Discussion Paper Series

The Macroeconomic Effects of Income and Consumption Tax Changes

By

AnhD.M.Nguyen, Luisanna Onnis, Raffaele Rossi

Centre for Growth and Business Cycle Research, Economic Studies,
University of Manchester, Manchester, M13 9PL, UK

December 2016
Number 227

The Macroeconomic Effects of Income and Consumption Tax Changes*

AnhD.M.Nguyen† Luisanna Onnis‡ Raffaele Rossi§

December 2016

Abstract

Do income and consumption tax changes affect the economy differently? We answer this question by estimating structural VARs, where we proxy the latent tax shocks with a newly constructed narrative account of income and consumption tax liability changes in the United Kingdom. We find that income tax shocks have large short run effects on GDP, private consumption and investment. The effects of consumption tax cuts are modest and not statistically different from zero on GDP and investment and only marginally expansionary on private consumption. These results indicate that i) it is crucial to distinguish between direct and indirect taxation when studying the transmission mechanism of fiscal policy, and ii) consistent with conventional public finance theories, consumption taxes are less distortive than income taxes.

JEL codes: E62, H24, H25, H31.
Keywords: fiscal policy, narrative account, consumption taxation, income taxation, Proxy-SVAR.

*We thank James Banks, Paul Beaudry, James Cloyne, Sarolta Laczó, Roberto Pancrazi, Ivan Petrella, Silvana Tenreyro, Akos Valentinyi, Eugenia Vella, Mazhar Waseem, Matthias Weber and the participants at various seminars and conferences for useful comments. The views expressed in this paper are not those of the Bank of Lithuania or the Eurosystem. The latest version of this paper can always be found at https://sites.google.com/site/raffaelerossiwebpage/Home/research.

†Applied Macroeconomic Research Division, Economics Department, Bank of Lithuania, Totoriu str. 4, LT 01121 Vilnius, Lithuania. Email: anguyen@lb.lt

‡Department of Economics, University of Sheffield, S1 4DT, 9 Mappin Street, Sheffield, United Kingdom. Email: l.onnis@sheffield.ac.uk

§Department of Economics, University of Manchester, M13 9PL, United Kingdom. Email: raffaele.rossi@manchester.ac.uk.

1 Introduction

Do consumption and income tax changes affect economic aggregates such as income, private consumption and investment differently? The answer to this age-old question is central for the setting of economic policy and its role in pursuing economic stabilisation, efficiency and growth. Conventional public finance theories argue that consumption taxes are less distortive than income taxes.[1] From this, two main theoretical results emerge: i) income tax multipliers are larger than consumption tax multipliers (Fernández-Villaverde, 2010; and Sims and Wolff, 2016 amongst others) and ii) shifting the burden of taxation from income to consumption is expansionary (Coleman, 2000; Altig, Auerbach, Koltikoff, Smetters, and Walliser, 2001; and Correia, 2010 amongst others). However, despite considerable research in the theoretical literature, little is known from an empirical perspective about the joint macroeconomic effects of changes in income taxes as opposed to changes in consumption taxes. Our main contribution consists in providing new estimates of these effects. We identify exogenous tax changes in a VAR model by proxying the latent tax shocks with a novel narrative measure of consumption and income tax liabilities changes for the United Kingdom, over the period 1973-2009.

We find that unexpected changes in average income tax (AIT), defined as an aggregate of personal and corporate taxes, have large short run effects on GDP, private consumption and investment. A percentage point cut in AIT raises GDP on impact between 0.6 and 0.9 percent and leads to maximal present value multipliers ranging from 2.2 to 2.7 two years after the shock. We also show that AIT cuts 'Starve the Beast', i.e. government spending shrinks in response to income tax decreases. Differently, the effects of cuts in average consumption tax (ACT), e.g. VAT and consumption duties, are

[1] See Summers (1981), Auerbach, Kotlikoff, and Skinner (1983), Jones, Manuelli, and Rossi (1993) and Trabandt and Uhlig (2011) amongst others. For instance, a key difference between consumption and income-based taxes is in their treatment of capital income, which is taxed under the income tax but exempt under a consumption tax. This difference emerges because the purchases of investment goods are not included in a consumption tax, either due to exemptions or credits under the VAT schemes. Such treatment is more generous than that offered under an income tax, which allows only deductions for economic depreciation.

modest and not significantly different than zero on GDP, investment, and government spending and only marginally expansionary on private consumption. Furthermore, we find that revenue neutral policies that increase consumption taxes and decrease income taxes, have a short but significant expansionary effect on GDP. To this end, our estimates are consistent with the results emerging from the theoretical literature on the effects of consumption and income tax changes.

On the policy side, our findings support the fiscal reforms and the ongoing proposals aimed to shift the burden of taxation from income to consumption. During the period 1979-1996, the UK government legislated several fiscal policy changes where cuts of income based taxation were coupled with increases in consumption taxation. In 2011, China started the implementation of its biggest tax reform since 1994, aimed to replace the income based Business Tax with a Value Added Tax (see Cui, 2014). In the US, there have been recent proposals to replace part of Federal Business Tax with a broad based Federal Consumption Tax.[2] Overall, most OECD countries have shifted their tax burden from income to consumption taxation during the last four decades (see OECD, 2014).

Measuring the effects of a tax change is particularly challenging for (at least) two reasons. First of all, there is a well known problem of endogeneity, as fiscal policy is rarely the result of random experimentation. Tax changes are likely to contemporaneously affect various expenditure components of GDP. At the same time, tax changes are also contemporaneously driven by GDP and its components. Second, distinct tax instruments are likely to affect the economy through separate channels, and hence lead to important differences in the transmission mechanism of fiscal policy. Partly due to these difficulties, existing empirical estimates of tax multipliers vary greatly from almost insignificant to very large.

Most of the literature deals with the problem of endogeneity by either constructing

[2]For example, in January 2005, the US President George W. Bush commissioned a panel to propose a comprehensive reform to the tax code (President's Advisory Panel on Federal Tax Reform, 2005, available at http://govinfo.library.unt.edu/taxreformpanel/index-2.html). The panel considered replacing the entire income-tax system with a national sales tax. However, the complexity of replacing the current tax regime with a broad-based consumption tax and the disagreement among panel members regarding the effects of such a reform inhibited its adoption.

narrative measures of exogenous policy changes, e.g. Romer and Romer (2010) and Cloyne (2013), or by imposing on a structural vector autoregression model (SVAR) a number of identifying restrictions, e.g. Blanchard and Perotti (2002) and Mountford and Uhlig (2009). The former method has the advantage of presenting rich information set behind the narrative accounts. However, it may suffer from measurement errors and arbitrary judgements about the precise nature of policy decisions. The latter method, on the other hand, has the advantage of giving parsimonious identification of the shock transmission mechanism but at the cost of parameter restrictions that are often questionable.

We adopt a third, more recent econometric technique that takes advantage of the desirable features of both the narrative accounting and the structural VARs. The idea behind this method is to use the narratively identified exogenous policy changes as instruments or 'proxies' for the true policy shocks in a structural VAR fashion, i.e. 'Proxy-SVAR', see Stock and Watson (2012), Mertens and Ravn (2013, 2014) and Caldara and Kamps (2017). This identification strategy relies on assuming that narrative measures of exogenous policy changes correlate with latent tax shocks but are orthogonal to other structural shocks. Hence, this approach uses the information provided by narrative datasets but it is robust to potential measurement errors as it does not require perfect correlation between the narrative measures and the latent structural tax shocks. At the same time, it avoids any identifying assumption or prior for the key structural elasticities. In turn, these elasticities are directly estimated within the econometric model. Given the scarcity of prior empirical studies on the dynamic effects of consumption tax changes as opposed to income tax changes, these features of the 'Proxy-SVAR' make it particularly appealing for our research question.

We present a novel narrative measure of consumption and income tax changes in the UK. Our dataset is constructed by using Cloyne (2012, 2013)'s account of the changes in total tax liabilities which we disaggregate into changes in consumption (VAT and other consumption duties) and income (personal and corporate) tax liabilities. We only use tax changes that we classify as 'exogenous', i.e. decisions that were taken for reasons uncorrelated with macroeconomic conditions. Then we show that our narrative dataset contains valuable information about the latent structural tax shocks and is exogenous to the remaining structural shocks. Finally, we use these narrative measures of exogenous AIT and ACT changes as proxies for the latent structural tax shocks in a SVAR model.

The literature has mainly studied the effects of aggregate tax shocks, e.g. Blanchard and Perotti (2002), Mountford and Uhlig (2009), Romer and Romer (2010), Favero and Giavazzi (2012) and Caldara and Kamps (2017). Nevertheless, as we show here, distinct tax components are likely to affect the economy through separate channels, and hence lead to important differences in the transmission mechanism of fiscal policy. These insights are generally missed in analyses with aggregate tax shocks. Mertens and Ravn (2013) analyse separately the impact of personal vs. corporate tax changes in the US economy. They find that both tax components have large short-run effects on GDP and investment, while they document a different transmission mechanism on private consumption and the labour market. However Mertens and Ravn (2013) focus on two homogeneous tax categories of income, i.e. both personal and corporate taxes are part of direct taxation. Differently, we concentrate on the comparison between direct (AIT) and indirect (ACT) taxation. To this end, the UK represents the ideal country of study, as consumption taxation represents, on average, around 30 percent of total tax revenues, while income tax accounts for around the 50 percent of total tax revenues.

Regarding income taxes, Barro and Redlick (2011) measure the effects of changes in the average marginal income tax (rather than, as here, on the average income tax rate) for the US and find tax multipliers slightly larger than one. Cloyne and Surico (2016) analyse the impact of exogenous tax changes in labour income in the UK, using narrative changes in this tax component and micro-data on British households. They find important heterogeneous effects when households are grouped by housing tenure, with the mortgagors exhibiting the largest and most significant response to fiscal policy shocks. Mertens (2015) estimates the dynamic effects of exogenous marginal tax rate changes on income in the US and finds large effects in the top 1 percent of the income distribution as well as in other income percentile brackets and for macroeconomic aggregates such as GDP. Zidar (2015) finds that the positive relationship between tax cuts and employment growth is largely driven by tax cuts for lower-income groups and that the effect of tax cuts for the top 10 percent on employment growth is small. While our findings on the expansionary effects of income tax changes are consistent with the mentioned literature, we do not analyse distributional issues of tax policy. Nevertheless, our novel dataset can be easily applied to study these topics. We believe this is an interesting avenue for future research.[3]

Most of the empirical literature on consumption taxation studies specific episodes of tax changes, such as the 1997 VAT hike in Japan (Cashin and Unayama, 2016) or the 2008 VAT cut in the UK (Blundell, 2009), and therefore it cannot be directly compared with our paper. A notable exception can be found in Riera-Crichton, Vegh, and Vuletin (2016), who adopt a narrative approach to estimate the effects of VAT changes for fiscal consolidation purposes in a panel of 14 countries and find that shocks in this tax component have large and significant expansionary effects on GDP. There are at least two reasons that can explain the difference between our and their results. First of all, our dataset includes many other tax changes, not just deficit measures. Secondly, and perhaps more importantly, Riera-Crichton, Vegh, and Vuletin (2016) analysis comprises only of exogenous VAT changes, thus they do not control for potential contemporaneous exogenous changes in other tax components. This is mostly due to data limitation. We show that we can obtain large ACT multipliers when we ignore the observed correlation between exogenous movements in consumption and income taxes. This finding is not surprising as changes in consumption and income taxes happens in the same quarter or in the same fiscal year. Therefore we conclude that controlling for simultaneous changes in income taxes is crucial for a correct evaluation of the transmission mechanism of consumption taxation.

The rest of this paper is organised as follows. Section 2 describes the econometric identification of our 'Proxy-SVAR'. Section 3 presents our narrative measure of income

[3]There is rich theoretical literature on the redistributive effects of consumption and income taxes, e.g. Krusell, Quadrini, and Ríos-Rull (1996), Nishiyama and Smetters (2005) and Correia (2010).

and consumption tax policy changes. Section 4 describes our results. Finally Section 5 concludes.

2 The Econometric Model

Proposed by Stock and Watson (2012) and Mertens and Ravn (2013, 2014), our econometric methodology consists of using the narrative measure of tax changes as 'instrument' or 'proxy' to identify the latent structural policy shocks in the VAR model, i.e. 'Proxy SVAR'. In what follows, we briefly describe this methodology. We postulate that our n 'observables' X_t follow the dynamic vector autoregressive model of order P,

$$X_t = \sum_{i=1}^{P} \beta_i X_{t-i} + \mathbf{\Gamma}\epsilon_t, \text{ with } i = 1, 2, ..., P, \tag{1}$$

where β_i are the $n \times n$ matrix of coefficients and $\mathbf{\Gamma}$ is a $n \times n$ non-singular matrix of coefficients. The term ϵ_t represents the $n \times 1$ vector of structural shocks which are assumed to satisfy the following conditions: i) $E(\epsilon_t) = 0$; ii) $E(\epsilon_t \epsilon_t') = I_n$, with I_n being the identity matrix of size n; and iii) $E(\epsilon_j \epsilon_k) = 0 \ \forall j \neq k$. Deterministic terms and exogenous regressors are omitted from (1) for notational brevity.

As with other SVAR models, the estimation of structural shocks ϵ_t requires the imposition of identifying assumptions on $\mathbf{\Gamma}$. To illustrate this point, let μ_t denote the reduced-form residuals from (1). These residuals are related to the structural shocks by

$$\mu_t = \mathbf{\Gamma}\epsilon_t. \tag{2}$$

One can estimate $E(\mu_t \mu_t')$, the covariance matrix of μ_t, via, for instance, the Ordinary Least Square method. Given that, by construction, $\epsilon_t = \mathbf{\Gamma}^{-1}\mu_t$, we can obtain $\frac{n(n+1)}{2}$ identifying restrictions from $E(\mu_t \mu_t') = \Sigma_{\mu\mu'} = \mathbf{\Gamma}\mathbf{\Gamma}'$.[4] However we need more restrictions in order to fully identify the elements of at least one of the columns of $\mathbf{\Gamma}$. Popular methods adopted to fulfil this objective include the policy reaction lags and institu-

[4]The number of restrictions follows from the fact that any covariance matrix is symmetric about the diagonal.

tional features as in Blanchard and Perotti (2002), or the theory-based sign restrictions as in Mountford and Uhlig (2009).

