 2003, which reduced marginal tax rates on individual income, capital gains and dividends and increased some tax expenditures. Each of these two pieces of legislation cut average personal income tax rates by more than one percentage point according to the narrative measure. The ACITR instead has fallen significantly over time from over 50 percent in the early 1950s to just above 20 percent at the end of the sample period. The narrative measure indicates several sizeable changes in corporate income taxes, the biggest one being a large increase in corporate tax liabilities associated with the repeal of the investment tax credit included in the Tax Reform Act of 1986.

We use the new tax narratives depicted in Figure 1 as proxies for structural tax shocks. Unless mentioned otherwise, the proxies are simply the demeaned narrative shocks. We checked whether lagged macro variables have predictive power for the narrative series but on the basis of standard Granger causality tests we found no such evidence.[8] We also tested for predictive power in regressions of only the nonzero observations of the measured tax shocks and lagged values of key variables but did not detect any statistical significance.

3 .2 Identifying Tax Shocks

To obtain valid covariance restrictions from the proxy variables m_t, it is essential that the measured tax changes are uncorrelated with lagged endogenous variables and non-tax structural shocks. It is however also important to consider whether measured changes in personal income taxes are uncorrelated with structural shocks to corporate taxes, and vice versa. If so, then each of the two proxy variables can be used sequentially to derive $n-1$ restrictions, or $2(n-1)$ in total. In combination with the residual covariance restrictions, each set of $n-1$ restrictions suffices to identify the impulse response to the respective tax shock, see appendix A. If we cannot impose zero cross-correlations between the measured tax changes and structural tax shocks, the identifying assumptions on the combined proxy series yield only $2(n-2)$ restrictions, which is insufficient to disentangle the causal effects of shocks to both types of taxes.

[8] Tests of the null hypothesis that the average tax rate, GDP, government spending and the tax base do not Granger cause the narrative shock measure have p-values of 0.70 for the PI tax shock measure and 0.76 for the CI tax shock measure. For the variables of our benchmark system, the p-values are 0.87 and 0.57. For these tests we used first differences for the variables as the test is problematic when the data is nonstationary. We also performed tests for a range of other variables such as municipal bonds spreads and government debt. The smallest p-value (0.23) we found was for the null hypothesis that the government debt to GDP ratio does not Granger cause the corporate narrative measure.

Conditional on a tax change taking place, the correlation between the PI and CI narrative tax rate changes in our sample is 0.42. Insofar that this positive correlation is not just due to chance or correlated measurement error, it appears inappropriate to treat the narrative PI (CI) tax changes as uncorrelated with exogenous shocks to the corporate (personal) tax rate. The positive correlation between the measured changes in personal and corporate taxes is natural for a number of reasons. The tax narratives record changes in tax liabilities for which the historical documents indicate that they were not explicitly motivated by countercyclical considerations. Yet they of course still occurred with certain objectives in mind, typically related to longer run goals for economic growth or debt reduction. When both personal and corporate income taxes are adjusted simultaneously, it is therefore not surprising that they are often adjusted in the same direction. Also, given that the tax narratives are based on actual legislative actions, the fixed costs of passing legislation naturally imply a temporal correlation of the changes in different types of taxes.

For isolating the causal effects of a change in only one of the tax rates, it is thus important to control for changes in the other tax rate, which in turn requires more restrictions. To see how we address this additional identification challenge, consider the following parametrization of the structural relationship between the VAR residuals u_t and the structural shocks ε_t:

$$\begin{aligned} u_{1t} &= \eta u_{2t} + S_1 \varepsilon_{1t} \ , \\ u_{2t} &= \zeta u_{1t} + S_2 \varepsilon_{2t} \ , \end{aligned} \tag{16}$$

where u_{1t} and ε_{1t} are the 2×1 vectors of reduced form and structural tax rate innovations, whereas the $(n-2) \times 1$ vectors u_{2t} and ε_{2t} contain the reduced form residuals and other structural shocks associated with an arbitrary number of additional variables. The matrices η, ζ, S_1 and S_2 contain the structural coefficients that underlie $\mathcal{B}$. In particular, the 2×2 non-singular matrix S_1 is not necessarily diagonal, capturing the potential contemporaneous interdependence of the tax instruments.

Obtaining the responses to ε_{1t} requires identification of β_1, the first two columns of $\mathcal{B}$, which is given by

$$\beta_1 = \begin{bmatrix} I+\eta(I-\zeta\eta)^{-1}\zeta \\ (I-\zeta\eta)^{-1}\zeta \end{bmatrix} S_1 \ .$$

In appendix A, we show that the narrative restrictions in (8) allow for the identification of the first term in square brackets, $\beta_1 S_1^{-1}$, as well as $S_1 S_1'$, the covariance of $S_1\varepsilon_{1t}$. The covariance restrictions are however not sufficient to obtain the structural decomposition of this covariance and obtain S_1. Ideally one would like to identify the structural relationship between exogenous changes in tax rates. This however would require, in our view, arbitrary assumptions on how personal income taxes respond contemporaneously to unanticipated changes in corporate taxes (beyond the indirect contemporaneous endogenous effects through u_{2t}), and vice versa. Fortunately, knowledge of $\beta_1 S_1^{-1}$ still permits economically meaningful structural responses to any linear combination of tax shocks ε_{1t}. We report the responses that result from a Choleski decomposition of $S_1 S_1'$, imposing that S_1 is lower triangular. Suppose for instance that the APITR is ordered before the ACITR. Then the response to a negative one percentage point ACITR shock is the response to an exogenous tax change that lowers the ACITR by one percentage point but leaves the APITR unchanged in 'cyclically adjusted' terms, i.e. after allowing for all the contemporaneous feedback from u_{2t}. A shock to the APITR on the other hand induces a change in the ACITR through feedback from u_{2t} as well as a direct response to the APITR shock that is determined by the identified correlation between both tax rates. If $S_1 S_1'$ is diagonal, the latter correlation is zero and the responses will be identical for different orderings of the tax rates. In summary, our strategy is to complement the identifying restrictions provided by m_t with a recursivity assumption and to verify robustness to the ordering of the tax rates.

3 .3 Benchmark Specification and Results

Our benchmark estimates for the dynamic output effects of tax changes are based on a VAR with seven variables: $Y_t = [T_t^{PI}, T_t^{CI}, \ln(B_t^{PI}), \ln(B_t^{CI}), \ln(G_t), \ln(GDP_t), \ln(DEBT_t)]$. T_t^i is the average tax rate of tax type $i = PI, CI$, i.e. federal personal and corporate income tax revenues as a fraction of the respective taxable income categories; B_t^i is the real per capita personal and corporate taxable incomes, respectively; G_t is real per capita government purchases of final goods; GDP_t is real per capita gross domestic product; Finally, $DEBT_t$ is the real federal government debt per capita.[9] All fiscal variables are for the government at the federal level. The sample consists of quarterly observations for the period 1950Q1-2006Q4. Precise data definitions are provided in appendix B. Based on the Akaike information criterion, the lag length in the VAR is set to four. All impulse responses are for a one percentage point decrease in either of the two tax rates over a forecast horizon of 20 quarters. Along with the point estimates, we report 95% percentile intervals computed using a recursive wild bootstrap, see Gonçalves and Kilian (2004), using 10,000 replications.[10]

Figures 2 and 3 show the effects of cuts in average personal and corporate income tax rates for each ordering of the tax rates. The correlation between the cyclically adjusted tax rate innovations $S_1\varepsilon_{1t}$ is small and estimated at -0.07 with a 95% confidence interval $[-0.41, 0.50]$. As a result, the responses are very close across the different tax rate orderings. This turns out to be a robust finding in sufficiently large VAR systems, in particular when they include a measure of government debt. When discussing a shock to a tax rate, for brevity we therefore only discuss the point estimates resulting from ordering that tax rate last, leaving the other tax rate unchanged in cyclically adjusted terms.

Figure 2 shows that after an initial one percentage point cut in personal income taxes, the APITR remains significantly below the level expected prior to the shock during the first year. Thereafter, the

[9]Government debt is a potentially important variable since any change in taxes (eventually) must lead to adjustments in the fiscal instruments. Especially if the reaction to debt is strong and relatively fast, it might be inappropriate not to explicitly allow for feedback from debt to taxes and spending.

[10]Letting $\hat{\delta}$ and $\hat{u}_t$ denote the OLS estimates, we generate bootstrap draws Y_t^b recursively using $\hat{\delta}$ and $\hat{u}_t e_t^b$, where e_t^b is the realization of a random variable taking on values of -1 or 1 with probability 0.5. We also generate a draw for the proxy variables $m_t^b = m_t e_t^b$, re-estimate the VAR for Y_t^b and apply the covariance restrictions implied by m_t^b. The intervals reported are 95% percentile intervals of the resulting distribution of impulse response coefficients. This procedure requires symmetric distributions for u_t and m_t but is robust to conditional heteroskedasticity. It also takes into account uncertainty about identification and measurement. This contrasts with the typical application of coefficient restrictions in SVARs as well as narrative specifications, which often treat m_t as deterministic. The standard residual bootstrap is problematic given that m_t contains many zero observations, which means that drawing with replacement from m_t yields zero vectors with positive probability.

APITR gradually converges to pre-shock expected levels in the longer run. The cut in the APITR sets off a significant increase in the personal income tax base which initially rises approximately 0.6 percent and peaks at 1.3 percent one year after the tax cut. Combining the responses of the tax base and the personal income tax rate, the decrease in the APITR implies a drop in personal income tax revenues of 5.4 percent upon impact.[11] Tax revenues remain relatively low until several years after the shock, but recover substantially from the initial drop during the first year. Despite the increase in the tax base we find that cuts in personal income taxes unambiguously lower personal tax revenues. Most importantly, cuts in average personal income taxes provide a substantial short run output stimulus. In our benchmark specification, we find that a one percentage point decrease in the APITR leads to an increase in output of 1.4 percent in the first quarter and a peak increase of 1.8 percent which occurs three quarters after the tax cut. The confidence intervals indicate a significant increase (at the 95% level) in economic activity within a two year window after the tax cut.

Figure 3 shows the effect of a one percentage point decrease in the average corporate income tax rate. The cut in the ACITR leads to a prolonged period of lower average corporate income tax rates. The cut in the ACITR induces a large and significant increase in the corporate income tax base which rises by up to 3.8 percent in the first 6 months. The increase in the tax base is sufficiently large such that there is only a very small decline in corporate income tax revenues in the first quarter and a surplus thereafter. The response of corporate tax revenues is however insignificant at every horizon. Hence, cuts in corporate income taxes appear to be approximately self-financing which is suggestive of particularly strong behavioral responses to changes in effective corporate tax rates.[12] The output effects of ACITR cuts are again significant and substantial. A one percentage point decrease leads to a rise in real GDP of around 0.4 percent rising up to 0.6 percent about one year after the cut.