Differently, we use the information from the narrative dataset to obtain co-variance restrictions for the structural shocks. First, we partition the vector of structural shocks as $\epsilon_t = [\epsilon_{1,t}\ \epsilon_{2,t}]'$, where $\epsilon_{1,t}$ is a $m \times 1$ vector of the structural shocks of study, while $\epsilon_{2,t}$ of size $(n-m) \times 1$, represents all the other structural shocks. Second, let q_t be a $m \times 1$ vector of proxies which is assumed to have zero mean, without any loss of generality. More importantly, in order to use it for the identification of $\mathbf{\Gamma}$, q_t needs to satisfy the following two conditions:

$$E(q_t \epsilon_{1,t}) = \mathbf{\Psi}, \tag{3}$$

$$E(q_t \epsilon_{2,t}) = 0. \tag{4}$$

These conditions imply that the 'proxies' are correlated with m structural shocks of interest by a $m \times m$ non-singular matrix of unknown coefficients $\mathbf{\Psi}$, but they are uncorrelated with the other $(n-m)$ structural shocks. As such, conditions (3) and (4) can be viewed as instrument validity conditions. When these conditions hold, the 'proxy' variables can be used to identify the latent structural shocks. Then, we can partition the matrix $\mathbf{\Gamma}$ as

$$\mathbf{\Gamma} = \left[\underbrace{\gamma_1}_{n\times m} \quad \underbrace{\gamma_2}_{n\times(n-m)} \right], \quad \gamma_1 = \left[\underbrace{\gamma_{11}'}_{m\times m} \quad \underbrace{\gamma_{21}'}_{m\times(n-m)} \right]', \quad \gamma_2 = \left[\underbrace{\gamma_{12}'}_{(n-m)\times m} \quad \underbrace{\gamma_{22}'}_{(n-m)\times(n-m)} \right]',$$

with the necessary conditions that γ_{11} and γ_{22} are non-singular. From this, it can be shown that

$$\mathbf{\Psi}\gamma_1' = \Sigma_{q\mu'}. \tag{5}$$

This last condition provides extra restrictions. However, (5) is based on the $m \times m$ unknown elements of $\mathbf{\Psi}$. Without any further assumption on $\mathbf{\Psi}$, (5) provides only $(n-m)m$ identification restrictions. Using the partition $\Sigma_{q\mu'} = [\underbrace{\Sigma_{q\mu_1'}}_{m\times m} \quad \underbrace{\Sigma_{q\mu_2'}}_{m\times(n-m)}]$ and (5), we can write

$$\gamma_{11} = \left(\Sigma_{q\mu_1'}^{-1}\Sigma_{q\mu_2'}\right)' \gamma_{21}. \tag{6}$$

Because $\Sigma_{q\mu_1'}^{-1}\Sigma_{q\mu_2'}$ can be estimated, say, with the two-stage least squares method by regressing from $\mu_{2,t}$ on $\mu_{1,t}$, using q_t as instruments for $\mu_{1,t}$, (6) imposes extra covariance restrictions. Altogether, the 'Proxy-SVAR' is estimated using the following steps:

1. Estimate the reduced-form VAR model by OLS.

2. Estimate $\Sigma_{q\mu_1'}^{-1}\Sigma_{q\mu_2'}$ by regressing the residuals of the reduced-form VAR on the proxies q_t.

3. Use (6) to estimate the objects of study.

4. Impose further restrictions when $m > 1$ and the instruments for structural shocks $\epsilon_{1,t}$ are correlated with each other.

There are at least two advantages in adopting this econometric specification. First of all, it requires the availability of proxies that comply with (3) and (4) and does not rely on specific assumptions about the elements of $\mathbf{\Gamma}$, as typically imposed in fiscal-SVAR applications, e.g. Blanchard and Perotti (2002). Second, it only needs non-singularity of $\mathbf{\Psi}$. Thus it does not necessitate that narrative changes correlate perfectly with latent policy shocks, as commonly assumed in the narrative literature, e.g. Romer and Romer (2010). This last property implies that the 'Proxy-SVAR' is robust to several measurement errors that naturally arise in narrative dataset.

3 A Narrative Account of Income and Consumption Tax Changes

This paper presents a narrative measure of legislated consumption and income tax liability changes in the United Kingdom at quarterly frequency, covering the sample 1973:III-2009:IV. Our measure of tax changes complements Cloyne (2012, 2013)'s dataset by decomposing the total tax liabilities changes into the following subcomponents: i) Value-added tax liabilities (VAT), ii) consumption-related duties tax liabilities (Duties), iii) personal income tax liabilities (PI), iv) corporate tax liabilities (CI), v) national insurance tax liabilities (NIC), vi) petroleum income tax (PetIT) and vii) a residual category with other revenue changing tax measures (OT).[5] We aggregate i) and ii) as consumption taxes, while we consider the total of iii), iv), v) and vi) as income taxes. Precise categorisation and grouping are reported in Appendix A (Table A.1).[6,7,8]

We then classify by motivation each tax change of the subcomponents and retain only the 'exogenous' tax changes, i.e. those uncorrelated with macroeconomic conditions and public spending. This is a fundamental step in order to comply with the identification conditions in (3) and (4). For this task, we use the classification reported in Cloyne (2013) and its companion paper Cloyne (2012), supplemented with additional information from the Financial Statements and Budget Reports (FSBR), the Pre-Budget documents (from 1997 onwards), the Budget speech by the Chancellor of the Exchequer and related entries, the minutes of the British parliament and the total liability effects of each Budget.[9] Furthermore, we discriminate between 'anticipated' and 'non-anticipated' tax shocks. We classify as 'non-anticipated' all those liability changes that were implemented during one quarter from their announcement date and retain only these 'non-anticipated' exogenous tax changes. By doing so we avoid possible antici-

[5]We group in the OT residual category, stamp duties, landfill tax, business rate, aggregate levy, climate change levy and other minor taxes.

[6]In the UK, VAT is levied on the sale of goods and services by registered businesses (those with annual turnover above some threshold level or who choose to register voluntarily). It applies to all sales, both to private consumers and other businesses. Under the 'invoice-credit' form of the VAT, registered businesses redeem the VAT they pay on their purchases ('input VAT') against the liability ('output VAT') on their sales, remitting only the net amount due. If this chain of output tax and input credit remains unbroken, no net revenue is collected from the taxation of intermediate goods sales, and that the ultimate base of the tax is final consumption. See Crawford, Keen, and Smith (2010) for further details.

[7]Cloyne (2012, 2013) provide a narrative accounting from 1945. However, available data on consumption tax, in particular VAT revenues, necessary to calculate our measure of ACT, are available only from 1973. Furthermore, we do not distinguish between personal and corporate taxation because precise data on the fiscal revenues for these two tax categories is not available prior to 1987. It is important to note that in terms of fiscal revenues, corporation taxes are small compared to personal income taxes. For example, in 2001 corporation tax revenues account for just a fifth of total income tax receipts.

[8]We consider NIC as part of income tax. As the Institute for Fiscal Studies (briefing note n. 9, 2009) explains: 'The link between the amount contributed and the benefit entitlement, which was once close, has now almost entirely gone and substantial progress has been made in aligning the NI rate structure and tax base with those of income tax. Most of this has occurred in the last 25 years'. This covers for most part of the sample considered here.

[9]The minutes of the British Parliament can be downloaded at `http://hansard.millbanksystems.com/`, while total the total liability effects of each Budget are taken from Cloyne (2012) and `www.jamescloyne.webspace.virginmedia.com/data.pdf`.

pation effects of tax reforms, see Mertens and Ravn (2012).[10] Appendix A (Table A.2) details the classification method of the legislated tax changes, while the detailed description of all the non-anticipated exogenous consumption and income tax changes can be found in the companion paper Nguyen, Onnis, and Rossi (2017).

We construct a quarterly dataset for each type of tax liability change by assigning the change in the projected full-year revenue of the corresponding type to the specific quarter of the implementation date. Counting multiple actions in a given quarter, our dataset consists of around 60 exogenous, non-anticipated consumption tax changes and 70 exogenous, non-anticipated income tax exogenous changes over the sample. Then we convert the exogenous series of income and consumption tax liability changes into the corresponding average tax change as

$$\Delta\tau_t^{i,n} = \frac{(NIC_t + PI_t + CI_t + PetIT_t)}{\text{Aggregate Income at time } t}; \tag{7}$$

$$\Delta\tau_t^{c,n} = \frac{(VAT_t + Duties_t)}{\text{Aggregate Consumption at time } t\text{-VAT and Duties receipts at time } t}, \tag{8}$$

where $\Delta\tau_t^{i,n}$ and $\Delta\tau_t^{c,n}$ represent the narrative average tax changes in income and consumption respectively as fraction of their corresponding contemporaneous tax bases. We use aggregate income, i.e. the sum of employment income (compensation of employees), self-employment income (mixed income) and profits (gross operating surplus) as the income tax base.[11] Results are virtually identical if we scale the narrative tax changes by their tax bases at time $t-1$ or if we use GDP as income tax base. The variables $\Delta\tau_t^{i,n}$ and $\Delta\tau_t^{c,n}$ are our 'proxies' used to identify the latent structural AIT and ACT shocks.

Our narrative dataset contains changes in tax liabilities for which the historical documents show that they were not directly motivated by countercyclical considerations or in conjunction with spending changes. Of course, these changes still occurred with

[10]Using 'anticipated' tax changes as an external instrument in the 'Proxy-SVAR' framework invalidates the interpretation of the VAR reduced form residuals as prediction errors. In fact, due to the presence of foresight, the conditioning variables may not span the information set of forward looking agents.

[11]Precise description of the data can be found in Appendix A.

certain political purposes. For example, during the Conservative Party tenure (1979-1996), the government shifted taxation from income to consumption in order to pursue long-run efficiency and growth. Therefore, it is not surprising that two narrative series $\Delta\tau_t^{i,n}$ and $\Delta\tau_t^{c,n}$ display a negative (and significant at 95 percent) correlation of -0.27. Given this correlation, it would not be appropriate to consider the narrative income (consumption) tax changes as uncorrelated with exogenous shocks to the consumption (income) tax rate. Indeed, the strong link between income and consumption taxation pushes for a joint analysis of these two policy instruments. As we show later, ignoring this correlation radically changes the impulse response functions (IRFs) of the two tax shocks.

We further construct a measure of average income tax (AIT) and average consumption tax (ACT) as

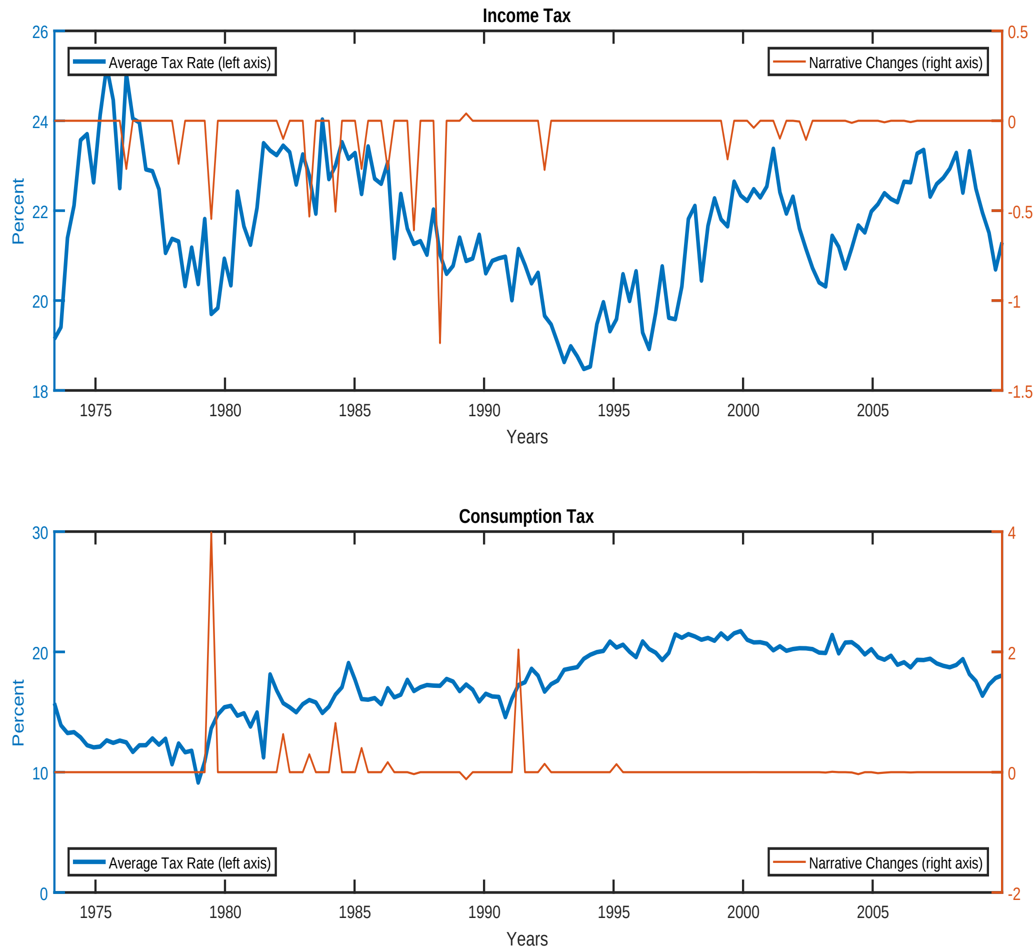
$$AIT_t = \frac{\text{Total income tax receipts at time } t}{\text{Aggregate Income at time } t}; \tag{9}$$

$$ACT_t = \frac{\text{VAT receipts at time } t + \text{Duties receipts at time } t}{\text{Aggregate Consumption at Time } t\text{-VAT and Duties receipts at time } t}. \tag{10}$$

Figure 1 presents the narratively identified policy shocks together with the average income (AIT) and consumption tax (ACT) rates as in (9) and (10), respectively. Both tax rates display a considerable variation over the sample under study. This variation is due to exogenous policy changes expressed in $\Delta\tau_t^{c,n}$ and $\Delta\tau_t^{i,n}$ but reflects also the endogeneity of average tax rates to macroeconomic conditions, distribution of income, changes of preferences in consumption *et cetera*. In particular, the AIT presents a cyclical pattern over the medium run, varying from almost the 25 percent in 1974 to as low as 19 percent in 1993. Differently, ACT presents a upper trend passing from as low as 9 percent in 1978 to around 20 percent after 1996.

The careful classification of the tax changes provided in Cloyne (2012) (that we complement), should guarantee that $\Delta\tau_t^{i,n}$ and $\Delta\tau_t^{c,n}$ comply with the exogeneity condition in (4). However, there could still exist cases where some 'exogenous' tax changes were rumoured or trailed prior to their announcements, such as in electoral manifestos, and macro variables may have responded to those rumours before the changes were officially implemented, thus leading to a correlation between the policy changes

Figure 1 – Average tax rates (Left Axis) and narrative shock measures (Right Axis), UK 1973:III-2009:IV.



and the lags of macro variables. In Appendix B we show that this is the case for the policy changes implemented right after two general elections, in 1979:III and in 1988:II. We therefore eliminate these two episodes from our narrative account. We then conduct Granger causality tests to check whether our resulting narrative measure of tax changes can be predicted from lagged macro variables such as GDP, consumption, investment and government spending.[12] Table 1 shows that the hypothesis of no

[12]We use four lags of the series, but results would be identical under different lag structures.

Granger causality cannot be rejected.[13]

Table 1 also presents a robustness check where we test whether or not the decision to act can be forecast from the past observations of macro variables using an ordered probit regression. As for the previous cases, the test indicate that we cannot reject the hypothesis of no predictability.

Table 1 – Exogeneity tests for $\Delta\tau_t^{c,n}$ and $\Delta\tau_t^{i,n}$.

Series	**Test Statistic**	***p*-value**
$\Delta\tau_t^{c,n}$		
Granger causality	10.06	0.86
Ordered probit	10.11	0.86
$\Delta\tau_t^{i,n}$		
Granger causality	10.83	0.82
Ordered probit	14.78	0.55

4 Results

4.1 Benchmark Specification

This section presents the main results of the paper. First we describe our benchmark econometric model and discuss a number of issues related to the validity of our estimation technique. Then we present a set of estimated impulse response functions (IRFs) to average income and consumption tax changes. Finally, we propose a revenue-neutral policy exercise where negative shocks in income taxes are compensated by positive shocks in consumption taxes.

Our benchmark econometric specification consists of a VAR with eight variables,

$$X_t = [ACT_t, AIT_t, ln(Y_t), ln(C_t), ln(I_t), ln(G_t), R_t, ln(P_t)]\,, \tag{11}$$

where ACT_t, AIT_t are the average consumption tax and income tax as presented above. Then, $ln(Y_t)$, $ln(C_t)$, $ln(I_t)$, $ln(G_t)$ are the GDP, private consumption, investment and

[13] We also test whether our constructed narrative series are Granger caused by movements in fiscal variables (government spending, ACT, and AIT), labour variables (wages and unemployment rate), inflation and monetary variables (interest rate and monetary base). None of these tests rejects the hypothesis of Granger non-causality of these variables to our narrative tax measures.

government spending respectively. These variables are expressed in logs, real per-capita. Finally, R_t represents the three-month treasury bill rate and $ln(P_t)$ is the (log of) prices, i.e. the deflator of household consumption expenditures, the closest measure to the Consumer Price Index (not available for the entire sample).[14] All variables are at quarterly frequency and our sample is 1973:III-2009:IV. Based on the Akaike Information Criterion and the Likelihood Ratio Test, the lag length in the VAR is set to three. However we have also experimented with different lags and results do not change, see Appendix C.

Given the observed correlation between the two tax categories, i.e. simultaneity, when measuring the effects of a change in one tax, one needs to control for contemporaneous exogenous changes in the other tax. We accomplish this task by adopting a standard Cholesky decomposition. To illustrate this point, we express the relationship between the VAR residuals of μ_t and the structural shocks ϵ_t as

$$\mu_{1,t} = \varrho \mu_{2,t} + B_1 \epsilon_{1,t} \tag{12}$$

$$\mu_{2,t} = \vartheta \mu_{1,t} + B_2 \epsilon_{2,t} \tag{13}$$

where $\mu_{1,t}$ and $\epsilon_{1,t}$ are the reduced-form innovations and the structural tax rate shocks, respectively, while $\mu_{2,t}$ and $\epsilon_{2,t}$ represent the reduced form innovations and the structural shocks of all other variables in our benchmark specification (11). The matrices ϱ, ϑ, B_1, and B_2 hold the structural coefficients of $\mathbf{\Gamma}$, introduced in (1). In order to study the response to $\epsilon_{1,t}$, we require the identification of the first two columns of $\mathbf{\Gamma}$, i.e. γ_1. By using (2), (12) and (13), we can derive γ_1 as

$$\gamma_1 = \begin{bmatrix} \mathbf{I} + \varrho(\mathbf{I} - \vartheta\varrho)^{-1}\vartheta \\ (\mathbf{I} - \vartheta\varrho)^{-1}\vartheta \end{bmatrix} B_1. \tag{14}$$

With the restrictions in (6), we identify the term in the square bracket $\gamma_1 B_1^{-1}$ and the covariance matrix $B_1 B_1'$. In particular, we estimate ϑ by using q_t as instrument for $\mu_{1,t}$. Once we have estimated ϑ, we use $(\mu_{2,t} - \vartheta\mu_{1,t})$ as instruments to estimate ϱ.

[14] See Appendix C for results with different definitions of the price index.

Then, the covariance of $(\mu_{1,t} - \varrho\mu_{2,t})$ can be used to estimate $B_1 B_1'$. However, the covariance restrictions are not sufficient to obtain B_1. To deal with this issue, we employ a Cholesky decomposition of $B_1 B_1'$ which results in a lower triangular B_1. Accordingly, the orderings of tax rates in the model could matter. For instance, when the ACT is ordered before the AIT, a negative one percentage point AIT shock lowers the AIT by one percentage point but leaves the ACT unchanged in cyclically adjusted terms, after controlling for contemporaneous response from $\mu_{2,t}$. In contrast, a negative one percentage ACT shock leads to a change in the AIT through not only feedback from $\mu_{2,t}$ but also a direct response to the ACT shock caused by the correlation between both tax rates (except when $B_1 B_1'$ is diagonal). Our empirical results show that the correlation between the cyclically adjusted tax rate innovations is -0.06 with 95 percent confidence bounds $[-0.89, 0.67]$, which implies that the responses to a specific tax shock are similar regardless of the ordering of tax rates within our 'Proxy-SVAR'. Therefore, for sake of brevity, all our IRFs to a tax rate shock show the estimates resulting from ordering that tax rate last, leaving the other tax rate unchanged in cyclically adjusted terms. The reader can find the estimates resulting from the alternative ordering of the tax rate of interest in Appendix C.

Our 'Proxy-SVAR' specification treats the narrative dataset as imperfect measures of the latent structural tax shocks. Thus, it produces estimates that are robust to potential measurement errors, as long as conditions (3) and (4) are satisfied. We discussed condition (4) in Section (3). In order to deal with condition (3), i.e. how much information our narrative dataset contains about the latent structural shocks, we adopt a reliability test as in Mertens and Ravn (2013, 2014). This test consists in calculating the eigenvalues of the statistic Ξ as,

$$\Xi = \frac{1}{d}\Sigma_{qq'}^{-1}\Sigma_{q\mu_1'}\left(\gamma_{11}\gamma_{11}'^{-1}\right)\Sigma_{q\mu_1'}', \tag{15}$$

where d is the fraction of uncensored observations of q_t. The eigenvalues of Ξ can be seen as measure of reliabilities of the principal components of the uncensored observations in q_t. In our benchmark specification, the eigenvalues of the estimated Ξ matrix are 0.26 and 0.60, with 95 percent bounds [0.14, 0.40] and [0.42, 0.85]. Therefore the correlations between the principal components of the narrative tax changes and the latent tax shocks are 0.51 and 0.77. The reliability matrix reveals that the proxies contain useful information for the identification of the structural tax shocks and that there is a fairly strong relation between the SVAR shocks and narrative changes to the tax components. However, the fact that the eigenvalues of Ξ are substantially smaller than 1, indicates that measurement error is a serious matter in practice. We also calculate the R^2 statistics for regressions of the reduced form residuals of AIT and ACT on non zero observations of the proxies q_t. We find values of 0.18 and 0.18, which show that the narrative changes explain a considerable fraction of the prediction error variance of the average income and consumption tax rates.

Now we turn our attention to the analysis of the IRFs. We report the point estimates of the response of the observables in (11) to a 1 percentage point cut in each tax rate, respectively. We also calculate the 90 and 95 percent confidence intervals with a recursive wild bootstrap, correcting for potential small sample bias, using 10 000 replications, see Kilian (1998) and Gonçalves and Kilian (2004). We start by exploring the effects of an unexpected AIT change. Figure 2 shows that a 1 percentage point cut in income tax implies for the AIT to remain significantly below zero for 5 quarters and then to slowly return to its long-run mean. The key finding from this exercise is that cuts in AIT generate large and persistent short-run expansions on total output and its (private) expenditure components. At the same time, we also find that public spending decreases significantly in response to a AIT cut. All these variables display a hump-shaped pattern, as typically found by the empirical macroeconomic literature. GDP increases significantly on impact by around 0.89 percent ($p = 0.002$), it peaks at 1.62 percent with 95 percent bounds [0.82, 3.45] in the fourth quarter after the tax shock and then gradually fades away. The effect of a AIT cut is significant from the period when the shock occurs until 7 quarters later, i.e. over a two-year window. Regarding the private expenditure components, households final consumption increases by 1.13 percent on impact ($p = 0.005$), it reaches its peak at 1.61 percent after 5 quarters, with 95 percent bounds [0.57, 3.69] and it remains significant both at 95 and 90 percent for more than two years. This suggests that income tax shocks have very similar effects on private consumption and on GDP. Private investments display the largest response to income tax shocks: a 1 percentage point cut in AIT generates a contemporaneous increase in investment of 3.93 percent ($p = 0.002$), with a maximum effect of 5.54 percent one year after the shock, with 95 percent bounds [2.32, 12.25]. The significance of the tax shock on this expenditure component lasts for 10 quarters.

Despite the boost in the overall economic activity, tax revenues decrease on impact by around 4 percent, and they remain significantly below their average for six months.[15] This result is expected, as the UK appears to be on the increasing side of the Laffer Curve for income tax components (a cut in average tax rates implies a decrease in revenues), see Trabandt and Uhlig (2011, 2012). More surprisingly, public spending decreases on impact by roughly 4 percent ($p \approx 0$) and remains below its average for several quarters, thus reflecting some sort of fiscal discipline: as the tax revenues decrease so does government consumption.[16] This result is particularly interesting, as the fiscal policy literature generally finds that aggregate tax cuts do not 'Starve the Beast', i.e. government spending does not respond to tax changes, e.g. Romer and Romer (2009). It is important to stress that our narrative account of tax shocks does not include spending-driven, deficit reduction or countercyclical motivated tax changes. These episodes are classified as endogenous and their inclusion would likely bias our results either towards or against a 'Starve the Beast' effect. Differently, here the negative response of government spending to income tax cuts is driven by exogenous policy changes that are not made in response to current macroeconomic conditions or in con-

[15]In order to calculate the response of fiscal revenues to a tax change, we combine the estimated responses of the specific tax base and tax rate. The formulas for the response of income tax revenue and consumption tax revenue read as, respectively

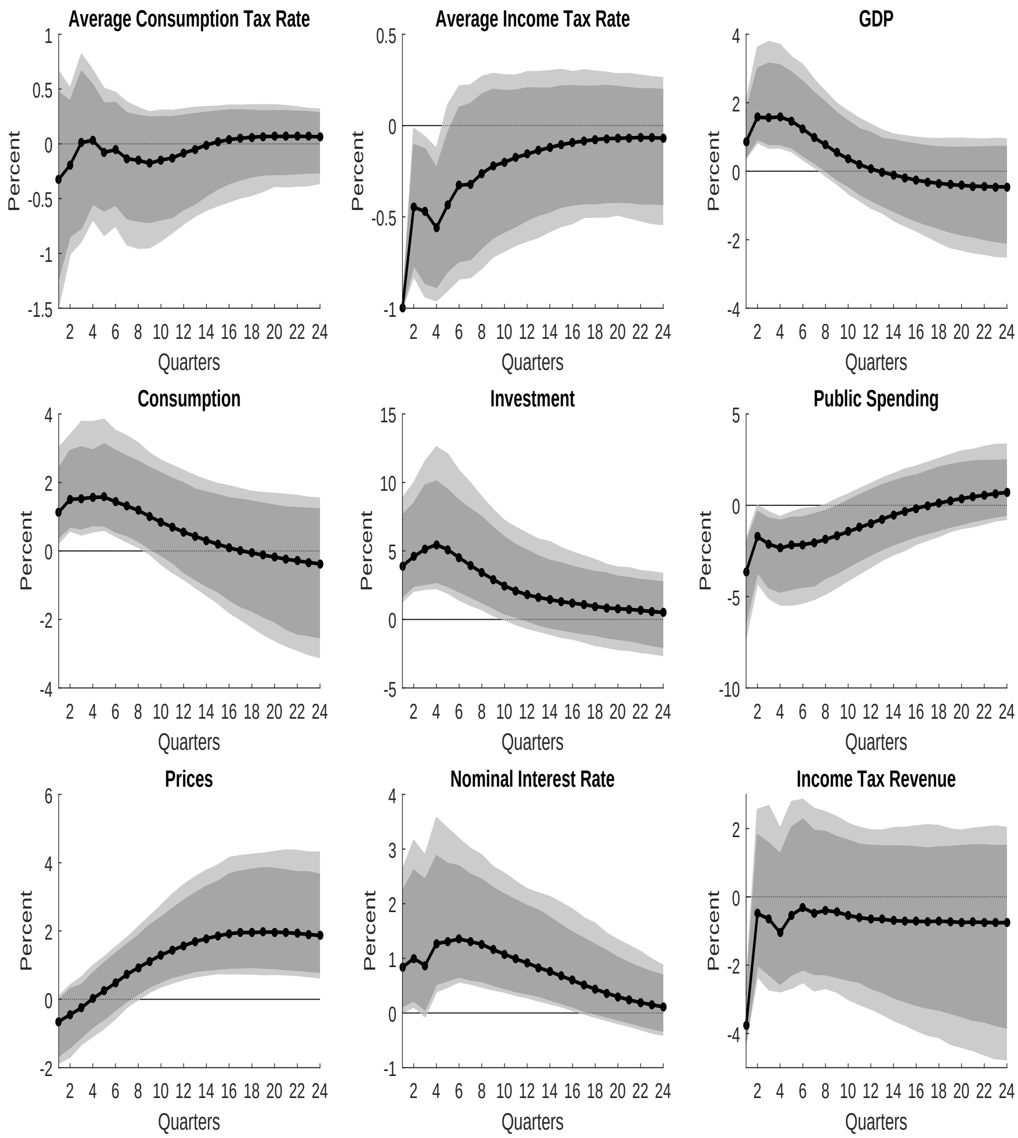
$$\widehat{tr}_t^{AIT} = \frac{\widehat{AIT}_t}{\overline{AIT}} + \widehat{y}_t,$$

$$\widehat{tr}_t^{ACT} = \frac{\widehat{ACT}_t}{\overline{ACT}(1 + \overline{ACT} + \widehat{ACT}_t)} + \widehat{c}_t,$$

where $\widehat{x}_t$ represents the impulse response of x_t, lower case letters represent logged variables, and overlined variables denote the mean average of the corresponding variable.