In accordance with Romer and Romer (2009b), we find little impact of either tax shocks on gov-

[11]The response of tax revenues are computed as $\widehat{tr}_t = \hat{T}^i_t/\bar{T}^i + \hat{b}^i{}_t$ where $\bar{T}^i$ is the mean average tax rate of type i in the sample, $\widehat{x}_t$ denotes the impulse response of x_t and lower case letters denote logged variables.

[12]See Trabandt and Uhlig (2009) for an argument based on a calibrated DSGE model for why there is little scope for raising tax revenues with capital income taxes in the US.

ernment spending. Figure 2 shows that the response of government spending to an APITR tax cut is not significantly different from zero at the 95% level at all forecast horizons. Similarly, there is little evidence that changes in the ACITR impact systematically on government spending. This is reassuring since it refutes the possibility that the responses to tax shocks are confounded with changes in government spending. We also find that cuts in one average tax rate lead to increases in the other average tax rate, although neither of these increases is significant. The mutual tax rate responses indicate that our orthogonalization scheme successfully disentangles the effects of different tax instruments. Government debt (not shown) increases significantly at the 95% level in the short run after an APITR cut, but does not change significantly after an ACITR cut. The debt response is more precisely estimated in specifications that include interest rates, which are discussed later.

Under the additional measurement error assumptions of section 2 .2, our procedure also allows for the identification of the reliabilities of the proxy variables. The estimated reliability matrix has eigenvalues of 0.30 and 0.69 with 95% confidence intervals $[0.16, 0.48]$ and $[0.47, 0.97]$. This implies that the principal components of the narrative tax changes have correlations with the true tax shocks of 0.55 and 0.83 respectively. The former number is also the smallest correlation of any linear combination of the proxy variables. These statistics indicate that the proxies contain valuable information for the identification of the structural tax shocks and that there is a reasonably strong connection between the SVAR shocks and historically documented legislative changes to the tax code. At the same time, the fact that the reliability matrix has eigenvalues substantially below unity is an indication that measurement error is a serious concern in practice.

Perhaps the most important result in this paper is that the estimated short run output effects of changes in average tax rates are large. Another common metric for these effects is the tax multiplier, defined as the dollar change in GDP per effective dollar loss in revenues. Multipliers can be obtained in our SVAR by rescaling the output response such that the implied drop in tax revenues is normalized to 1 percent of GDP. For the personal income tax we find a multiplier of 2.0 on impact rising to a maximum of 2.5 in the third quarter. Since for corporate taxes the impact on revenues is almost nonexistent, the corporate tax multiplier is poorly defined.[13]

3 .4 Discussion and Relationship to the Literature

In order to gain some further understanding of the benchmark results, we elaborate on several aspects of our estimation procedure. First, we discuss the importance of allowing for nonzero cross-correlations between the measured tax changes and structural tax shocks. Next, we compare our results to those that obtain from more standard approaches in the narrative identification literature. Finally, we analyze the role of using average versus marginal tax rates and compare our findings with some of the existing results in the literature.

Correlation Between the Proxies and Tax Shocks Given the positive correlation between the narrative measures, it is likely that the measured changes in one tax rate are correlated with shocks to both tax rates. The benchmark specification controls for simultaneous changes in both tax rates and resolves the shortage of identification restrictions with an additional recursivity condition. Here we analyze the consequences of replacing the recursivity assumption with the alternative assumption that each of the proxies is correlated with only a single tax shock. This assumption effectively ignores the observed correlation between the proxies, which is only valid if the correlation is due to chance or correlated measurement errors. In practice it means that each of the proxies can be used in isolation to identify the corresponding impulse response functions.

Figures 4 and 5 show the impulse responses of output following a one percentage point decrease in either of the two tax rates when estimating each of the tax shocks separately using only a single proxy at a time. The specification is otherwise identical to the benchmark. For comparison, both figures also show the impulse responses from the benchmark specification that result from ordering the tax rate that is shocked last, as well as the associated percentile intervals.

Figure 4 shows that a cut in the APITR identified with a single proxy leads to a persistent de-

[13] See also Barro and Redlick (2011) on this point.

crease in the APITR that is similar to the benchmark. Figure 5 shows the same is true for the ACITR cut. However, the output effects that follow the tax cuts depend importantly on whether one controls for the correlation between the proxies or not. When the correlation is ignored we find much larger effects of corporate income tax cuts than in the benchmark specification, while the opposite pattern is evident for the personal income tax cut. The sizeable differences suggest that, when using the disaggregated narrative tax changes for identification, it is important to control explicitly for the interactions between the different tax instruments. The impact of ignoring the correlation between the proxies is most dramatic when both average tax rates are included in the vector of observables, as is the case in the benchmark specification. In smaller specifications that include only the average tax rate and tax base associated with the tax of interest, the impulse responses identified with a single proxy are typically much closer to those of our benchmark specification.

Comparison with Traditional Narrative Approaches To demonstrate the relevance of our estimation methodology relative to standard narrative approaches, we compare our estimates to those of the following two specifications (omitting constants):

$$\Delta \ln(GDP_t) = \sum_{s=1}^{K} \beta_s \Delta T_{t-s+1}^{i,narr} + u_t \tag{17}$$

$$Y_t = \sum_{i=1}^{p} \delta_i Y_{t-i} + \beta \Delta T_t^{i,narr} + u_t \tag{18}$$

where $\Delta T_t^{i,narr}$ ($i = PI, CI$) are the narratively identified tax changes. The first of these specifications is a simple regression of output growth on the contemporaneous and lagged narrative, which is the approach of Romer and Romer (2010). The second specification in (18) adopts a reduced form VAR that includes the narrative as an exogenous regressor, as in for instance Favero and Giavazzi (2011). When estimating (17) we set $K = 12$. Figure 6 depicts the resulting impulse response functions to one percentage point cuts in $\Delta T_t^{i,narr}$ together with the results from the benchmark SVAR.

The models in (17)-(18) imply substantially smaller point estimates of the output effects of tax changes than the benchmark estimates. This is particularly evident for the corporate income tax cut where the output responses derived from (17) and (18) are close to zero at all forecast horizons and significantly smaller than our estimates. For the personal income tax, the output responses produced by (18) are smaller than our estimates at all forecast horizons and significantly so during the first 3 quarters after the tax shock. Specification (17) also delivers estimates of the impact of cuts in the average personal income tax rate that are considerably smaller at all horizons apart from a few quarters around the two year horizon.

The finding that our estimation approach yields larger output responses to tax cuts in the short run also extends to using the aggregate measures of tax shocks as Romer and Romer (2010) and Favero and Giavazzi (2011), see Mertens and Ravn (2011c). The main reason can be found in measurement problems. First, there is an important difference in scaling since we scale the shocks by their impact on effective average tax rates while the Romer and Romer (2010) multiplier estimates are based on *projected* tax liability calculations. These are in turn based on the assumption that output (and other determinants of tax revenue) does not respond to changes in taxes. Since we find that economic activity expands following a tax cut, the tax changes implicit in $\Delta T_t^{i,narr}$ are smaller than those assumed in the estimates we report. Secondly, our estimator allows for the presence of measurement error in the narrative accounts. We showed above how this can bias the estimated output responses, typically in a downward direction.[14] Our estimates of the reliability of the proxies indicate that measurement error bias is also quantitatively relevant. Interestingly, Perotti (2011) updates the Romer and Romer (2009a) series with the aim to improve measurement and as a result also finds tax multipliers that are relative larger.

Comparison with Existing Estimates in the Literature There are relatively few studies which we can use for direct comparison, as most macro estimates are for shocks to total taxes. A notable exception is Barro and Redlick (2011), who estimate the impact of changes in a measure of taxes

[14] In the context of our measurement equation assumptions, specification (18) necessarily suffers from attenuation bias. One should not jump to the conclusion that all narrative results in the literature are downward biased because of measurement error. When lagged or multiple narrative measures are included, measurement error can lead to attenuation or expansion bias. Some studies, such as Ramey (2011a), rescale impulse responses according to the impact on one of the observables, which can substantially mitigate the problem.

related to our APITR variable. Using annual data, they consider the output response to changes in average marginal income tax rates (AMTRs) which includes state taxes and excludes most forms of capital income taxes. In contrast, our measure excludes state income taxes, and includes capital income taxes that are not classified as corporate income taxes. Identification in Barro and Redlick (2011) relies on using the year-aggregated Romer and Romer (2009a) series for exogenous total tax liability tax changes at the federal level as an instrument for AMTR shocks, without removing pre-announced tax changes. Based on annual data they find that a one percentage point cut in the AMTR increases next year GDP by 0.5 percent, corresponding to a tax multiplier of around 1.1. Our benchmark estimates indicate output effects that are considerably larger for changes in federal average personal income tax rates.

The shocks to average tax rates that we identify reflect changes to marginal tax rates, as well as tax brackets and tax expenditures, all of which in principle have distinct influences on economic decisions. Shocks to average marginal rates arguably have a more straightforward structural interpretation. The drawback of using marginal rates is the annual frequency and that, to our knowledge, no good data is available for corporate taxes. Figure 7 plots the annual NIPA-based APITR variable as well as the average marginal tax rate constructed by Barro and Redlick (2011). For a better comparison, we exclude the contribution of state taxes from their AMTR variable. The two tax rates are highly correlated: 0.90 in levels and 0.62 in first differences. To assess the role of using average versus marginal rates, we identify shocks to personal income tax rates in an SVAR with annual data and two lags of the endogenous variables. To keep the dimension of the VAR manageable as well as mitigate concerns about the correlation between the tax changes, we include the benchmark variables but omit the corporate tax rate and base. As the tax rate measure T_t^{PI}, we sequentially use the APITR and AMTR variables depicted in Figure 7, and rely on the time aggregated narrative APITR series for identification.

Figure 8 compares the effect of a one percentage point cut in the tax rates. The output response to a marginal rate cut is highly significant and very similar in size to our benchmark estimates.

The output response to the average rate cut is somewhat larger in the annual data. Overall, using marginal rates delivers results that are broadly similar to our specifications with quarterly frequency and both average rates. Interesting differences are that the decline in the marginal rate is more persistent and that the confidence intervals are much narrower when using the marginal rate. Besides other methodological differences, one possible explanation for why our estimates are higher than in Barro and Redlick (2011) is that including pre-announced tax changes leads to a downward bias. This is because forward looking agents and intertemporal substitution motives generate a tendency for pre-announced cuts in income taxes to lower output prior to implementation, see Yang (2005), Mertens and Ravn (2011a,b,c) and Leeper, Walker and Yang (2011) for theory and evidence.[15]

Blanchard and Perotti (2002) estimate the impact of shocks to total tax revenues using an SVAR estimator. They find an impact multiplier of 0.69 and a peak multiplier of 0.78 in quarterly US data for the sample period 1947-1997. Even though they include tax revenues at all levels of government, our estimates imply significantly larger effects on economic activity. Mertens and Ravn (2011c) provide a detailed analysis of this result and argue that the key discrepancy relates to the elasticity of tax revenues to output.[16] Mountford and Uhlig (2009) also analyze shocks to aggregate tax revenues identified using sign restrictions. In response to a deficit financed tax cut, they estimate multipliers of 0.29 on impact, 0.93 after one year and up to 3.41 at twelve quarters. These numbers are much larger at longer horizons, but similar to Blanchard and Perotti (2002) in the short run. This contrasts with our finding of large output effects in the shorter run. Romer and Romer (2010) estimate the impact of innovations to their aggregate tax liability narrative and find that a one percent drop in legislated tax liabilities relative to GDP leads to an increase in GDP of less than half a percent on impact growing steadily to a 3 percent increase at the 10 quarter horizon. Again, these estimates are not directly comparable to ours since we consider disaggregated taxes, but as with the SVAR based estimates the main difference is that we find large output effects in the short run.