[16]This implies a reduction of the government spending-to-GDP ratio of around 0.9 percent.

Figure 2 – Benchmark specification: 1% cut in AIT. Dark and light shaded areas represent the 90 and 95 percent confidence intervals, respectively.



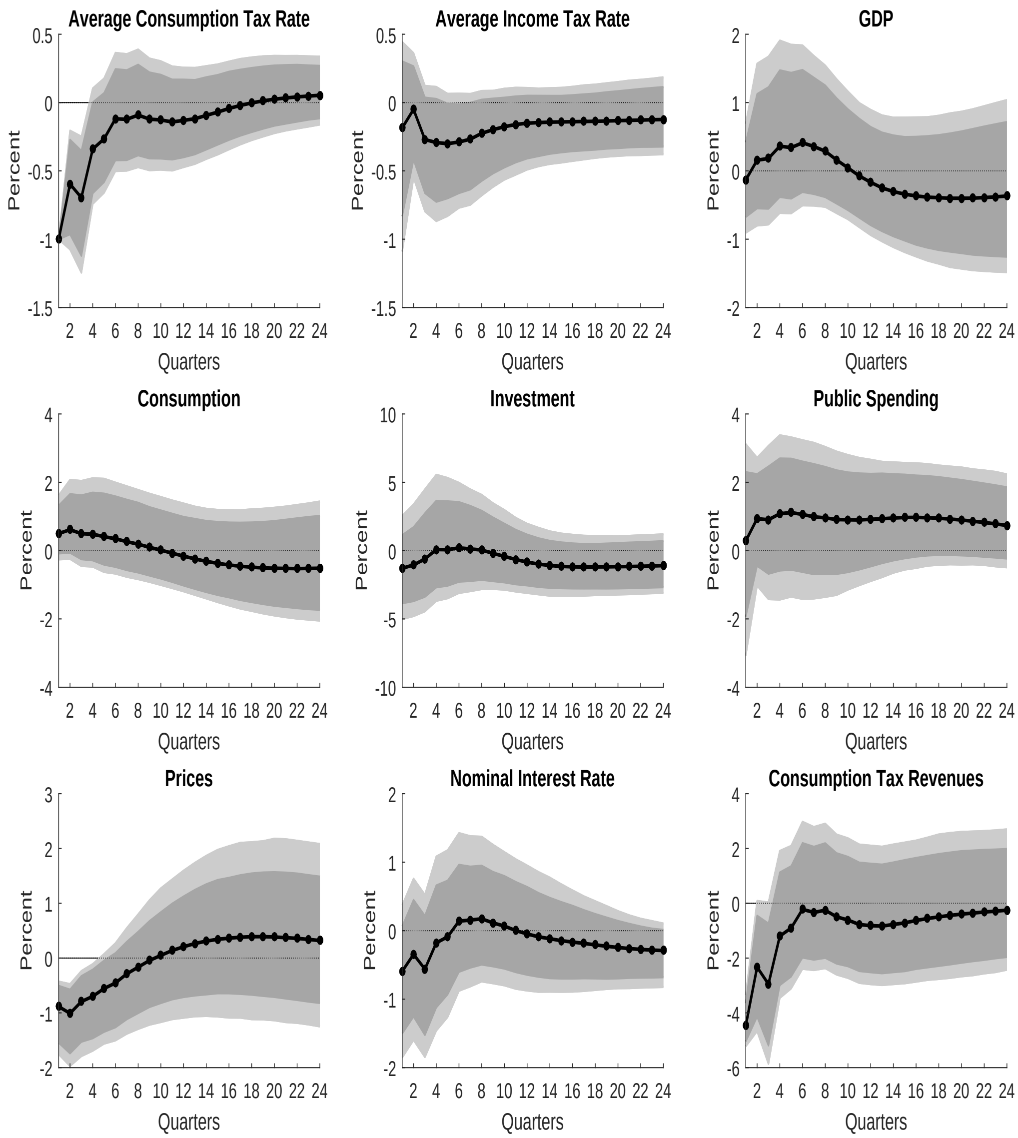
junction with spending changes. Despite this, there might still be a bias that works in the direction of a 'Starve the Beast' effect: a desire for smaller government. However the motivations of our exogenous tax changes rarely refer to future adjustments in government spending. Therefore this source of bias is likely to be limited.

Our benchmark econometric specification also controls for a measure of the price level and for monetary policy. This is important for at least two reasons. First, the sign of the price level response is indicative of whether the expansionary effects of income tax cuts are primarily derived from an increase in aggregate demand or in supply of final goods. Secondly, the response of monetary policy is generally quite important in shaping the equilibrium effects of fiscal policy, see Leeper (1991) and Woodford (1995), among others. We find that prices fall on impact, -0.65 percent ($p = 0.041$), underlying a possible supply-side effect of tax shocks. Our estimate also indicate that prices increase over time, peaking at 1.97 percent ($p \approx 0$) 19 quarters after the shock and thus suggesting a demand-side effect over both the medium and long-run. The nominal interest rate is significantly above zero for 16 quarters after the tax change, with a maximum response at 1.39 percent ($p \approx 0$) in the sixth quarter, revealing some monetary policy concerns about the increase in the price level. Remarkably, our finding on prices and interest rate are similar to the effects of an aggregate tax shock presented in Cloyne (2013).

Now we turn our attention to consumption tax shocks. Figure 3 reports the IRFs to a one percentage point cut in ACT, together with the 90 and 95 percent confidence bounds. The decrease in the tax rate, implies for the ACT to remain significantly below zero for 3 quarters and then to slowly return to its long-run mean. Fiscal revenues for this tax component decrease significantly by around 4.5 percentage points. This result is consistent with the notion that the Laffer curve for consumption taxation peaks at infinity, i.e. the marginal increase in consumption tax revenues implied by a tax hike is always greater than its marginal distortionary effect on its tax base, see Trabandt and Uhlig (2011). We also register a pass-through of the consumption tax shock to the price level. This result is mainly due to the accounting measure of the price level, which includes consumption taxes. The nominal interest rate does not respond significantly, suggesting that monetary policy reacts weakly to price changes caused by direct variations of consumption taxes.

The main result of Figure 3 is that, in contrast to the large and persistent expan-

Figure 3 – Benchmark specification: 1% cut in ACT. Dark and light shaded areas represent the 90 and 95 percent confidence intervals, respectively.



sionary effects found under income tax changes, shocks to consumption taxes have a modest effect on GDP and its components. This is a robust feature of our analysis and it suggests that, as predicted by public finance theories, the transmission mechanisms of direct and indirect tax components are fundamentally different. As we will discuss in detail in 4.2, these results rely crucially on the identification method adopted that allows for measurement error and for correlation between income and consumption tax shocks. Failing to account for any of these two features would alter radically our IRFs analysis.

Facing an ACT cut, private consumption expands, with a maximal response of 0.61 percent in the quarter after the shock and a borderline significance level ($p = 0.065$). Meanwhile, the effects on total output and its other expenditure components are modest and not statistically different from zero for standard significance levels. The response of GDP is insignificant at all forecast horizons and peaks below 0.5 percent six quarters after the shock ($p = 0.18$). The response of private investment is negative for part of the horizon, with a maximal reduction of 1.30 percent. However this response is not statistically different from zero ($p = 0.14$). We also register a positive response of government spending, but again, this effect is not significant at standard levels.

In order to gain a deeper quantitative understanding of the effects of income and consumption tax shocks, we calculate the impact of a tax change along the entire path of the response, up to a given period, see, for example, Mountford and Uhlig (2009). To this end, Figure 4 shows the discounted cumulative response of GDP and fiscal variables to a 1 percentage point cut in either AIT or ACT in the benchmark model (11).[17] Furthermore, in Table 2, we also report the present value multiplier at a generic lag k (PVM_k), where we normalise the cumulative discounted response of GDP by the movement of the tax rate along the response path, i.e.

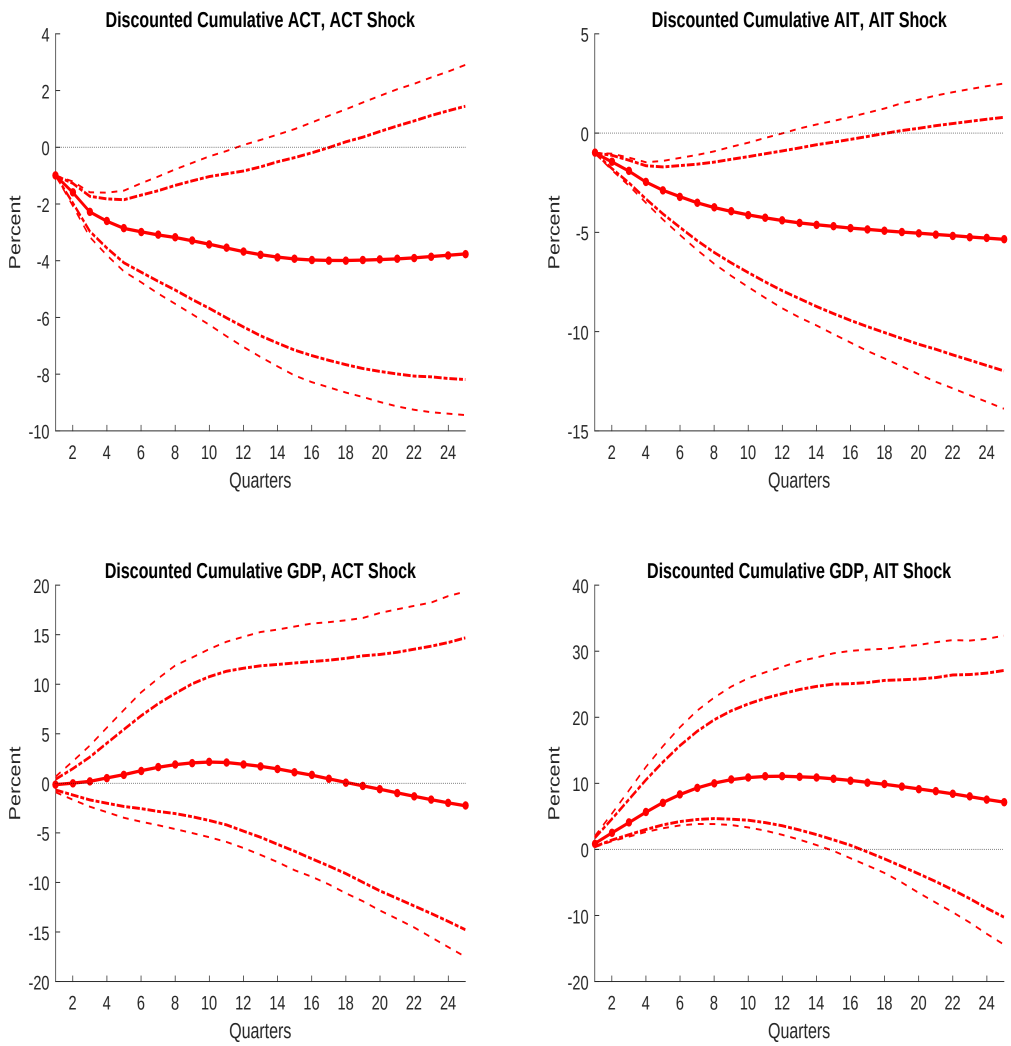
$$PVM_k = \frac{\sum_{i=0}^{k} (1+\bar{r})^{-i} y_i}{\sum_{i=0}^{k} (1+\bar{r})^{-i} f_i}, \tag{16}$$

where y_i is the response of GDP at period i, f_i is the response of the tax rate at period i, $\bar{r}$ is the average quarterly real interest rate over the sample.

Figure 4 and Table 2 tell the same story. They show that income tax cuts have a much greater effect on GDP than consumption tax. The cumulative discounted response of GDP to a ACT cut is never significant, whereas that for the AIT is signifi-

[17] We compute the discount factor as the inverse of the sample average gross quarterly real interest rate.

Figure 4 – Cumulative discounted response of ACT and AIT shocks. Broken lines indicate 90 and 95 percent confidence intervals, respectively.



cantly positive throughout up to four years after the shock, with a maximal cumulative response of around 11 percent. Figure 4 also shows that the 90 and 95 percent standard errors of the cumulative response become vary large at the later periods. Table 2 indicates that the maximal present value multiplier (PVM) on GDP of a one percentage point decrease in AIT is 2.68 percent, with a 5 percent significance level. This occurs two years after the shock. Differently, the maximum present value multiplier on GDP

Table 2 – Present value multipliers of different tax shocks

	1 qrt	4 qrts	8 qrts	12 qrts	20 qrts	Maximum
AIT Shock	0.89***	2.29***	2.68**	2.52*	1.81	2.68** (qrt 9)
ACT Shock	-0.13	0.21	0.59	0.52	-0.15	0.63 (qrt 10)

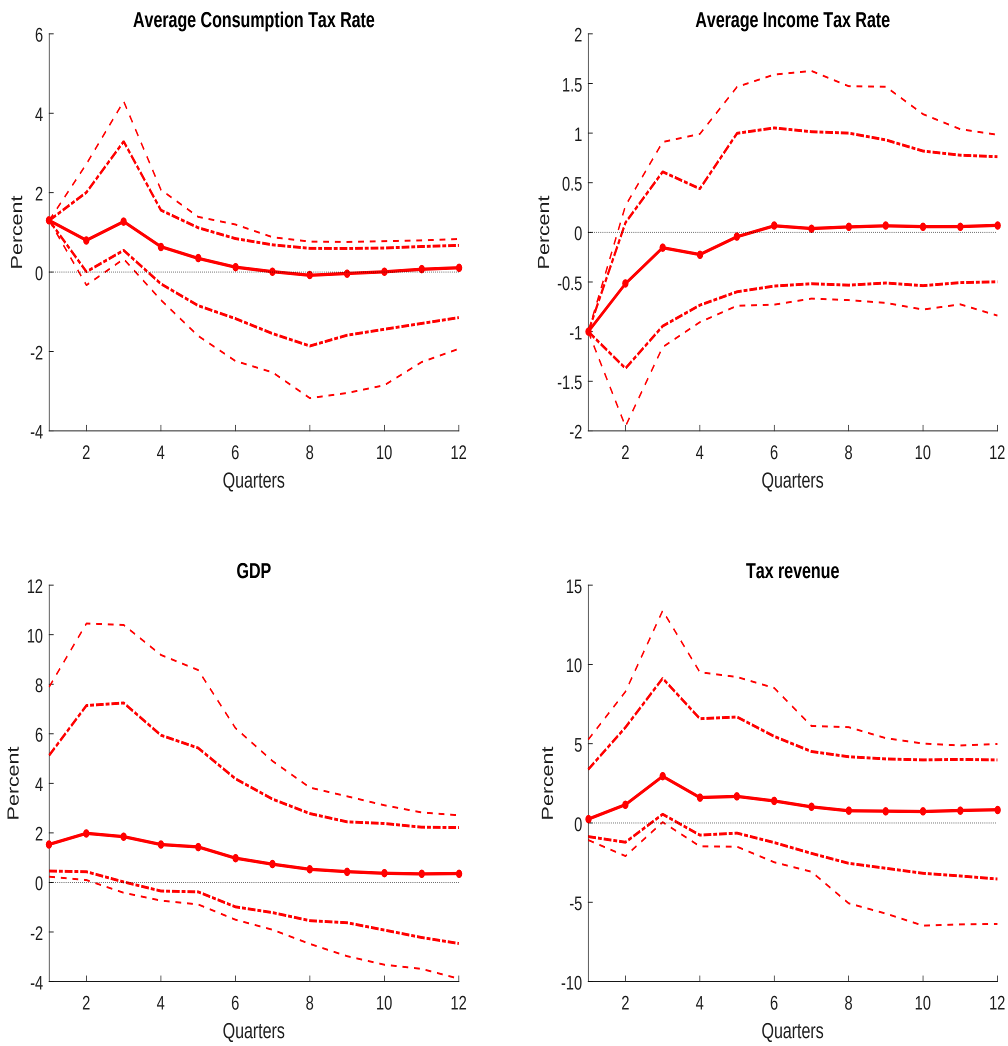
Notes: Present value multipliers for one percentage point cut in AIT and ACT. Asterisks ***, ** and * denote 1, 5 and 10 percent significance, respectively.

of a one percentage point decrease in ACT is rather small, i.e. 0.63 percent, and it is not statistically different than zero for standard significance levels.

Having explored the quantitative outcome of AIT and ACT changes, we now turn our attention to the effects of shifting the burden of taxation from income to consumption. Figure 5 presents a revenue neutral policy experiment where an exogenous decrease in AIT is compensated by an exogenous increase in ACT so that at the time of the implementation of the policy, aggregate fiscal revenues stay constant. The analysis is conducted under our benchmark identification technique with the set of observables in (11). This estimation aims to mimic the fiscal reforms implemented in the UK during the period 1979-1996, where cuts in personal and corporate income taxes were often balanced with increases in VAT and other consumption duties such as on fuel, alcohol and tobacco. Our exercise is designed in such a way that at the time of the reform, AIT exogenously decreases by one percentage point and ACT increases by 1.3 percentage points. In this way fiscal revenues remain constant. ACT needs to increase by more than AIT as the consumption tax base is smaller than the income tax base.

The responses of average income and consumption taxes remain significant for two quarters after the reform. Crucially, this revenue neutral policy has a short (2 quarters) but significant (at 95 percent) expansionary effect on GDP. This result is important as it gives empirical support to the predictions of conventional public finance theories: revenue neutral policies where cuts in income taxes are balanced by increases in consumption taxes increase efficiency and therefore are expansionary, see inter alia Coleman (2000), Altig, Auerbach, Koltikoff, Smetters, and Walliser (2001) and Correia (2010).

Figure 5 – Revenue neutral policy experiment. Broken lines indicate 90 and 95 percent confidence intervals, respectively.



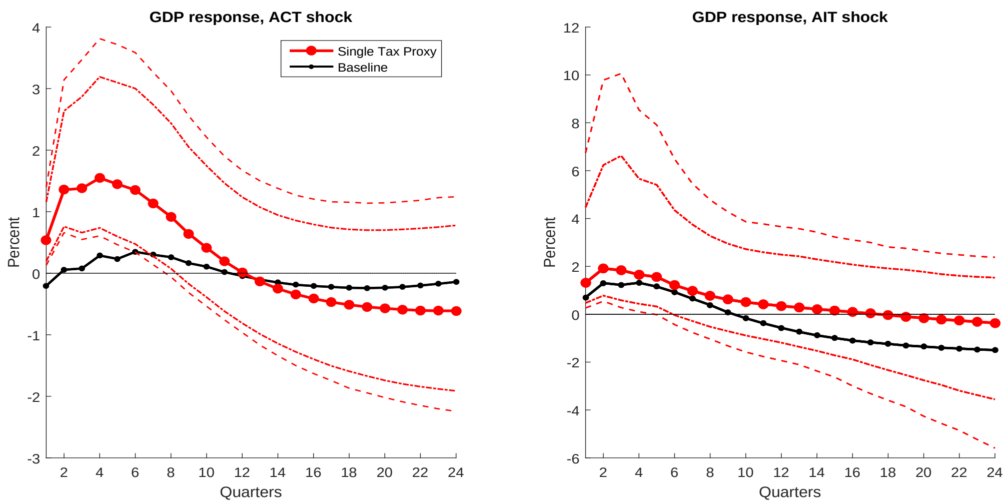
4.2 Alternative Identifications

This section discusses in further details several aspects of our identification procedure. Our goal is to obtain a deeper insight of our benchmark findings and how these relates to the existing literature. To this end, we start by highlighting the importance of allowing for non zero cross-correlations between the measured tax changes and structural tax shocks. Next, we compare our results to those using identification procedures commonly adopted in the narrative literature. Finally we discuss the potential problem of timing errors in our constructed series.

As described earlier in the paper, our narratively identified changes in income and consumption taxes correlate with each other. Therefore it is probable that changes in one tax component correlate with shocks in both taxes. As explained above, we control for contemporaneous variation in both taxes and we rely on standard Cholesky decomposition to deal with the insufficient number of identifying restrictions. Here we discuss the importance of controlling for the correlation between the two tax changes. We do so by making the alternative assumption that each of the proxy correlates with only one tax shock, e.g. the proxy on consumption tax changes correlates only with ACT. This means that each tax shock can be identified by using a single proxy. The underlying assumption behind this alternative specification is that the correlation between our narrative measures is random or solely due to measurement error. The econometric model is otherwise isomorphic to the benchmark. In this way we can fully isolate the consequences of ignoring the observed negative correlation between the proxies.

Figure 6 presents the IRFs for our benchmark econometric model under the alternative specification where each tax shock is identified with a single proxy. For comparison we also report the IRFs of our benchmark case. Bounds are calculated as before and kept at 90 and 95 percent. The key finding of this experiment is that a one percentage point decrease in ACT leads to a large and significant increase in GDP, with an impact response of 0.54 percent ($p = 0.007$) and a peak response of 1.55 percent in the fourth quarter, with 95 confidence bounds of [0.61, 3.81]. The implied maximal PVM of a one percentage point cut in ACT is 3.07, see Table 3. This outcome is in sharp contrast with the estimates obtained under our benchmark identification, where a shock to ACT has no significant effect on aggregate output. The sizeable difference obtained under the two approaches indicates that it is critical to control explicitly for the interactions between consumption and income tax components when measuring the effects of an ACT shock. Furthermore, this insight is particularly instructive because it enables us to understand the differences between our benchmark results described in 4.1 and the ones presented in Riera-Crichton, Vegh, and Vuletin (2016), who estimate large and significant short run expansionary effects of consumption tax cuts but, as here, impose zero correlation between the exogenous changes in consumption taxes and changes in other tax components.

Figure 6 – Baseline model and single tax proxy. Broken lines indicate 90 and 95 percent confidence intervals, respectively.



The estimates are somehow more similar to the benchmark identification for the case of AIT shocks. A one percentage point decrease in AIT results in a large and significant increase in GDP, with an impact response of 1.31 percent ($p = 0.018$) and a peak response of 1.91 percent in the second quarter, with 95 confidence bounds of [0.55, 9.78] and a maximal PVM of 3.10 (Table 3). However the confidence bounds get larger with a single proxy approach. This suggests that controlling for the correlation between different tax components is important also for the analysis of the transmission mechanism of income tax shocks.

Next we compare our benchmark results with an alternative identification strategy where the narrative series are added to the econometric model in (11) as exogenous regressors, see inter alia Romer and Romer (2010). In order to do so, we estimate the

Table 3 – Summary statistics for Proxy-SVAR models

Model	AIT Shock		ACT Shock		Reliability
	Impact	$\text{PVM}^{\max}$	Impact	$\text{PVM}^{\max}$	
Benchmark (Figures 2 and 3)	0.89***	2.68**	-0.13	0.63	0.26, 0.60
Single Proxy (Figure 6)	1.31***	3.10*	0.56***	3.07**	0.36, 0.55
Timing Error (Figure 8)	0.88***	2.37***	-0.12	0.62	0.08, 0.29
With Unemployment (Figure 9)	0.61***	2.20**	-0.15	0.68	0.29, 0.62
With Debt (Figure 10)	0.71***	2.25**	-0.20	0.49	0.31, 0.65

Notes: $\textbf{PVM}^{\max}$ identifies the maximal present value multipliers for one percentage point cut in either AIT or ACT. Asterisks ***, ** and * denote 1, 5 and 10 percent significance, respectively.

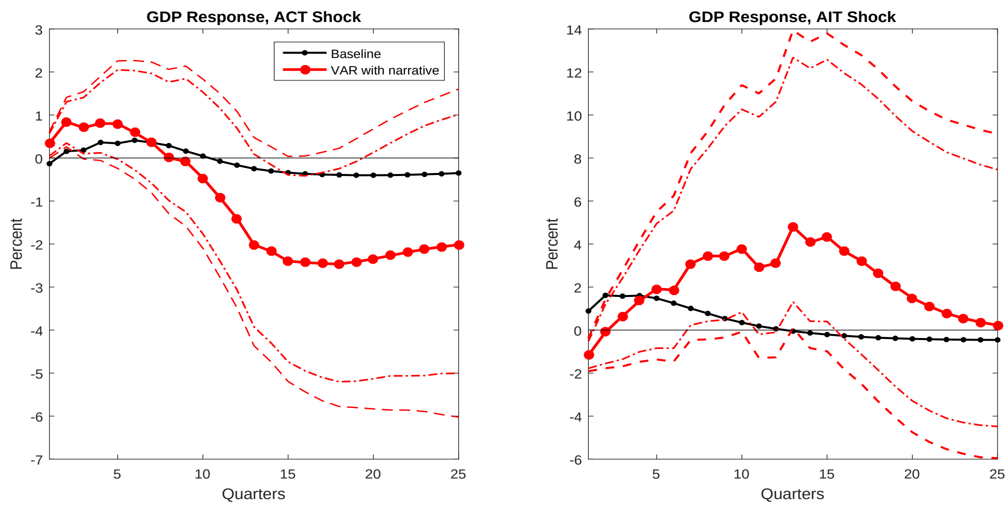
following model:

$$X_t = \sum_{i=1}^{P} \beta_i X_{t-i} + \sum_{s=0}^{Q} \zeta_s \Delta\tau_{t-s}^{j,n} + \mu_t, \text{ with } i = 1, 2, ..., P \text{ and } s = 0, 1, 2, ..., Q; \qquad (17)$$

where $\Delta\tau_t^{j,n}$ ($j = i, c$) are the narratively identified tax changes defined in (7) and (8), X_t is the set of observables as in (11), μ_t represents the reduced form innovations, $P = 3$ and $Q = 12$. Figure 7 reports the IRF to a one percentage point cut in $\Delta\tau_t^{j,n}$ as well as our benchmark results for comparison. The two approaches lead to sizeable differences for both ACT and AIT cuts. In response to a one percentage point cut in consumption taxation, GDP expands significantly above zero for four (six) quarters at 95 (90) percent, with an impact response of 0.34 percent ($p = 0.027$) and maximal response in the second quarter of 0.84 percent and 95 percent bounds of [0.26, 1.41].

Moving to average income tax shocks, we find that a one percentage point cut in this tax component leads to a reduction in GDP on impact of 1.17 percent ($p = 0.005$). Aggregate output increases over time, turning positive at 90 percent confidence one year and a half after the tax shock occurs. The overall expansionary effect on GDP is much larger than the one obtained under our baseline identification, with a maximal effect of 4.80 percent, with wide 95 percent confidence bounds [0.09, 13.93]. Furthermore, compared to our benchmark 'Proxy-SVAR' estimates, the peak increase in GDP occurs at a later horizon. There two main reasons that can help explaining the

Figure 7 – Proxy-SVAR and VAR incl. narrative. Broken lines indicate 90 and 95 percent confidence intervals, respectively.



differences obtained under the two identification strategies. First of all, our baseline estimator is robust to the existence of random measurement errors, while the identification in (17) is not. As we discussed in Section (4.1), our estimates of the reliability of the proxies indicate that measurement error is quantitatively important. Second, estimates from (17) do not control for the observed negative correlation between the proxies and therefore, as for the case of a single proxy identification, may bias upwards the response to tax shocks. All in all, allowing for measurement errors and for the correlation between the two tax changes appear to be extremely important.