[15] The output response to a marginal rate cut is somewhat closer to Barro and Redlick (2011) when we do not remove state taxes. The first-year output response in that case is 0.7 percent, rising to 1.7 percent in the third year.

[16] Blanchard and Perotti (2002) calibrate the output elasticity of tax revenues to 2.08 while in Mertens and Ravn (2011c) we estimate a larger elasticity of 3.13 based on the narrative data. The discrepancy explains the entire difference between tax multiplier estimates.

3 .5 Robustness

We have investigated the robustness of our main results with respect to several issues. For brevity we refer to an online appendix for the figures and more detail.

The SVARs are estimated in log levels and the responses at long forecast horizons are typically imprecisely estimated. It is possible to make more specific assumptions about the long run statistical properties of the time series and SVAR results can be somewhat sensitive to different assumptions about trends, as in for instance Blanchard and Perotti (2002). We verified our results for a specification with the observables in first differences and another with a deterministic linear-quadratic time trend. The key features of the short and medium run effects of tax shocks, our primary focus, are insensitive to these alternatives. However, different trend assumptions are important at longer forecast horizons and determine whether tax changes are permanent or temporary. In terms of economic theory, whether displacements in tax rates are perceived by agents as permanent or transitory does matter importantly, see for instance Chetty, Guren, Manoli and Weber (2012).

To avoid problems related to anticipation effects, we eliminated all tax liability changes that were implemented more than 90 days after the relevant tax changes became law. In Mertens and Ravn (2011a) we find no significant effects in the quarters leading up to aggregate tax changes that we classified as unanticipated. One might still worry that we do not fully address the possibility of tax foresight as tax changes may have been anticipated even before legislation. The mistiming of shocks and/or the omission of an important variable can potentially lead to misleading results, see Leeper, Walker and Yang (2011), Ramey (2011a) and Mertens and Ravn (2010). We looked at measures of expected future taxes derived from municipal bond prices constructed by Leeper et al. (2011). Municipal bonds are exempt from federal income taxation in the US and the spread between the yields on municipal bonds and similar tax nonexempt bonds may therefore contain information about the market expectation of the present value of income taxes over the maturity of the bond, see for instance Poterba (1988) and Fortune (1996). A measure of implicit expected future taxes can be derived in combination with a no arbitrage assumption, see Leeper et al. (2011) for details. We used their measure for bonds with maturity of one and five years and added them as additional controls to the benchmark specification. We found no evidence that the large output effects of tax cuts are sensitive to controlling for municipal yield spreads. We also used the spreads in Granger causality tests and as explanatory variables in regressions with nonzero narrative tax changes, but we did not detect any significant predictive power. As a final check, we also ran the benchmark specification after first regressing the nonzero observations of our narrative tax measures on lags of the implicit expected tax rate variables and then using the residuals as the proxies for the structural shocks. The point estimates derived from these alternative proxies remain similar to the benchmark specification. In conclusion, our approach of eliminating pre-announced tax changes from the narrative tax measures on the basis of implementations lags appears successful in dealing with tax foresight.

Anticipation effects may be relevant not only for tax changes but also for government spending. While our interest is in estimating the impact of tax shocks, pre-announced changes in government spending that are not controlled for may also give rise to problems of omitted variable bias and misalignment of the information sets of the econometrician and economic agents. Ramey (2011a) for instance argues that anticipation effects are crucial for the identification of government spending shocks. We extended the vector of observables of the benchmark specification with an asset price that is likely to contain information about future changes in government spending. In particular, we included a series for the accumulated excess returns of large US military contractors constructed by Fisher and Peters (2010). Alternatively, we included Ramey's (2011a) defense spending news variable in the vector of observables. This narrative is based on professional forecasters' projections of the path of future military spending. Including these sources of information about future changes in government spending did not lead to dramatic changes in the output responses.

One may worry that the narrative tax changes are correlated with the inherited level of government debt, especially since a few of the legislative changes were explicitly motivated by budgetary concerns. In Granger causality tests and regressions of nonzero tax shock observations on lagged debt-to-GDP, we did not find any formal evidence for a significant relationship. Using the residual from a regression of the tax shocks on debt-to-GDP as alternative proxies also has very little effect on the results.

A different potential measurement problem is error in the timing of the tax changes. We verified the sensitivity of our benchmark estimates with respect to this issue by conducting simulation experiments similar to Ramey (2011a). The estimated output responses remain fairly stable when we assume that up to 50% of the measured tax change is randomly mistimed by one quarter, either as a lead or a lag. Note that if some of the narrative tax changes pre-date the true tax shocks, none of our identifying assumptions are violated. Our approach is therefore already robust to this type of timing error, which merely results in a loss in precision and lower reliability statistics. Post-dating of the true tax shocks is more serious because it leads to a violation of the identifying assumptions. This type of error is closely related to the problem of tax foresight which we already addressed above.

4 The Wider Macroeconomic Effects of Tax Changes

One advantage of the narrative identification approach is that it is straightforward to estimate the effects of shocks on other macroeconomic variables. Looking beyond the impact of tax changes on output or revenues allows us to gain further insight into how tax changes are transmitted to the economy and into possible differences between both tax components. In this section we consider a set of alternative VAR systems. Each of these consists of a fixed set of five baseline variables containing the two average tax rates, output, public debt and government spending, and varying set of additional variables. We consider in turn variables related to monetary policy, the labor market and private consumption and investment. As in the benchmark specification, the estimates and confidence intervals are always very similar for different orderings of the tax rates. For brevity, we only report the response to a shock to a tax rate resulting from ordering that tax rate last, leaving the other unchanged in cyclically adjusted terms.

4 .1 Monetary Policy and Inflation

Changes in taxes may impact on costs of production and, to the extent that cost changes are passed into prices, may affect inflation. The sign of the inflation response is indicative of whether the expansionary effects of tax cuts are primarily derived from increased demand or supply for final goods. The impact of taxes on inflation is also relevant because it may lead to monetary policy adjustments that in theoretical models are typically very important in determining the ultimate effects of fiscal shocks. Thus, it is important to investigate the extent to which explicitly including monetary policy instruments in the vector observables impacts on the results. To this end, we added the following series to the five baseline variables: the effective federal funds rate, the (log) level of nonborrowed reserves and the (log) level of the price index for personal consumption expenditures. Figure 9 depicts the impact of tax changes in the SVAR that includes the monetary variables, along with both 90% and 95% percentile intervals.

The first row of Figure 9 shows that the output stimuli provided by both types of tax cuts are similar in size and timing to the benchmark specification, and our main conclusions are insensitive to the inclusion of the monetary policy instruments. The second row reports the response of real federal government debt per capita, which turns out to be more precisely estimated with the inclusion of the monetary variables.[17] Government debt increases persistently after an APITR cut although the effect is only statistically significant at the 95% level in the first two quarters. Consistent with the absence of any impact on revenues, there is no significant effect on debt from a cut in the corporate tax.

A cut in the APITR is mildly disinflationary on impact and briefly inflationary in the third quarter, but none of these effects are significant at the 90% or 95% levels. We find a stronger negative impact of a cut in the ACITR on the inflation rate in the short run and, in contrast to the results for the APITR, the decline in inflation is persistent and statistically significant at the 95% percent level in the first two quarters. The short run disinflationary effects of corporate tax cuts are robust to using alternative measures of the nominal price level, such as the GDP deflator or the BLS consumer price

[17] Our interpretation is that including a nominal interest rate leads to better estimates of government debt dynamics.

index. The drop in inflation after a corporate tax cut is consistent with a fall in marginal costs and dominating supply side effects. The evidence for changes in personal income taxes is inconclusive.

There is no strong evidence that changes in either of the two tax rates impact significantly on the short term nominal interest rate, as measured by the funds rate, and we found the same when using the 3 month T-Bill rate. For the APITR this result is not too surprising given there is no clear impact on the inflation rate. For the ACITR instead, the short run decline in the inflation rate following a tax cut might instead have been expected to trigger a stronger monetary policy accommodation. There are various possible explanations including that the drop in inflation is accompanied by an increase in aggregate activity and that the impact on inflation is transitory. In any case, there is no evidence for a systematic short run interaction between monetary and discretionary tax policies in our sample.

4 .2 Labor Market

The labor market often takes center stage in discussions on fiscal policy. Romer and Bernstein (2009), for example, argue that "*Tax cuts, especially temporary ones, and fiscal relief to the states are likely to create fewer jobs than direct increases in government purchases.*" However, systematic empirical evidence on the dynamic effects of fiscal interventions on employment is surprisingly scarce. Ravn and Simonelli (2007) and Monacelli, Perotti and Trigari (2010) find that positive shocks to government spending impact negatively on the unemployment rate, but the response is very slow. Monacelli, Perotti and Trigari (2010) investigate the effects of tax shocks on unemployment and other labor market variables and find that tax cuts lead to delayed but sizeable reductions in unemployment.

To investigate the impact of tax changes on the labor market we add the following three variables to the baseline vector of observables: the log of total employment per capita, the log of hours worked per worker and the log of the labor force relative to population, all for the aggregate business, government (including military) and non profits sectors (see the appendix for precise data definitions). Combining these variables, we can also derive estimates of the impact of tax shocks on the unemployment rate. Figure 10 depicts the impact of a one percent cut in the APITR (left column) and in the ACITR (right column) on the new variables. The responses of the other variables, including output, are comparable to the benchmark and are therefore not shown.

Cuts in personal income taxes boost employment and do so relatively quickly. A one percentage point decrease in the APITR leads to a statistically significant rise in employment per capita of 0.3 percent on impact. The employment response peaks at around 0.8 percent five quarters after the tax stimulus. The labor input response to an APITR tax cut is however not restricted to the extensive margin. The number of hours worked per worker also rises significantly on impact by 0.4 percent and the impact remains significantly positive for the first year. In contrast to the fairly elastic short run responses of labor input at both the intensive and extensive margins, we find no evidence for significant effects on labor force participation at any horizon. This is perhaps not surprising given that, the reduction in the APITR is fairly transitory, and may therefore provide only limited incentives to enter the labor market. The increase in employment and lack of any effect on participation together imply a decrease in the unemployment rate of 0.3 percentage points on impact and a maximum decrease of slightly more than 0.5 percentage points in the fifth quarter after the tax cut.

The results for the ACITR depicted in the right column of Figure 10 indicate that changes in corporate taxes have much less pronounced effects on the labor market. In contrast to the personal income tax cut, there is no evidence that a cut in corporate taxes is associated with any significant impact on employment, despite the considerable and significant immediate increase in output. Instead, there is a gradual rise in employment that however never becomes statistically significant. The maximum increase in employment after a one percent cut in the ACITR is 0.3 percent. Another difference with the cut in personal income taxes is that there is no significant impact on hours worked per worker at any horizon. As was the case with the APITR cut, labor force participation is unaffected. We find that a cut in corporate taxes lowers the rate of unemployment after a few quarters, but the effect is very gradual and never statistically significant.

An interesting question is how the labor market effects are distributed across the public and private sector. We repeated the analysis above for employment in the two sectors (see the on-line appendix for details) and found that the positive response of total employment to a cut in average personal income taxes is composed of a more strongly positive private sector employment response and a temporary drop in public sector employment. The private sector employment response to a cut in corporate taxes is close to the response of total employment, while public sector employment drops marginally for two quarters after the tax cut.

We draw two conclusions from our study of the labor market effects of tax changes. First, there are important differences in how personal and corporate income tax changes affect the labor market. Studies that focus exclusively on total average tax rates or revenues are therefore only of limited use for assessing the ability of tax policy to affect employment at various horizons. The second conclusion is that when the prime policy objective is to create jobs relatively fast, cuts in personal income taxes are probably the best fiscal instrument.[18] The employment effects of cuts in corporate taxes are more delayed and less certain. The studies cited above suggest that the same is true for government spending increases.

4 .3 Private Expenditure Components

Changes in taxes are often implemented with the aim of stimulating private consumption or setting the economy on a path of higher investment and higher prosperity in the long run. Thus, it is interesting to examine how tax changes affect private sector spending and saving. For estimating the responses of private consumption, we add consumption of nondurable goods and services, durable goods purchases and personal taxable income to the baseline variables. For investment, we add nonresidential investment and residential investment as well as corporate profits.

---

[18]Monacelli, Perotti and Trigari (2011) also separately estimate the effects of business and labor taxes. When expressed in terms of multipliers, our results are entirely consistent with their finding that the effects of business taxes on employment are larger than those of labor taxes. Relative to their estimates, our results imply larger effects on unemployment which in the case of labor taxes are also more immediate.

Figure 11 shows the responses of the private consumption and investment expenditure components following a one percentage point cut in the APITR (left column) and in the ACITR (right column), respectively. In response to a cut in the APITR, nondurable and services consumption rises by 0.1 percent on impact and subsequently increases gradually to a peak response of just above 0.4 percent which occurs around 2 years after the tax cut. The consumption response appears roughly consistent with permanent income predictions for persistent changes in disposable income: it is more muted and smoother relative to the response of personal income. However the response is imprecisely estimated and not statistically significant. Durable goods purchases rise on impact by 3.6 percent and remain higher at 5 percent for two years after the tax stimulus. The positive response of nondurable purchases is significant at the 95% level for more than a year after the cut in the APITR.

The positive consumption response to an APITR cut contrasts with the response to a cut in the ACITR, which induces a decline in nondurable and services consumption that is marginally statistically significant at the 90% level on impact, but not thereafter. Durable goods purchases decline slightly but insignificantly so. Since a corporate tax cut more or less directly increases the return on saving, the consumption decline is indicative of substitution effects dominating income effects. We also looked at the response of the personal savings rate, which increases after both types of tax cuts.

The impact on private nonresidential investment is more uniform across the two tax components. A one percentage point cut in the APITR sets off a 2.1 percent increase in nonresidential investment in the quarter of the tax cut rising to a maximum of 4 percent after one year. The corresponding numbers for the ACITR are an impact increase in nonresidential investment of 0.5 percent and a peak increase of 2.3 percent after six quarters. Relative to the size of the output response, these numbers imply a stronger investment response to the ACITR than the APITR. In both cases the response of nonresidential investment is statistically significant for multiple quarters. Residential investment also responds positively to cuts in both types of taxes, although only significantly so for the ACITR.

In summary, changes in taxes impact importantly on the key spending components but again there is an important difference between personal and corporate income taxes. Changes in either type of taxes boost investment but only personal income tax cuts have short run positive effects on consumption expenditures, whereas corporate tax cuts do not affect or even lower consumption expenditures. We emphasize though that the estimates for consumption are relatively imprecise.

5 Concluding Remarks

Our analysis shows that changes in taxes have important consequences for the economy. This is important given the current debate on the efficacy of fiscal policy and on the possible consequences of the fiscal consolidation that is bound to take place over the coming years. The evidence we contribute in this paper is supportive for (i) relatively large and immediate output effects following changes in average tax rates (ii) tax multipliers that are larger than most estimates of government spending multipliers (iii) personal income tax cuts being more effective in creating jobs and stimulating consumption in the short run than cuts to corporate profit taxes and (iv) changes in corporate tax rates being approximately revenue neutral.

A key finding is that there are important differences in the effects on various macroeconomic aggregates after distinguishing between different types of taxes. Studies that focus on changes in total tax revenues alone can therefore only provide limited insight into a complex tax transmission mechanism and offer little guidance for judging the relative merits of different types of tax changes. On the other hand, the shocks to average tax rates that we identify still reflect changes to marginal tax rates as well as other tax policy instruments. The main benefit of such aggregation is that it allows for controlling for macroeconomic conditions as traditionally emphasized in the macro literature. This approach is complementary to single event studies of macro data, such as House and Shapiro (2006) or those surveyed in Chetty, Guren, Manoli and Weber (2012), that do not explicitly control for macroeconomic conditions but can incorporate much greater legislative detail.

There are several interesting avenues for future research. First, we believe that it would be interesting to apply the methodology to data from other countries. Tax narratives are becoming increasingly available, see e.g. Cloyne (2011) for a UK tax narrative and the International Monetary Fund (2010) for a tax narrative for a broad selection of countries. It is likely that measurement errors are systematic features of these accounts making our approach attractive. Secondly, it would be interesting to confront the evidence that we have uncovered with macroeconomic models in order to examine its congruence with economic theory. Another possible direction is to allow for time-varying effects of fiscal shocks, as in Auerbach and Gorodnichenko (2011). Finally, the methodology that we propose lends itself to applications to government spending and monetary policy where narrative policy measures are available. The methodology can also be used without availability of narrative measures as long as other proxies are available. Such applications could be very helpful in bringing about further evidence about the impact of structural shocks.

References

Auerbach, Alan J., and Yuriy Gorodnichenko, 2011, "Measuring the Output Responses to Fiscal Policy," forthcoming, American Economic Journal: Economic Policy.

Barro, Robert J. and Charles J. Redlick, 2011, " The Macroeconomic Effects of Government Purchases and Taxes", Quarterly Journal of Economics 126(1), 51-102.

Blanchard, Olivier and Roberto Perotti, 2002, "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output", Quarterly Journal of Economics 117(4), pp. 1329—1368.

Bollen, Kenneth A., 1989, "Structural Equations with Latent Variables", Wiley Series in Probability and Mathematical Statistics.

Burnside, Craig, Martin Eichenbaum and Jonas D.M. Fisher, 2004, "Fiscal Shocks and Their Consequences", Journal of Economic Theory 115(1), pp. 89–117.

Chetty, Raj, Adam Guren, Day Manoli and Andrea Weber, 2012, "Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities", NBER Macroeconomics Annual 2012 forthcoming .

Cloyne, James, 2011, "What Are the Effects of Tax Changes in the United Kingdom? New Evidence from a Narrative Evaluation", manuscript, University College London.

Evans, Charles L. and David A. Marshall, 2009, "Fundamental Economic Shocks and the Macroeconomy", Journal of Money, Credit and Banking 41(8), pp. 1515–1555.

Favero, Carlo and Francesco Giavazzi, 2011,“Measuring Tax Multipliers. The Narrative Method in Fiscal VARs”. American Economic Journal: Economic Policy, forthcoming.

Fisher, Jonas D.M. and Ryan Peters, 2010, “Using Stock Returns to Identify Government Spending Shocks”, Economic Journal 120(544), pp. 414–436.

Fortune, Peter, 1996, “Do Municipal Bond Yields Forecast Tax Policy?”, New England Economic Review, Federal Reserve Bank of Boston, September, pp. 29–48.

Francis, Neville and Valerie A. Ramey, 2009, “Measures of per Capita Hours and Their Implications for the Technology-Hours Debate”, Journal of Money Credit and Banking 41(6), pp. 1071–1097.

Gleser, Leon J., 1992, “The Importance of Assessing Measurement Reliability in Multivariate Regression”, Journal of the American Statistical Association, 87(419) Theory and Methods: pp. 696–707.

Gonçalves, Silvia and Lutz Kilian, 2004, “Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form”, Journal of Econometrics 123(1), pp. 89–120.

House, Christopher L. and Matthew D. Shapiro, 2006, “ Phased-In Tax Cuts and Economic Activity”, American Economic Review 96(5), pp. 1835-1849.

Hausman, Jerry A. and William E. Taylor, 1983,“Identification in Linear Simultaneous Equation Models with Covariance Restrictions: An Instrumental Variables Interpretation.”, Econometrica 51(5), pp. 1527–1549.

International Monetary Fund, 2010, “Will It Hurt? Macroeconomic Effects of Fiscal Consolidations”, Chapter 3 in: World Economic Outlook 2010.

Jones, John Bailey, 2002, “Has Fiscal Policy Helped Stabilize the Postwar U.S. Economy?”, Journal of Monetary Economics 49(4), pp. 709–746.

Leeper, Eric M., Todd B. Walker and Shu-Chun Susan Yang, 2011, “Foresight and Information Flows”, manuscript, Indiana University Bloomington.

Mendoza, Enrique G., Assaf Razin and Linda L. Tesar, 1994, “Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption”, Journal of Monetary Economics 34(3), pp. 297–323.

Mertens, Karel and Morten O. Ravn, 2010, “Measuring the Impact of Fiscal Policy in the Face of Anticipation: a Structural VAR Approach”, Economic Journal, 120(544), pp. 393–413.

Mertens, Karel and Morten O. Ravn, 2011a, “Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated U.S. Tax Policy Shocks”, American Economic Journal: Economic Policy, forthcoming.

Mertens, Karel and Morten O. Ravn, 2011b, “Understanding the Aggregate Effects of Anticipated and Unanticipated Tax Policy Shocks”, Review of Economic Dynamics, 14(1), pages 27–54.

Mertens, Karel and Morten O. Ravn, 2011c, “A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers”, manuscript, Cornell University and University College London.

Monacelli, Tommaso, Roberto Perotti and Antonella Trigari, 2010, “Unemployment Fiscal Multipliers”, Journal of Monetary Economics 57(5), pp. 531–53.

Monacelli, Tommaso, Roberto Perotti and Antonella Trigari, 2011, “Taxes and the Labor Market”, in: L. Céspedes and J. Galí, Editors, Series on Central Banking, Analysis, and Economic Performance, Central Bank of Chile.