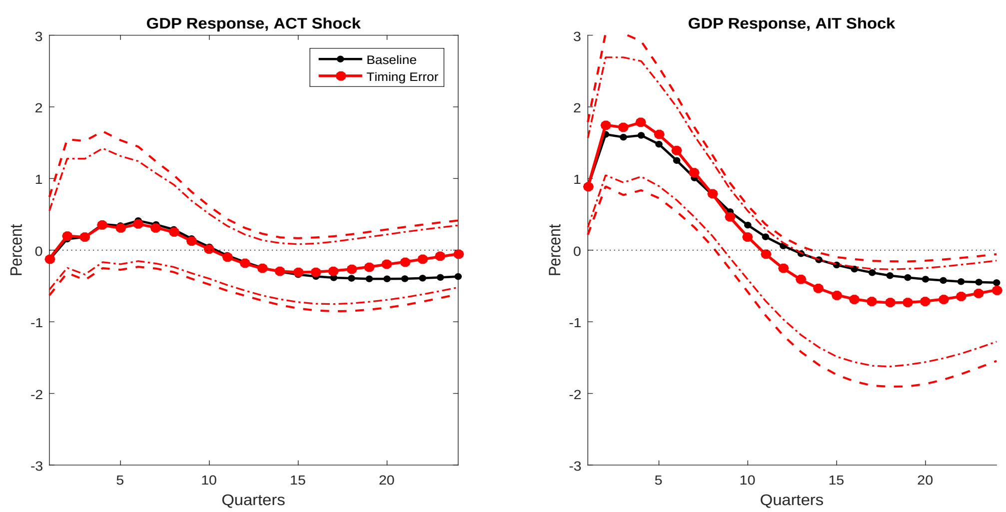
Finally we control for the possibility of timing errors in the construction of our narrative account of tax changes. In order to do this, we perform a simulation in the vein of Ramey (2011). This consists of constructing alternative narrative series $\Delta\tilde{\tau}_t^{j,n}$, with $j = i, c$ such that,

$$\Delta\tilde{\tau}_t^{j,n} = \left(1 - \varpi^j\right)\Delta\tau_t^{j,n} + \rho^j\varpi^j\Delta\tau_{t-1}^{j,n} + \left(1 - \rho^j\right)\varpi^j\Delta\tau_{t+1}^{j,n}, \tag{18}$$

with $\varpi^j \sim U\,(0, 0.5)$, $\rho^j \sim B\,(0.5)$. Measurement error is added by allowing up to 50 percent of the value of each observation of the two proxies to be mistimed by a quarter, so that ϖ^j is uniformly distributed between 0 and 0.5 and ρ^j takes the value of 0 with

50 percent probability, and 1 with 50 percent probability. In this way there is an equal chance of mistiming of leading and/or lagging a quarter. We ran (18) for 10000 times. For each replication we estimated the baseline Proxy-SVAR with the new generated instruments $\Delta\tilde{\tau}_t^{j,n}$ and calculate the IRFs. Figure 8 presents the median values of the estimated responses. For comparison we also report the benchmark specification. Not surprisingly, adding time uncertainty changes the response of GDP to AIT and ACT changes. The major noticeable difference is that output becomes slightly negative several quarters after an income tax cut. However the main conclusions of the paper are unaffected, see Table 3. This is somehow expected as our benchmark identification is robust to this type of random timing errors. Obviously, adding timing error decreases sensibly the reliability of our proxies, i.e. the eigenvalues of the statistic Ξ for $\Delta\tilde{\tau}_t^{j,n}$ drop to [0.08, 0.29].

Figure 8 – Proxy-SVAR with timing error. Broken lines indicate 90 and 95 percent confidence intervals, respectively.



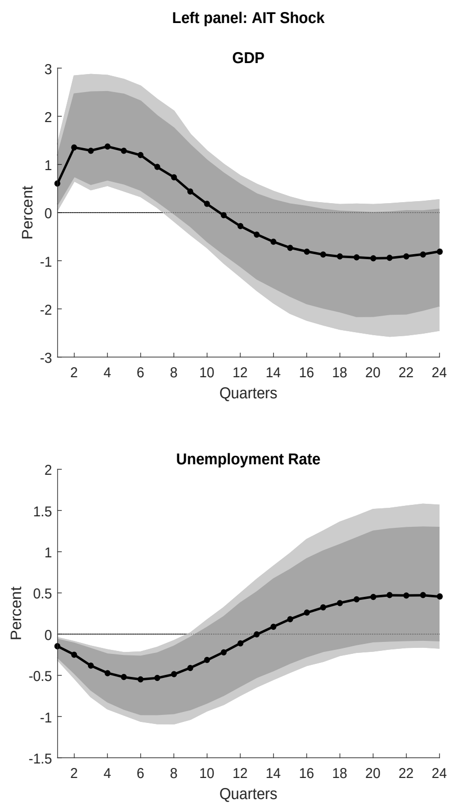
4.3 Further Results

This section presents a set of findings resulting from expanding the set of observable variables in our econometric model, while keeping unchanged our baseline identification strategy. For the sake of brevity, for each extension of the baseline model we report only the IRFs of the variables of interest in addition to the GDP. Moreover, we present the response to a shock to a tax component resulting from ordering that tax rate last, leaving the other unchanged in cyclically adjusted terms. Appendix C provides a more comprehensive number of robustness checks.

The first extension consists in including the unemployment rate in our baseline model. Controlling for unemployment is important as this variable is often the primary target of short and long run macroeconomic policies. Ideally one would like to analyse the effects of a fiscal shock on a wider set of labour market variables, but doing so would have implied losing too many degrees of freedom. Results from this estimation are presented in Figure 9. The GDP point estimates of a cut in AIT are remarkably similar to the benchmark, with a maximal increase of 1.40 percent and 95 percent confidence bounds of [0.63, 2.92]. The maximal PVM is slightly smaller than the benchmark case, but still above 2 (Table 3). The key result from this exercise is that a 1 percentage point decrease in AIT generates a significant reduction of the unemployment rate, with a peak effect of -0.56 ($p \approx 0$) 6 quarters after the tax cut. Differently, a one percentage point cut in ACT has a small and not significant effect on the unemployment rate and, as in the benchmark case, a modest and not significant effect on GDP at all forecast horizons, with a maximal PVM of 0.68. Therefore this experiment confirms the main conclusions of the paper. Cutting income taxation is very effective in stimulating the economy, while decreasing consumption taxation is not. Finally we note that the reliability of the proxies increases slightly compared to the benchmark, i.e. with this specification, the eigenvalues of the Ξ statistic are 0.29 and 0.62.

Next we expand our baseline model by inserting public debt. Government debt is a potentially important variable since any change in taxes eventually must lead to adjustments in the fiscal instruments. Unfortunately, UK data on public debt are only available at annual frequency prior to 1997. In order to deal with this problem, we construct a quarterly series of government debt-to-GDP ratio. We proceed as follows. First we construct the year-on-year quarterly public debt growth for the period 1956-2009 with missing values for Q2, Q3 and Q4 of each year up to 1997 and full data af-

Figure 9 – A percentage point cut in AIT (left column) and ACT (right column) on GDP and the unemployment rate. Dark and light shaded areas represent the 90 and 95 percent confidence intervals, respectively.



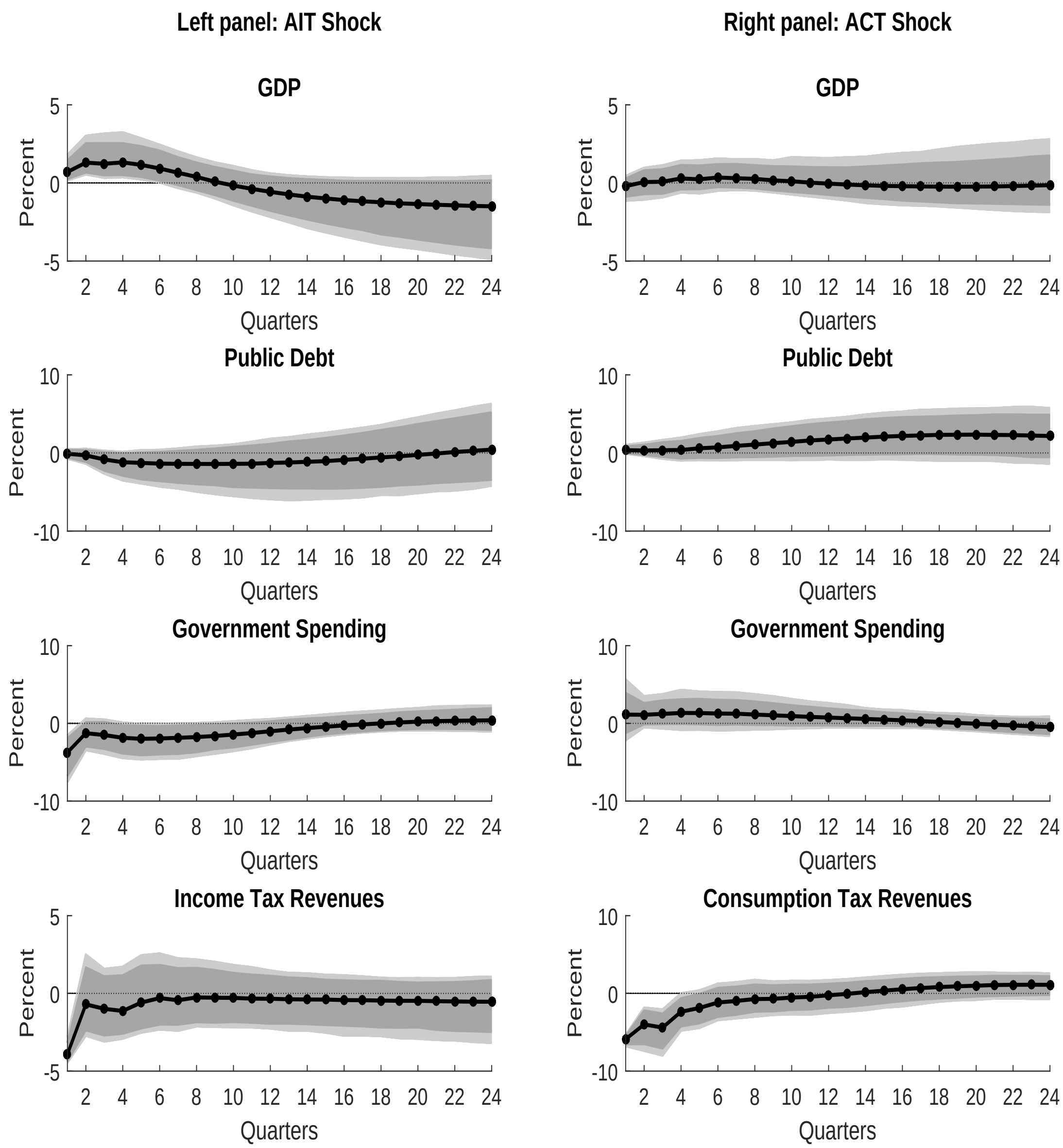
terwards. Then we estimate the missing quarterly values via Kalman Filter through a VAR with GDP growth, inflation and public debt growth. Finally, we use the smoothed series of debt growth to compute the Debt-to-GDP ratio at quarterly frequency for the 1973Q3-1997Q4 sample. For further details on this method, see Casals, Jerez, and Sotoca (2000). Figure 10 presents the estimates on the variables of interest of adding our constructed quarterly series of debt-to-GDP to the set of observables in (11). First, controlling for public debt leads to similar result as in our benchmark model. AIT shocks have large and persistent expansionary effect on GDP, with a maximal effect of 1.33 percent ($p = 0.002$) and a maximal PVM of 2.24 (Table 3). Moreover, as in the benchmark case, AIT shocks 'Starve the Beast', i.e. government spending decreases with the tax rate cut. At the same time, ACT shocks are not significant on GDP and government spending at all forecast horizons. Second, public debt remains roughly constant under income tax shocks. This is mainly due to the fact that most of the revenue consequences of the policy change are absorbed by public spending. Differently, the stock of public debt slowly builds up under consumption tax shocks, but it never becomes statistically different than zero at 90 and 95 percent. For this model, the values of the reliability test are [0.31, 0.65].

5 Conclusions

We analyse the aggregate effects of unexpected changes in consumption and income taxes in the United Kingdom and present a novel narrative account of exogenous tax liability changes in these two components for the period 1973-2009. We estimate 'Proxy-SVARs' by using these narrative tax changes as proxies for the latent structural tax shocks. We find important differences in the size and the significance of the two tax multipliers. The short run effects of income tax changes are large, significant and persistent on output, private consumption and investment. On the same variables, consumption tax changes have effects that are modest and in most cases, not statistically different from zero. Furthermore, we find that income tax cuts 'Starve the Beast', i.e. government spending decreases in response to shocks in this tax component. This effect is not present under consumption tax shocks. We think these results are important for two reasons. First they show that is crucial to distinguish between direct and indirect taxation when studying the transmission mechanism of fiscal policy as changes in these two tax categories have very different effects on key macroeconomic variables. Studies that focus exclusively on changes in total tax revenues generally miss these insights and therefore they can offer only a limited perspective in the relative merits of different tax components. Second, our results provide empirical support for the theoretical predictions of conventional public finance theories, where consumption taxes are less distortive than income taxes.

An important limit of our analysis is that we analyse two broad tax categories. For example, our definition of income taxation aggregates changes in personal and

Figure 10 – A Percentage point cut in AIT (left column) and ACT (right column) on fiscal variables including public debt. Dark and light shaded areas represent the 90 and 95 percent confidence intervals, respectively.



corporate taxation. This is due to data limitation. However, while both personal and corporate tax cuts are expansionary on GDP, their short-run effects are likely to be different on private consumption and the labour market, see Mertens and Ravn (2013).

When more data on fiscal revenues becomes available, it will be possible to study the effects of each narrowly defined tax categories in isolation.

There are several important avenues for future research. First, it would be interesting to extend our analysis on the effects of income and consumption taxation for different income groups as in Mertens (2015) or by grouping consumers by housing tenure as in Cloyne and Surico (2016). Second, it would be interesting to allow for time-varying effects of tax shocks in the vein of Auerbach and Gorodnichenko (2012) and Ramey and Zubairy (2014). In this way one could check whether the transmission mechanism of consumption and income taxation is different in good and in bad times or when the nominal interest rate is close to its zero lower bound. Finally, our dataset allows to analyse the effects of tax uncertainty on the economy, thus extending on the fiscal dimension the recent literature on volatility shocks, e.g. Caggiano, Castelnuovo, and Groshenny (2014) and Mumtaz and Theodoridis (2016).

References

Altig, D., A. J. Auerbach, L. J. Koltikoff, K. A. Smetters, and J. Walliser (2001). Simulating fundamental tax reform in the united states. *American Economic Review 91*(3), 574–595.

Auerbach, A. J. and Y. Gorodnichenko (2012, May). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy 4*(2), 1–27.

Auerbach, A. J., L. J. Kotlikoff, and J. Skinner (1983). The efficiency gains from dynamic tax reform. *International Economic Review 24*(1), pp. 81–100.

Barro, R. J. and C. J. Redlick (2011). Macroeconomic effects from government purchases and taxes. *The Quarterly Journal of Economics 126*(1), 51–102.

Blanchard, O. and R. Perotti (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics 117*(4), 1329–1368.