Mountford, Andrew, and Harald Uhlig, 2009, “What Are the Effects of Fiscal Policy Shocks?” Journal of Applied Econometrics 24(6), pp. 960–992.

Nevo, Aviv, and Adam M. Rosen, 2010, “Identification with Imperfect Instruments”, Review of Economics and Statistics, forthcoming.

Perotti, Roberto, 2011, “The Effects of Tax Shocks on Output: Not So Large, But Not Small Either”, American Economic Journal: Economic Policy, forthcoming.

Poterba, James M., 1988, “Are Consumers Forward Looking? Evidence From Fiscal Experiments”, American Economic Review 78(2), pp. 413–18.

Ramey, Valerie A., 2011a, “Identifying Government Spending Shocks: It’s All in the Timing”, Quarterly Journal of Economics 126(1), pp. 10-50.

Ramey, Valerie A., 2011b, “Can Government Purchases Stimulate the Economy?”, Journal of Economic Literature, forthcoming.

Ramey, Valerie A. and Matthew D. Shapiro, 1998, “Costly Capital Reallocation and the Effects of Government Spending”, Carnegie-Rochester Conference Series on Public Policy, vol. 48(1), pp.145–194.

Ravn, Morten O., and Saverio Simonelli, 2007, “Labor Market Dynamics and Business Cycles: Structural Evidence for the United States”, Scandinavian Journal of Economics 109(4), p.743–77.

Romer, Christina D., and Jared Bernstein, 2009, “The Job Market Impact of the American Economic Recovery and Investment Plan”, manuscript, Council of Economic Advisors.

Romer, Christina D., and David H. Romer, 1989, “Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz”, in: Blanchard, O.J. and Fischer, S., Editors, NBER Macroeconomics Annual 1989, MIT Press, Cambridge, MA, pp.121–170.

Romer, Christina D., and David H. Romer, 2009a, “A Narrative Analysis of Postwar Tax Changes” manuscript, University of California, Berkeley.

Romer, Christina D., and David H. Romer, 2009b, “Do Tax Cuts Starve the Beast? The Effect of Tax Changes on Government spending”, Brookings Papers on Economic Activity 40(1), pp. 139–214.

Romer, Christina D., and David H. Romer, 2010, “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks” , American Economic Review 100(3), pp. 763–801.

Rudebusch, Glenn, 1998, “Do Measures of Monetary Policy in a VAR Make Sense?”, International Economic Review 39(4), pp. 907-931.

Yang, Shu-Chun Susan, 2005, "Quantifying Tax Effects under Policy Foresight", Journal of Monetary Economics 52(8), pp.1557–1568.

A Identification

In this appendix, we provide the identification of the impulse response functions and reliability statistics in terms of observable data moments $\Sigma_{uu'}$, $\Sigma_{mu'}$ and $\Sigma_{mm'}$. The identifying covariance restrictions are $\Sigma_{uu'} = \mathcal{B}\mathcal{B}'$ and (8). These restrictions yield the following closed form solutions

$$\begin{aligned}
\beta_{11}S_1^{-1} &= \left(I-\beta_{12}\beta_{22}^{-1}\beta_{21}\beta_{11}^{-1}\right)^{-1} && (19)\\
\beta_{21}S_1^{-1} &= \beta_{21}\beta_{11}^{-1}\left(I-\beta_{12}\beta_{22}^{-1}\beta_{21}\beta_{11}^{-1}\right)^{-1} && (20)\\
S_1S_1' &= \left(I-\beta_{12}\beta_{22}^{-1}\beta_{21}\beta_{11}^{-1}\right)\beta_{11}\beta_{11}'\left(I-\beta_{12}\beta_{22}^{-1}\beta_{21}\beta_{11}^{-1}\right)' && (21)
\end{aligned}$$

where

$$\begin{aligned}
\beta_{21}\beta_{11}^{-1} &= (\Sigma_{mu_1'}^{-1}\Sigma_{mu_2'})'\\
\beta_{12}\beta_{22}^{-1} &= \left(\beta_{12}\beta_{12}'(\beta_{21}\beta_{11}^{-1})' + \left(\Sigma_{21}-\beta_{21}\beta_{11}^{-1}\Sigma_{11}\right)'\right)(\beta_{22}\beta_{22}')^{-1}\\
\beta_{12}\beta_{12}' &= \left(\Sigma_{21}-\beta_{21}\beta_{11}^{-1}\Sigma_{11}\right)'\mathcal{Z}^{-1}\left(\Sigma_{21}-\beta_{21}\beta_{11}^{-1}\Sigma_{11}\right)\\
\beta_{22}\beta_{22}' &= \Sigma_{22}+\beta_{21}\beta_{11}^{-1}\left(\beta_{12}\beta_{12}'-\Sigma_{11}\right)(\beta_{21}\beta_{11}^{-1})'\\
\beta_{11}\beta_{11}' &= \Sigma_{11}-\beta_{12}\beta_{12}'\\
\mathcal{Z} &= \beta_{21}\beta_{11}^{-1}\Sigma_{11}(\beta_{21}\beta_{11}^{-1})' - \left(\Sigma_{21}(\beta_{21}\beta_{11}^{-1})' + \beta_{21}\beta_{11}^{-1}\Sigma_{21}'\right) + \Sigma_{22}
\end{aligned}$$

where the Σ_{ij}'s denote the elements of the appropriate partitioning of $\Sigma_{uu'}$. When a single proxy is used, i.e. $k = 1$, the first column of $\mathcal{B}$ is determined (up to a signing convention) since S_1S_1' is a scalar. With multiple proxies $k > 1$, the identification of the structural impulse responses is completed by a Choleski decomposition of S_1S_1'.

Under the additional restrictions of the measurement error model, the reliability matrix is identified by

$$\Lambda = \frac{1}{d}\Sigma_{mm}^{-1}\Sigma_{mu_1'}(\beta_{11}\beta_{11}')^{-1}\Sigma_{mu_1'}' \tag{22}$$

where d is the fraction of uncensored observations of m_t. For the univariate case ($k = 1$), β_{11} and the shocks ε_{1t} are identified. The scalar reliability of m_t can in that case also be estimated in a sample of length T by,

$$\Lambda = \left(\Gamma^2\sum_{t=1}^{T} D_t\varepsilon_{1t}^2 + \sum_{t=1}^{T} D_t(m_t - \Gamma\varepsilon_{1t})^2\right)^{-1}\Gamma^2\sum_{t=1}^{T} D_t\varepsilon_{1t}^2\ , \tag{23}$$

where $\Gamma = \left(\sum_{t=1}^{T} D_t m_t u_{1t} / \sum_{t=1}^{T} D_t\right)/\beta_{11}$. The advantage of (23) over (22) is that it always lies in the unit interval. We therefore prefer this estimator when $k = 1$.

B Data Definitions and Sources

Benchmark Variables

Population is the total population over age 16 from Francis and Ramey (2009) (*nipop16*); **Output** is Real GDP in line 1 from NIPA Table 1.1.3 divided by population; **Government spending** is Real Federal Government Consumption Expenditures and Gross Investment in line 22 from NIPA Table 1.1.3 divided by population; The **personal income tax base** is NIPA personal income (Table 2.1 line 1) less government transfers (Table 2.1 line 17) plus contributions for government social insurance (Table 3.2 line 11); The **corporate income tax base** is NIPA corporate profits (Table 1.12 line 13) less Federal Reserve Bank Profits (Tables 6.16 B-C-D). The tax bases are all deflated by the GDP deflator in line 1 from Table 1.1.9 and by population; The **average personal income tax rate** is the sum of federal personal current taxes (Table 3.2 line 3) and contributions for government social insurance divided by the personal income tax base; The **average corporate income tax rate** is federal taxes on corporate income excluding Federal Reserve banks (Table 3.2 line 9) divided by corporate profits (excluding Fed profits). **Debt** is Federal Debt Held by the Public from Favero and Giavazzi (2011) (*DEBTHP*), divided by the GDP deflator and population. All NIPA tables were downloaded 1/23/2012.

Other Variables

Employment/Population is total economy employment from Francis and Ramey (2009), divided by population; The **Labor Force/Population** is the sum of employment and the number of unemployed (FRED, series UNEMPLOY) divided by population; **Hours per worker** is the total economy hours worked series from Francis and Ramey (2009), which includes the government and non-profits sectors, divided by employment. The **price level** is the implicit deflator for Personal Consumption Expenditures (NIPA Table 1.1.9 line 2) and inflation is the annualized quarterly percentage change in the price level; The **federal funds rate** is the effective federal funds rate series from Romer and Romer (2010) which they extended back to 1950Q1 ; **Nonborrowed Reserves** is from FRED (series BOGNONBR), extended back to 1950Q1 by subtracting borrowed reserves (FRED: BORROW) from total reserve balances (FRED: RESBALNS) after adjusting for changes in reserve requirements using the reserve adjustment magnitude from the St. Louis Fed. **Consumption of Nondurable Goods And Services** is the aggregated chained nondurable and services consumption obtained using data from NIPA Tables 1.1.5 and 1.1.9, divided by the population ; **Durable Goods Purchases**, **Nonresidential and Residential investment** are from NIPA Table 1.1.3 (lines 4, 9 and 12) and were divided by the population. All NIPA and FRED tables were downloaded 1/23/2012.

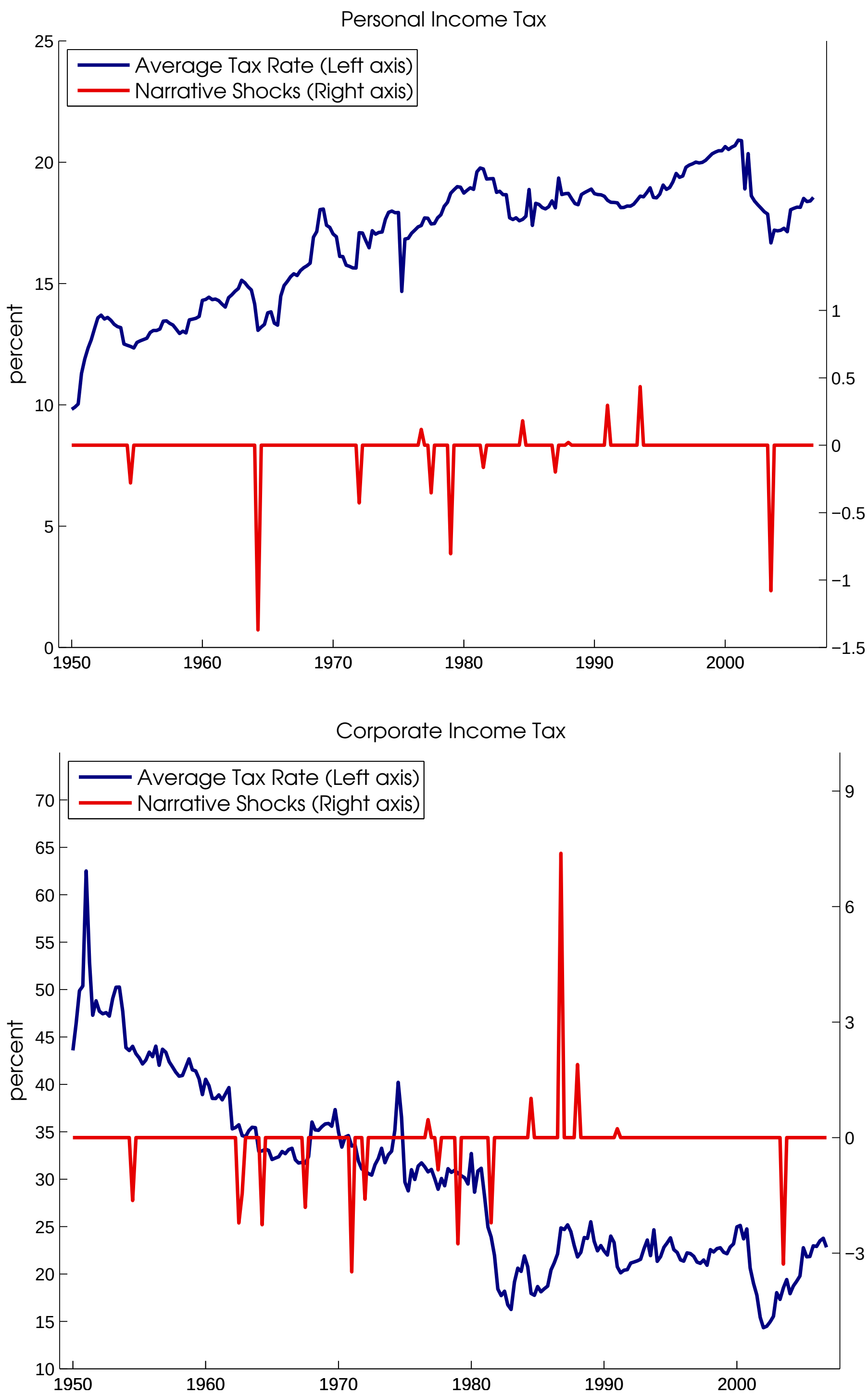

Figure 1 Average Tax Rates and Narrative Shock Measures for the US 1950Q1-2006Q4

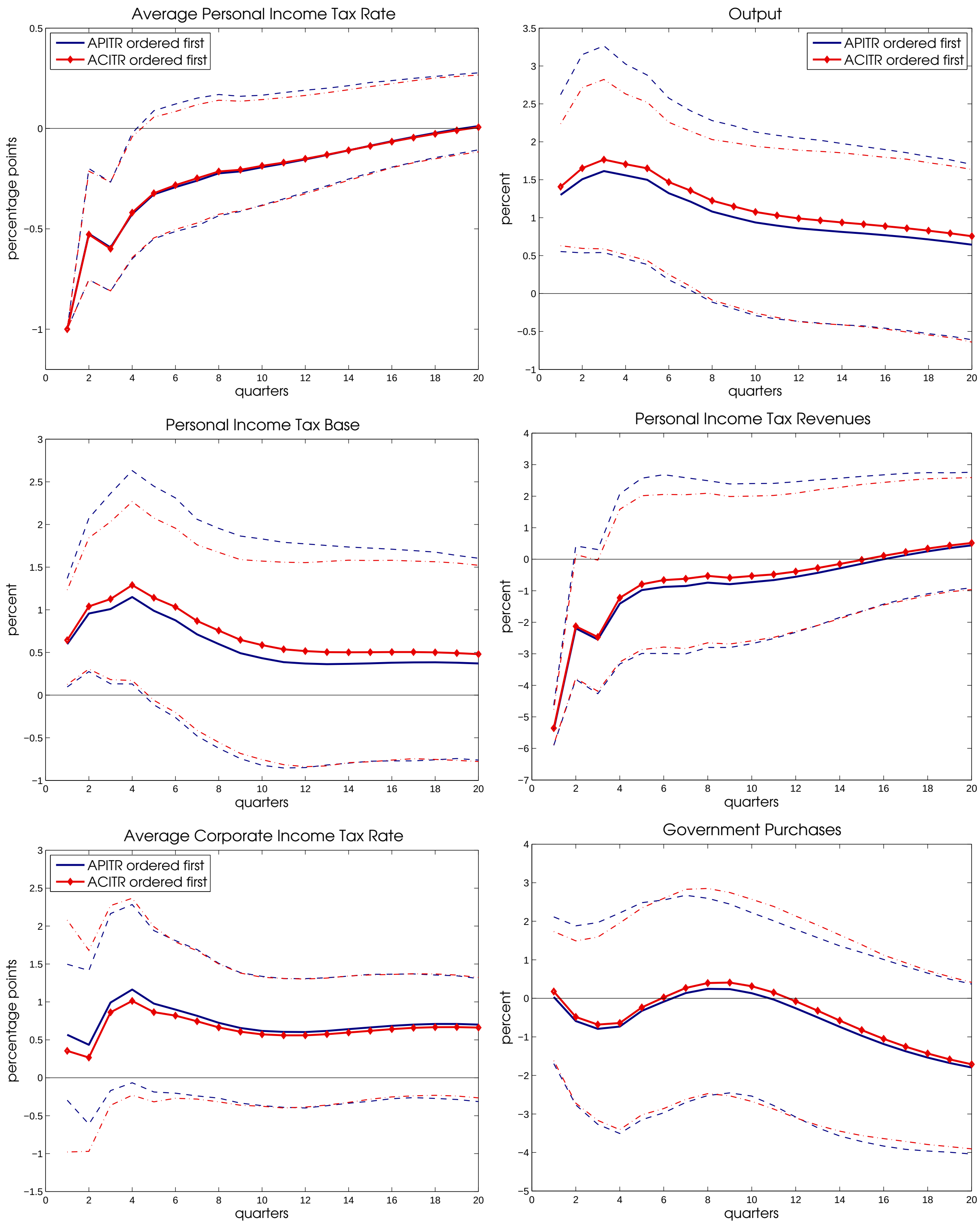

Figure 2 Benchmark Specification: Response to One Percentage Point Cut In **Average Personal Income Tax Rate**. Broken lines are 95% percentile intervals.

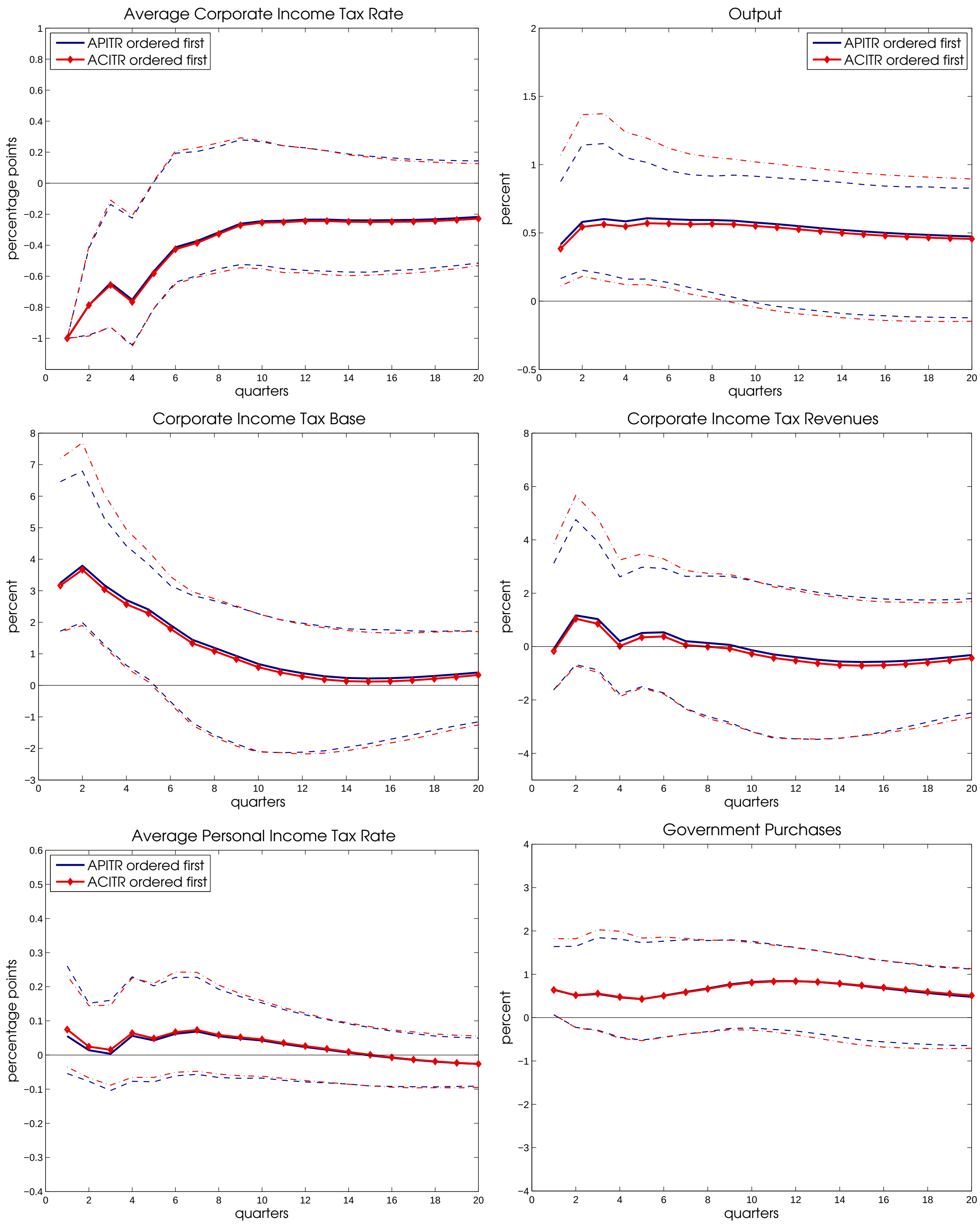

Figure 3 Benchmark Specification: Response to One Percentage Point Cut In **Average Corporate Income Tax Rate**. Broken lines are 95% percentile intervals.