Blundell, R. (2009). Assessing the temporary vat cut policy in the uk. *Fiscal Studies 30*(1), 31–38.

Caggiano, G., E. Castelnuovo, and N. Groshenny (2014). Uncertainty shocks and unemployment dynamics in u.s. recessions. *Journal of Monetary Economics 67*, 78 – 92.

Caldara, D. and C. Kamps (2017). The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers. *The Review of Economic Studies*.

Casals, J., M. Jerez, and S. Sotoca (2000). Exact smoothing for stationary and nonstationary time series. *International Journal of Forecasting 16*(1), 59 – 69.

Cashin, D. and T. Unayama (2016). Measuring intertemporal substitution in consumption: Evidence from a vat increase in japan. *Review of Economics and Statistics 98*(2), 285–297.

Cloyne, J. (2012). Discretionary tax shocks in the United Kingdom 1945-2009: a narrative account and dataset. Working Paper 37739, MPRA.

Cloyne, J. (2013). Discretionary tax changes and the macroeconomy: New narrative evidence from the united kingdom. *American Economic Review 103*(4), 1507–28.

Cloyne, J. S. and P. Surico (2016). Household debt and the dynamic effects of income tax changes. *The Review of Economic Studies*.

Coleman, W. J. I. (2000). Welfare and Optimum Dynamic Taxation of Consumption and Income. *Journal of Public Economics 76*(1), 1–39.

Correia, I. (2010). Consumption Taxes and Redistribution. *American Economic Review 100*(4), 1673–1694.

Crawford, I., M. Keen, and S. Smith (2010, Apr). *Value added tax and excises*.

Cui, W. (2014). China's business-tax-to-vat reform: An interim assessment. *British Tax Review 5*, 617–641.

Engle, R. F. and D. F. Hendry (1993, March). Testing superexogeneity and invariance in regression models. *Journal of Econometrics 56*(1-2), 119–139.

Favero, C. and F. Giavazzi (2012). Measuring tax multipliers: The narrative method in fiscal vars. *American Economic Journal: Economic Policy 4*(2), 69–94.

Fernández-Villaverde, J. (2010). Fiscal policy in a model with financial frictions. *American Economic Review 100*(2), 35–40.

Gonçalves, S. and L. Kilian (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics 123*(1), 89 – 120.

Hendry, D. F. and G. E. Mizon (1993). Evaluating dynamic econometric models by encompassing the var. In P. Phillips (Ed.), *Models, Methods and Applications of Econometrics*, pp. 272–300. Blackwell.

Jones, L. E., R. E. Manuelli, and P. E. Rossi (1993). Optimal taxation in models of endogenous growth. *Journal of Political Economy 101*(3), pp. 485–517.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *The American Economic Review 95*(1), 161–182.

Kilian, L. (1998, May). Small-Sample Confidence Intervals For Impulse Response Functions. *The Review of Economics and Statistics 80*(2), 218–230.

Krusell, P., V. Quadrini, and J.-V. Ríos-Rull (1996). Are consumption taxes really better than income taxes? *Journal of Monetary Economics 37*(3), 475 – 503.

Leeper, E. M. (1991, February). Equilibria under 'active' and 'passive' monetary and fiscal policies. *Journal of Monetary Economics 27*(1), 129–147.

Mertens, K. (2015, June). Marginal tax rates and income: New time series evidence. Working Paper 19171, National Bureau of Economic Research.

Mertens, K. and M. O. Ravn (2012). Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated US Tax Policy Shocks. *American Economic Journal: Economic Policy 4*(2), 145–81.

Mertens, K. and M. O. Ravn (2013). The dynamic effects of personal and corporate income tax changes in the united states. *American Economic Review 103*(4), 1212–47.

Mertens, K. and M. O. Ravn (2014). A reconciliation of SVAR and narrative estimates of tax multipliers. *Journal of Monetary Economics 68*(S), S1–S19.

Mountford, A. and H. Uhlig (2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics 24*(6), 960–992.

Mumtaz, H. and K. Theodoridis (2016). The changing transmission of uncertainty shocks in the us: An empirical analysis. *Journal of Business & Economic Statistics 0*(ja), 1–39.

Nguyen, A. D., L. Onnis, and R. Rossi (2017). A Narrative Account of Income and Consumption Tax Changes in the United Kingdom 1973-2009. Technical appendix, University of Manchester.

Nishiyama, S. and K. Smetters (2005). Consumption taxes and economic efficiency with idiosyncratic wage shocks. *Journal of Political Economy 113*(5), 1088–1115.

OECD (2014). Consumption Tax Trends 2014. *OECD Policy Notes 80*(2).

Ramey, V. A. (2011). Identifying Government Spending Shocks: It's all in the Timing. *The Quarterly Journal of Economics 126*(1), 1–50.

Ramey, V. A. and S. Zubairy (2014, November). Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data. NBER Working Papers 20719, National Bureau of Economic Research, Inc.

Riera-Crichton, D., C. A. Vegh, and G. Vuletin (2016). Tax multipliers: Pitfalls in measurement and identification. *Journal of Monetary Economics 79*, 30 – 48.

Romer, C. D. and D. H. Romer (2009). Do Tax Cuts Starve the Beast? The Effect of Tax Changes on Government Spending. *Brookings Papers on Economic Activity 40*(1 (Spring), 139–214.

Romer, C. D. and D. H. Romer (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review 100*(3), 763–801.

Sims, E. and J. Wolff (2016, May). The State-Dependent Effects of Tax Shocks.

Stock, J. H. and M. W. Watson (2012, May). Disentangling the channels of the 2007-2009 recession. Working Paper 18094, National Bureau of Economic Research.

Summers, L. H. (1981, September). Capital Taxation and Accumulation in a Life Cycle Growth Model. *American Economic Review 71*(4), 533–44.

Trabandt, M. and H. Uhlig (2011). The Laffer curve revisited. *Journal of Monetary Economics 58*(4), 305–327.

Trabandt, M. and H. Uhlig (2012). How Do Laffer Curves Differ across Countries? In *Fiscal Policy after the Financial Crisis*, NBER Chapters, pp. 211–249. National Bureau of Economic Research, Inc.

Woodford, M. (1995). Price-level determinacy without control of a monetary aggregate. *Carnegie-Rochester Conference Series on Public Policy 43*, 1 – 46.

Zidar, O. M. (2015, March). Tax cuts for whom? heterogeneous effects of income tax changes on growth and employment. Working Paper 21035, National Bureau of Economic Research.

Appendices

A Data Description

When not otherwise mentioned, the data source is the Office for National Statistics, the national statistical institute of the UK.

- **Population** is UK total population, code EBAQ.
- **Output** is the real GDP, code ABMI.
- **Consumption** is the final household consumption expenditure, code ABJR.
- **Investment** is the gross fixed capital formation, code NPQT.
- **Government Spending** is the central government total managed expenditure calculated as the sum of Government Consumption Expenditures and Gross Investment, divided by GDP deflator, codes (ANLP+ANNS+NSRN)/YBGB.
- **Unemployment rate** ONS unemployment rate (Age 16 and over) Claimant count and ILO measure, code MGSX.
- **Nominal Interest Rate** is the three-month treasury bill, source Federal Reserve Economic Data.
- **NER** is the effective nominal exchange rate, source BIS (code NNGB).
- **Imports** are the total imports in goods and services, code IKBL.
- **Exports** are the total exports in goods and services, code IKBK.
- **GDP deflator** is the implicit price deflator for GDP, code YBGB.
- **RPI** is the retail price index, code CHAW (for data from 1987Q1). The series from 1987Q1 backwards are extrapolated based on the series of percentage change in RPI, code CZBH.
- **Households Consumption Expenditure Deflator** is the implicit deflator for final consumption expenditure by households and NPISH, code YBFS.
- **Taxes on Income** are the receipts of our income tax categories, code NMCU, and National Insurance Contribution, code AIIH.
- **Total Consumption Tax Receipts** is the sum of VAT (code NZGF), Tobacco duties (code GTAO), Fuel duties (code CUDG), Vehicle duties, i.e. road tax (code CDDZ and EKED), Alcohol duties available from 1993 (code MF6V), (the following data are taken from HM Revenue and Customs) Air duties (APD) from 1995, Betting and Gaming duties from 1986, Insurance Premium tax IPT, from 1995.

- **Income Tax Base** is the sum of total compensation of employees (code DTWM), gross operating surplus (code CGBZ) and mixed (self-employed) income (code CGBX).

Table A.1 – Inside Income and Consumption Taxes

Group	Sub-category
Income	1. Self assessed income tax.
	2. Capital gains tax.
	3. National Insurance Contributions.
	4. Pay As You Earn (PAYE) Income Tax.
	5. Other personal income taxes.
	6. Corporation tax.
	7. Petroleum revenue tax.
	8. Miscellaneous.
Consumption	1. Value Added Tax (VAT).
	2. Fuel Duties.
	3. Alcohol Duties.
	4. Tobacco Duties.
	5. Other Duties.

Notes: The category 'Other personal income taxes' mainly consists of repayments and those tax credits recorded as negative taxes plus company IT and TDSI (tax deduction scheme for interest). The category 'Other duties' includes: Vehicle excise duties (VED), taxes on betting, gaming, lottery, Camelot payments to National Lottery (from 1986 onwards), air passenger duty (from 1995 onwards), insurance premium tax (from 1995 onwards).

Table A.2 – Classification of different tax changes, Cloyne (2012, 2013) and Nguyen, Onnis, and Rossi (2017).

Group	Sub-category	Explanation and examples
Endogenous	1. Demand Management (**DM**)	1.a Targeting the aggregate level of demand e.g. to boost investment, consumption, growth, or curb inflation. 1.b Specific help to households and individuals by stimulating disposable income. 1.c Dealing with a balance of payments crisis via demand.
	2. Supply Stimulus (**SS**)	2.a Certain help for businesses during a downturn (e.g. NIC cut). 2.b Short term sector support (e.g. targeted tax cuts for a sector).
	3. Deficit Reduction (**DR**)	3.a Direct measures to deal with a budget or external deficit.
	4. Spending Driven (**SD**)	4.a Taxes which fund specific spending commitments contemporaneously caused.
Exogenous	1. Long-Run performance (**LR**)	1.a Measures to improve competitiveness, productivity, efficiency and long-run growth (but not taken to offset a shock). 1.b Simplification and deregulation measures. 1.c Long-term support for business or sectors of the economy.
	2. Ideological (**IL**)	2.a Long-term social or political goals, independent of their effect on performance and not to offset current shocks. 2.b Some anti-avoidance measures (where no other motive is given).
	3. External (**ET**)	3.a Court rulings and enforcement of directives.
	4. Deficit Consolidation (**DC**)	4.a Measures to lower inherited deficit for reasons of economic philosophy or to offset current actions in the future. 4.b Does not include actions forced on the government, or decisions contemporaneous motivated by a current shock.

B Properties of the Narrative Series

The 'exogeneity' of our narrative tax changes is a crucial assumption for the validity of our identification strategy. As shown in the main paper, we found no evidence that lagged macro variables Granger cause our narrative measure of tax changes $\Delta\tau_t^{i,n}$ and $\Delta\tau_t^{c,n}$. As we discuss here, this finding relies crucially on excluding from our narrative account, the tax changes announced in 1979:III and in 1988:II that we, as Cloyne (2012), classify as exogenous.

In the 1979 Budget (FSBR, 12 June 1979), there were major tax measures in both consumption and income taxes, aimed to switch part of the tax burden from earnings to spending. In the 1988 Budget (FSBR, 15 March 1988), there were changes in income taxes, mainly a reduction of basic rate and the abolition of higher rates of income tax. It is important to stress that we cannot precisely pinpoint the reason why these tax changes where predicted by lagged macro variables. However, given that these measures where introduced in the Budgets right after two general elections (3 May 1979 and 11 June 1987), it is likely that they were announced in electoral manifestos and thus expected by economic agents.[18]

In order to identify the incriminated policy measures, we proceed as follows. We run a set of Granger causality tests on each macro variable (GDP, consumption, investments and government spending) and jointly together, by excluding one tax change at a time from our narrative series. For the narrative series of consumption taxes, i.e. $\Delta\tau_t^{c,n}$, the tests reject non-causality in all cases but when the 1979:III tax changes are excluded. We therefore eliminate this consumption tax change from the series and, for consistency, we also exclude the exogenous income tax changes happening in the same budget. We repeat the same exercise for $\Delta\tau_t^{i,n}$ and found that the tests reject non-causality in all cases but when the 1988:II tax changes are excluded.

[18]On the 12 of June 1979, during the presentation of the annual budget, the Chancellor of the Exchequer (Sir Geoffrey Howe) announced: 'We made it clear in our *manifesto* that we intended to switch some of the tax burden from taxes on earnings to taxes on spending. This is the only way that we can restore incentives and make it more worth while to work and, at the same time, increase the freedom of choice of the individual. We must make a start now', HC Deb 12 June 1979 vol 968 cc249-250. Similarly, on the 15 of March 1988, during the presentation of the annual budget, the Chancellor of the Exchequer (Sir Nigel Lawson) announced: 'In our general *election manifesto* last year, we committed ourselves to reducing the basic rate of income tax to 25 pence in the pound as soon as it was prudent to do so. This pledge followed a reduction of twopence in the pound to 27 pence in last year's Budget.' HC Deb 15 March 1988 vol 129 cc1006-13.

Table B.1 – Granger Non-Causality Tests for Narrative Consumption Tax Series

	With 1979:III		**Without 1979:III**	
	Test statistic	p-value	Test statistic	p-value
$\Delta ln(Y_t)$	10.74	0.03	0.32	0.99
$\Delta ln(C_t)$	26.80	0.00	0.91	0.92
$\Delta ln(I_t)$	2.98	0.56	0.93	0.92
$\Delta ln(G_t)$	4.72	0.32	7.18	0.13
All	57.60	0.00	10.06	0.86

Notes: The table shows the test statistic and its corresponding p-value of the test that the variable in the first column does not Granger cause narrative income tax series.

For the sake of brevity, Table B.1 and Table B.2 present the results from the Granger tests on each narrative variable with and without the mentioned policy measures. As required by the test, we use the first (log) difference for non-stationary variables. The test rejects the null hypothesis that lagged macro variables do not Granger cause our narrative measure of consumption tax changes when the 1979:III episode is included (Table B.1). In particular the consumption tax changes can be predicted by GDP ($p = 0.03$), private consumption ($p \approx 0$) and when we consider all macro variables together ($p \approx 0$). On the other hand, when we exclude the tax changes in 1979:III, the same test cannot reject the null on each macro variable and when they are tested together. Similarly, the test rejects the null hypothesis that lagged macro variables do not Granger cause our narrative measure of income tax changes when the 1988:II episode is included (Table B.2). In this case our narrative account of income tax changes can be predicted by past investments ($p = 0.04$). Once this episode is removed, we find no evidence that lagged macro variables Granger cause our narrative measure of income tax changes.