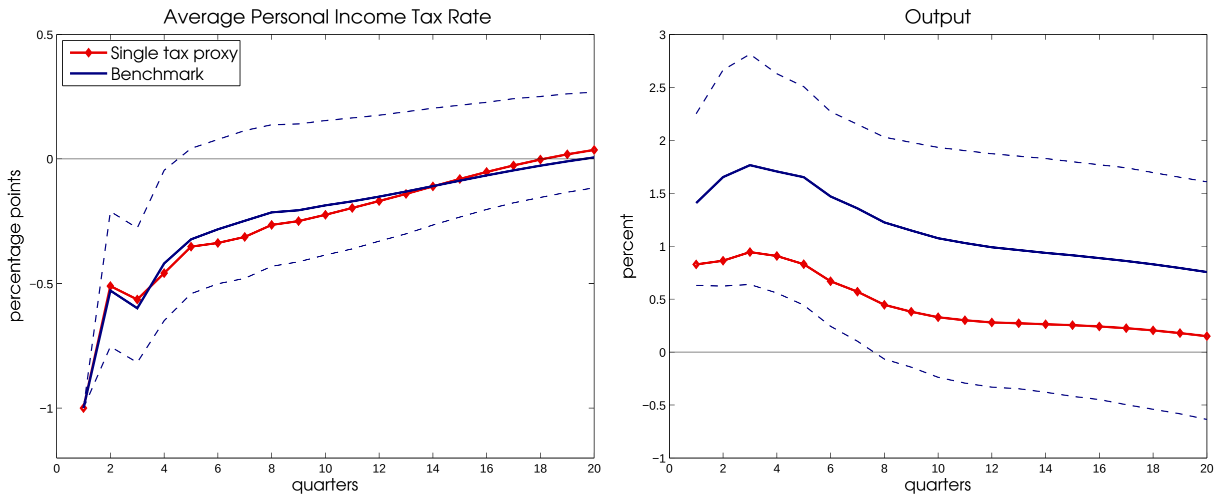

Figure 4 One Percentage Point Cut In **Average Personal Income Tax Rate**. Broken lines are 95% percentile intervals.

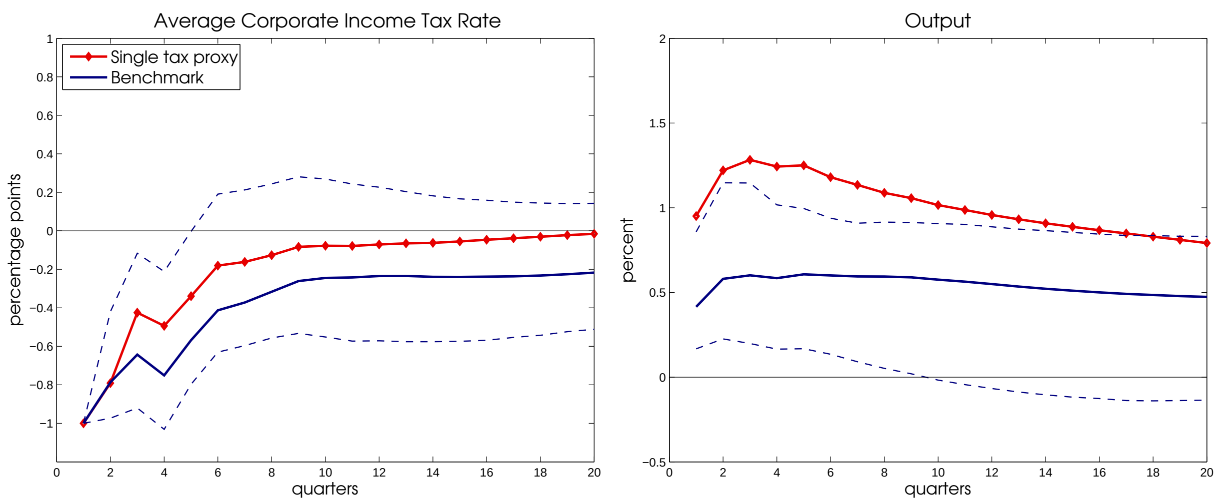

Figure 5 One Percentage Point Cut In **Average Corporate Income Tax Rate**. Broken lines are 95% percentile intervals.

(A) Personal Income Tax Cut

(B) Corporate Income Tax Cut

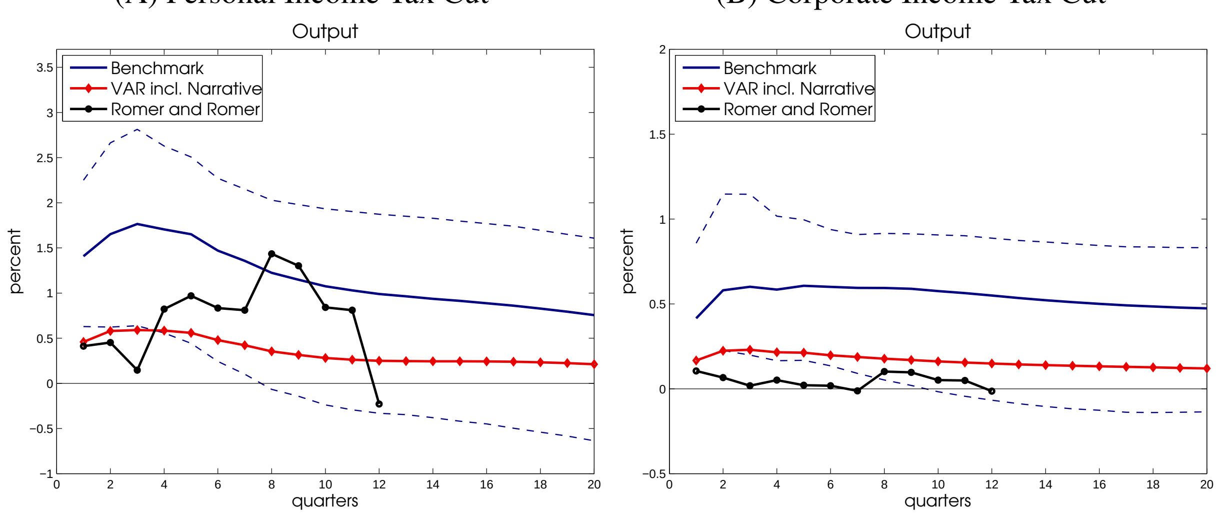

Figure 6 Comparing to **Alternative Empirical Specifications**

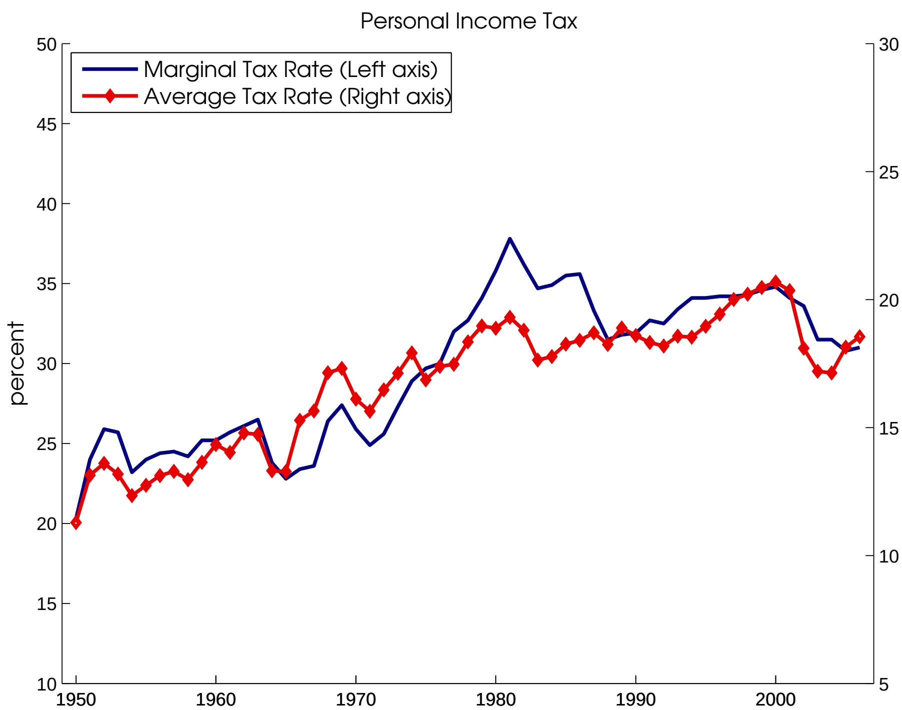

Figure 7 Annual Data for Average and Marginal Rates

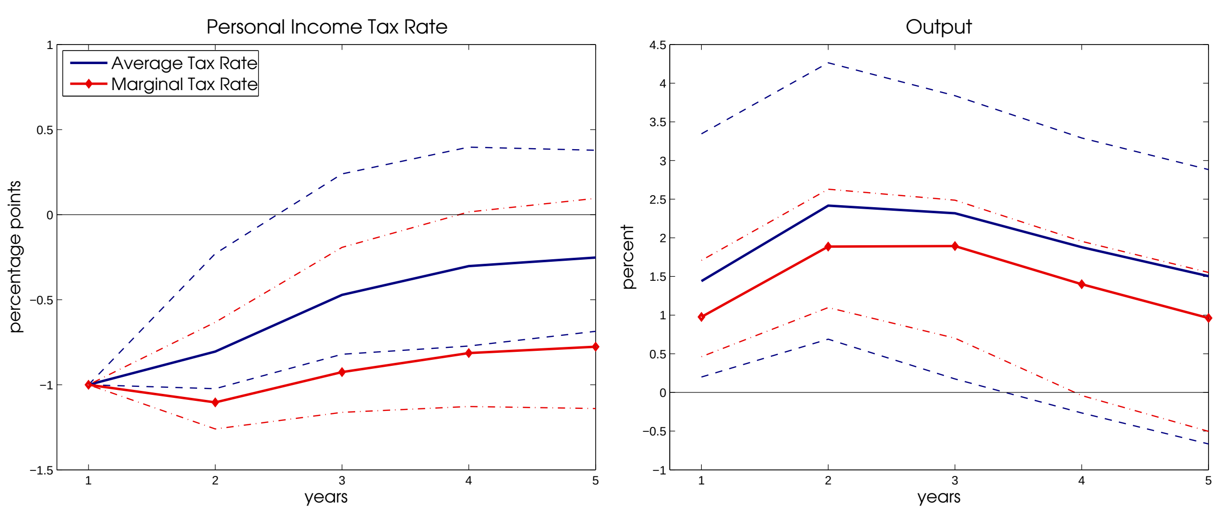

Figure 8 Annual VAR: Response to One Percentage Point Cut In **Marginal or Average Personal Income Tax Rate**. Broken lines are 95% percentile intervals.

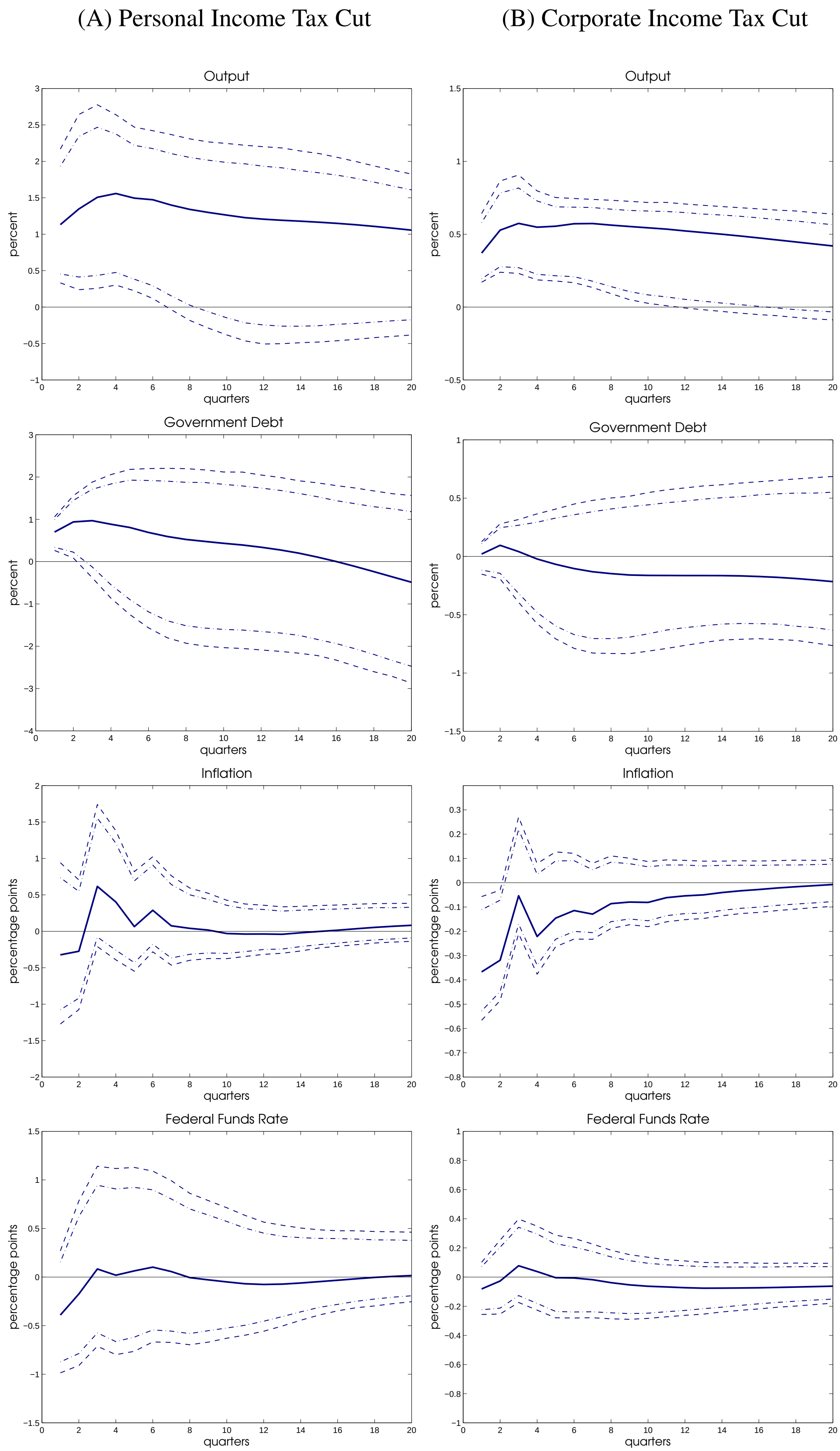

Figure 9 Monetary Policy and Inflation: Response to One Percentage Point Cut In **Average Tax Rate**. Broken lines are 90% and 95% percentile intervals.

(A) Personal Income Tax Cut

(B) Corporate Income Tax Cut

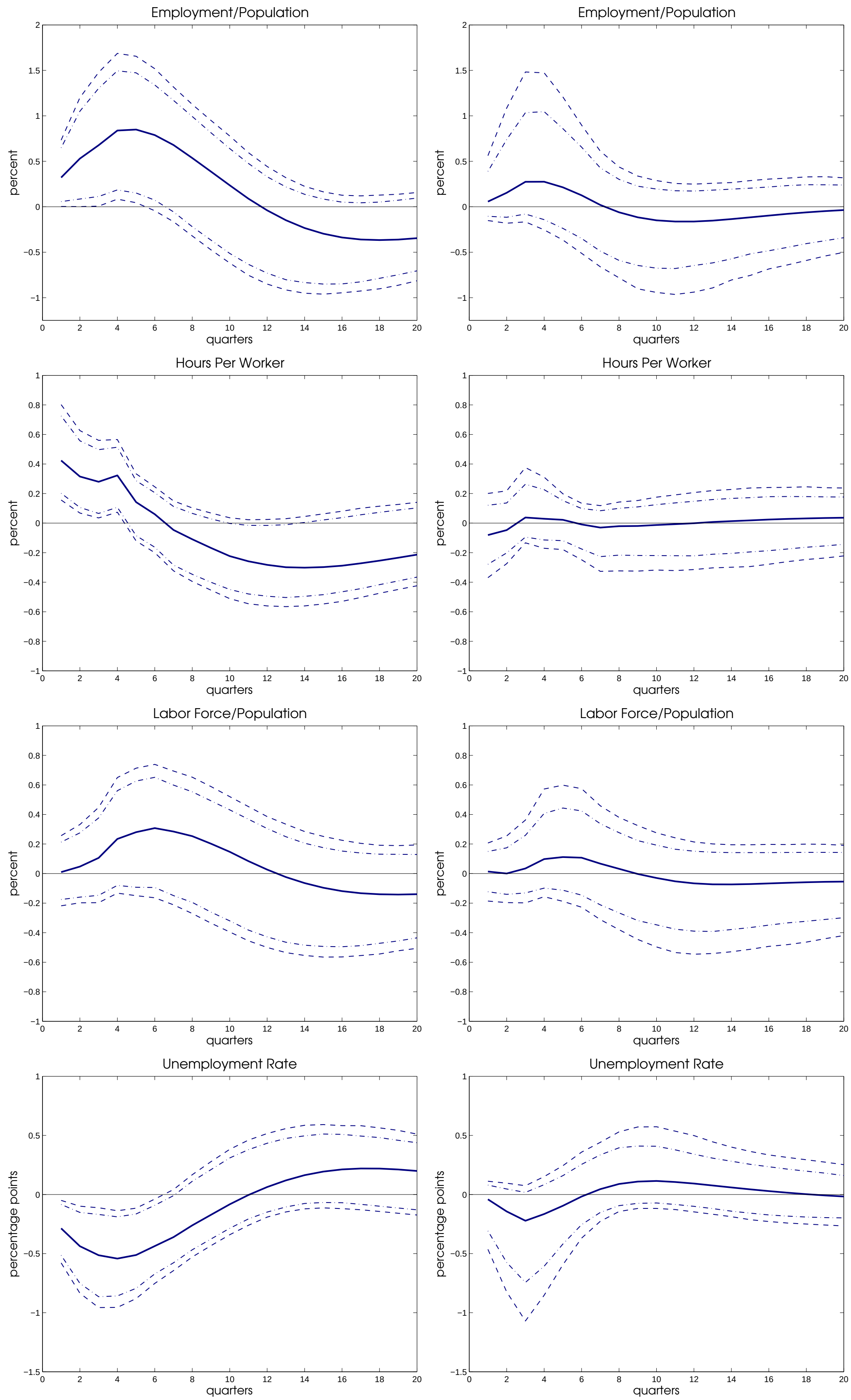

Figure 10 Labor Market: Response to One Percentage Point Cut In **Average Tax Rate**. Broken lines are 90% and 95% percentile intervals.

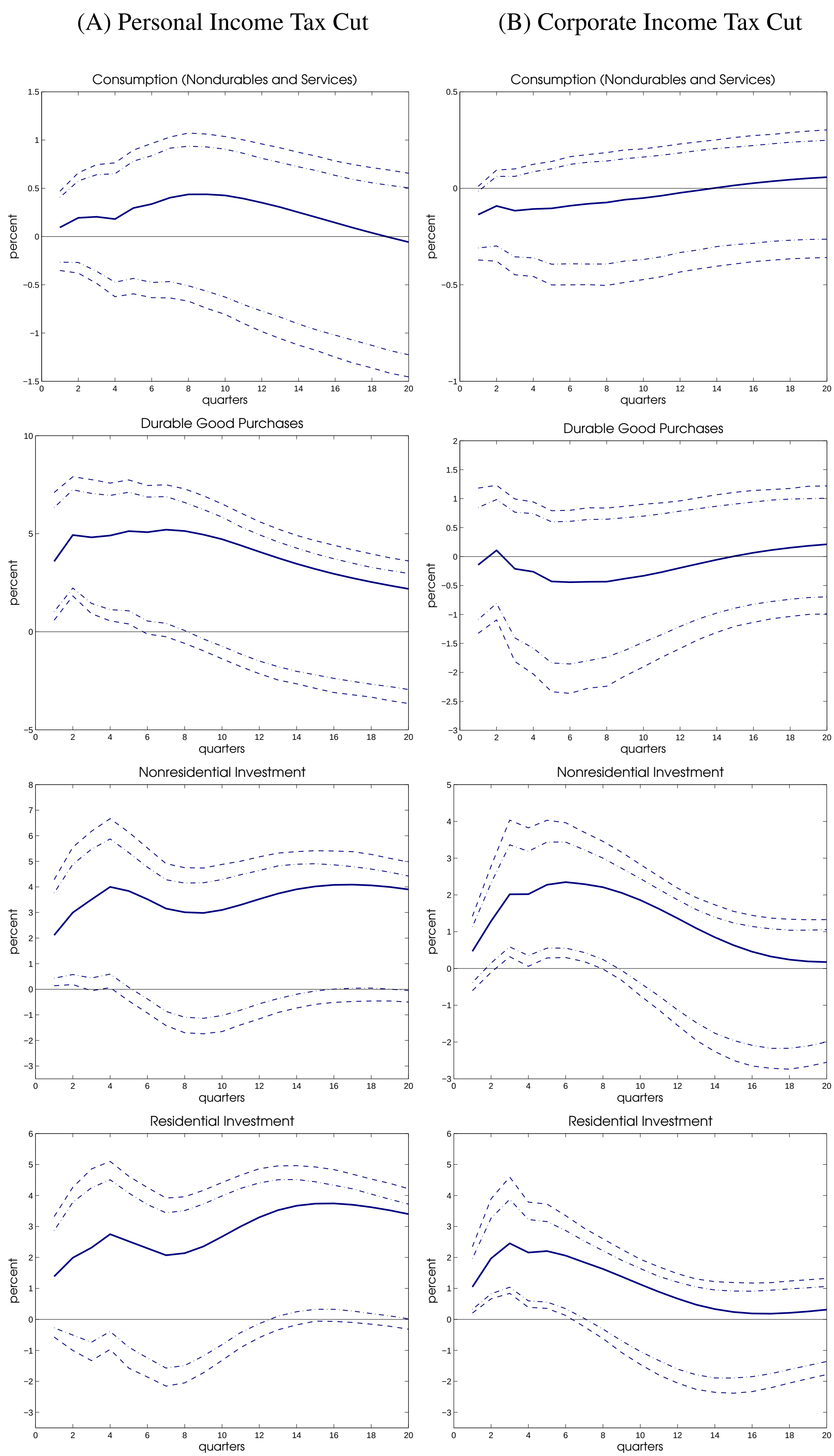

Figure 11 Major Private Expenditure Components: Response to One Percentage Point Cut In **Average Tax Rate**. Broken lines are 90% and 95% percentile intervals.