Table B.2 – Granger Non-Causality Tests for Narrative Income Tax Series

	With 1988:II		**Without 1988:II**	
	Test statistic	p-value	Test statistic	p-value
$\Delta ln(Y_t)$	2.20	0.70	5.31	0.26
$\Delta ln(C_t)$	0.83	0.94	1.36	0.85
$\Delta ln(I_t)$	10.28	0.04	3.06	0.55
$\Delta ln(G_t)$	5.69	0.22	0.88	0.93
All	21.16	0.17	10.83	0.82

Notes: The table shows the test statistic and its corresponding p-value of the test that the variable in the first column does not Granger cause narrative income tax series.

In addition, we find that the eigenvalues of Ξ are [0.08, 0.51] when the policy measures at 1979:III and 1988:II are included in our narrative account. These numbers are substantially smaller than the corresponding ones of the benchmark specification. When we substitute the narrative at 1979:III and 1988:II by their associated residuals obtained from regressing the narrative measures on the above lagged macro variables, the eigenvalues of Ξ improve to [0.12, 0.58], but they are still substantially smaller than those in the benchmark.

To sum up, we do not include the tax measures legislated at 1979:III and 1988:II into our narrative series in order to comply with 'exogeneity' condition, necessary for our identification strategy. In doing so, we also improve the reliability of the proxy and the correlation between our narrative measures and the latent tax shocks.

C Robustness

C.1 Different Tax Ordering

We start this section by providing robustness regarding the tax ordering. As described in the main text, we find that the correlation between the cyclically adjusted tax rate innovations is -0.06 with 95 percent confidence bounds [-0.89, 0.67], which implies that the responses to a specific tax shock are similar independently of the ordering of tax rates within our 'Proxy-SVAR'. For the sake of completeness, we report the IRFs of a one percent cut of AIT and ACT when each tax shock is ordered first. Results from this exercise are reported in Figure C.1 and Figure C.2. The only noticeable difference we find is that the confidence bounds of the response of private consumption to an AIT shock are larger than in the benchmark case, but we retain the 95 percent significance. We therefore conclude that the benchmark results are robust to tax ordering.

Figure C.1 – Benchmark specification, alternative ordering: 1% cut in Average Income Taxation. Dark and light shaded areas represent the 90 and 95 percent confidence intervals, respectively.

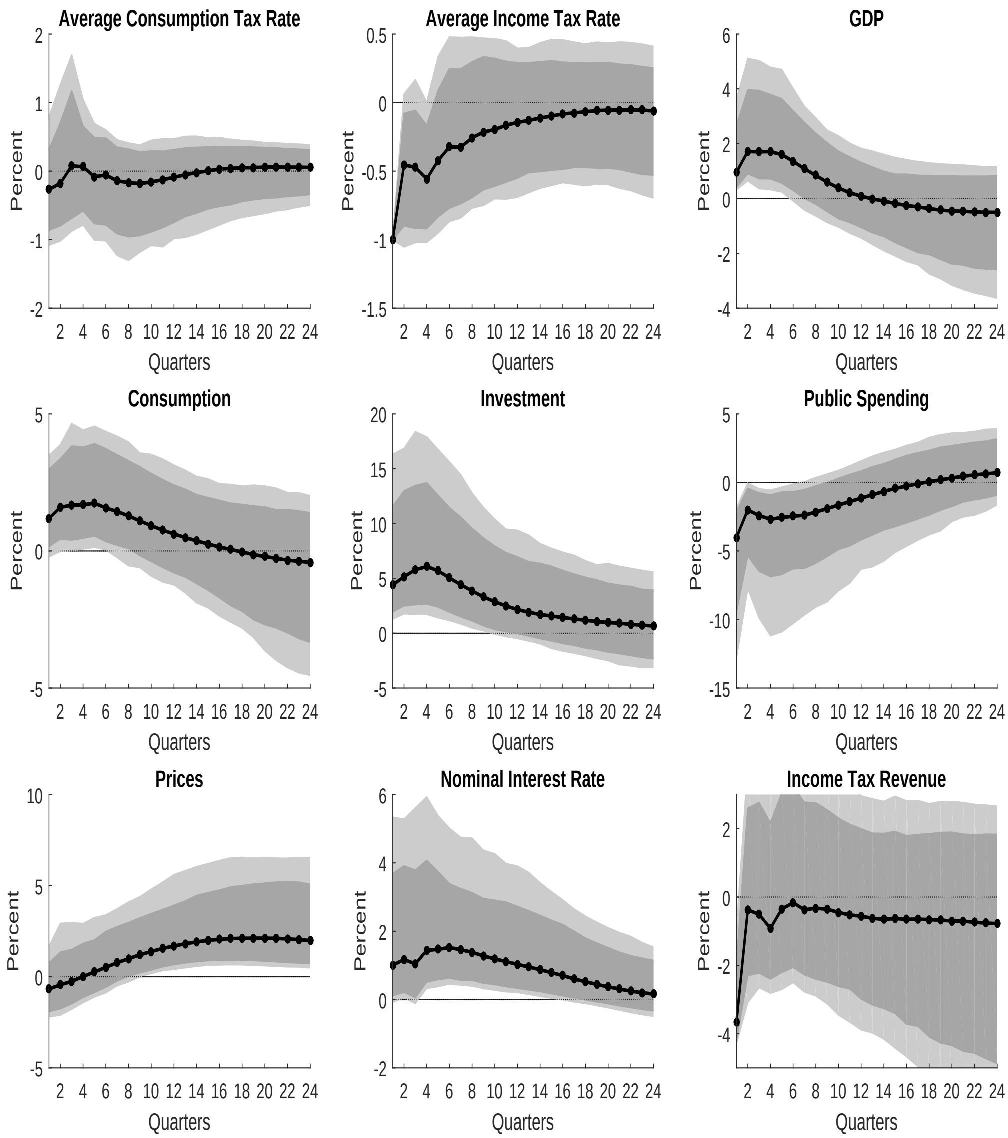
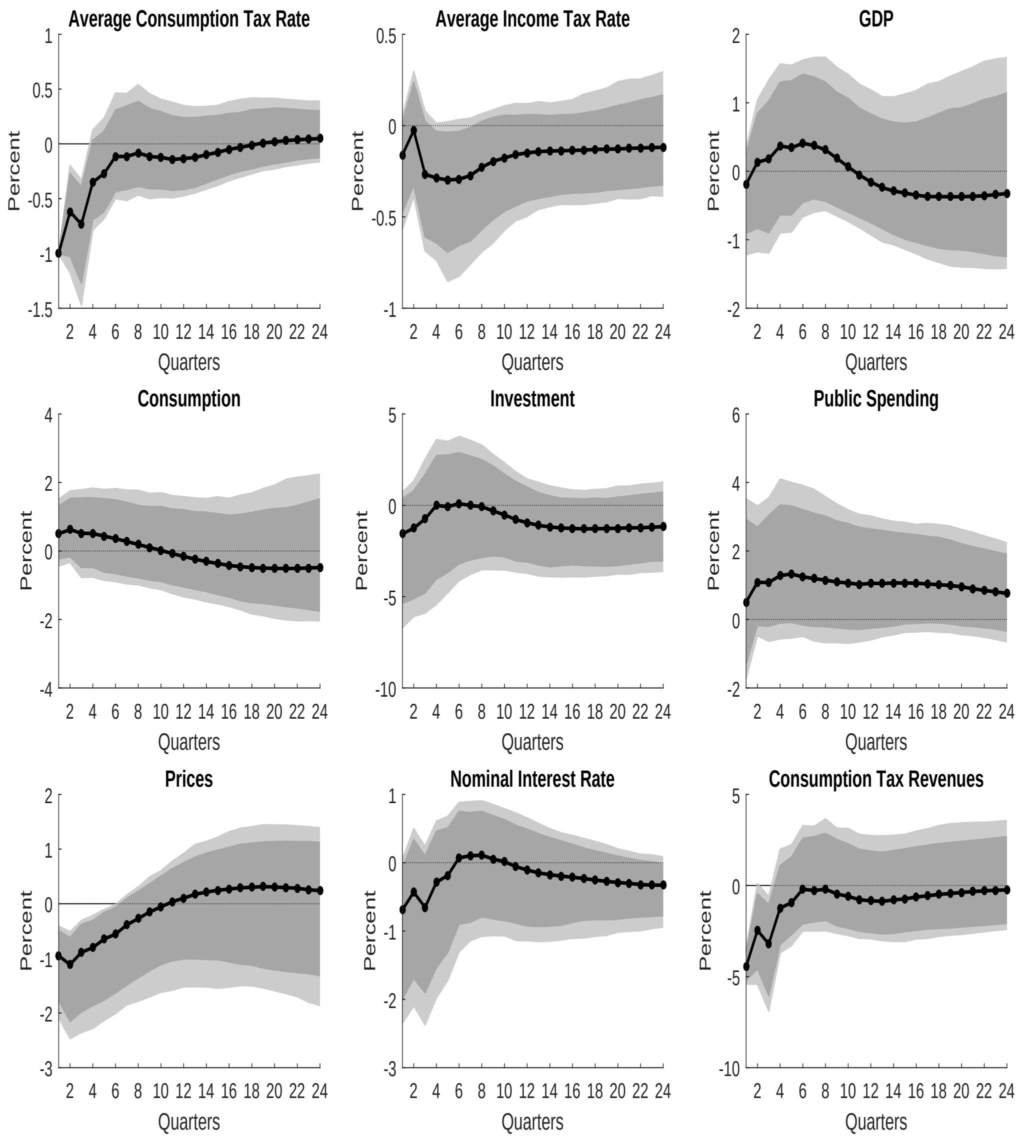


Figure C.2 – Benchmark specification, alternative ordering: 1% Cut in Average Consumption Taxation. Dark and light shaded areas represent the 90 and 95 percent confidence intervals, respectively.



C.2 Trend and Growth Rates

Figure C.3 presents the results from adding a linear-quadratic time trend to the benchmark specification. Including a deterministic time trend does not affect the response on impact of GDP to a income tax shock, but it yields to slightly smaller peak effect, which is reduced to around 1.3. At the same time, the linear-quadratic time trend improves the relative performance of consumption tax shock on GDP, which becomes almost significant at 10 percent for one period, two years after the shock occurs. In this case the reliability test gives values that are somewhat lower than in the benchmark, i.e. the eigenvalues of Ξ are [0.21, 0.58].

Figure C.4 presents the estimates from using stationary data in (11), i.e. first differences of GDP and prices and ratios of consumption, investment and government spending to total output. Estimates show a larger output response to an ACT cut at all horizons although the response remains statistically not different than zero, and slightly lower response to an AIT cut. The reliability test gives values that are slightly smaller than the benchmark one, i.e. the eigenvalues of Ξ are [0.24, 0.53].

Figure C.3 – Benchmark specification with linear-quadratic trend. Broken lines represent the 90 and 95 percent confidence intervals, respectively.

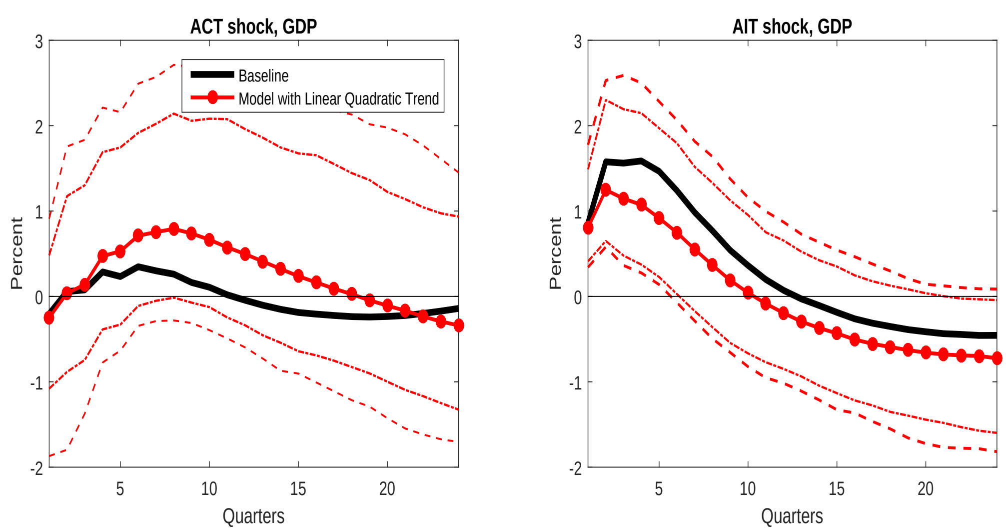
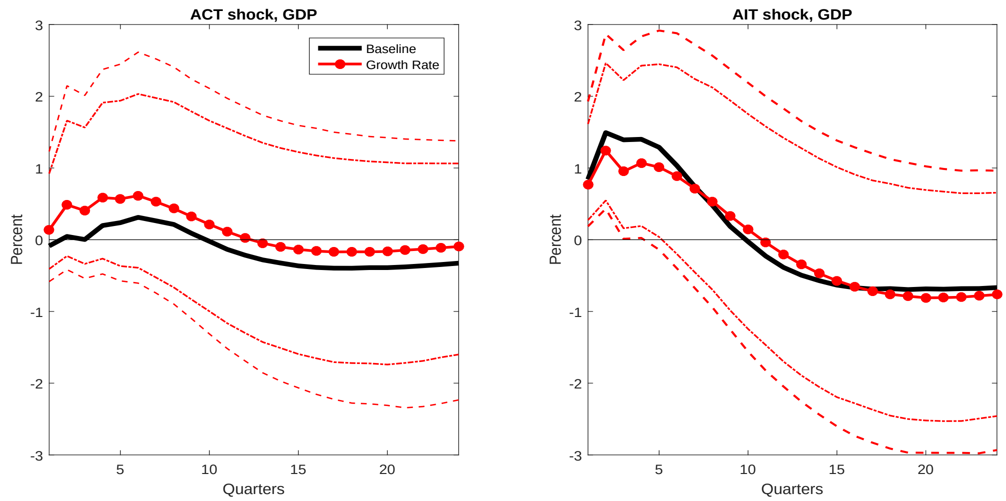


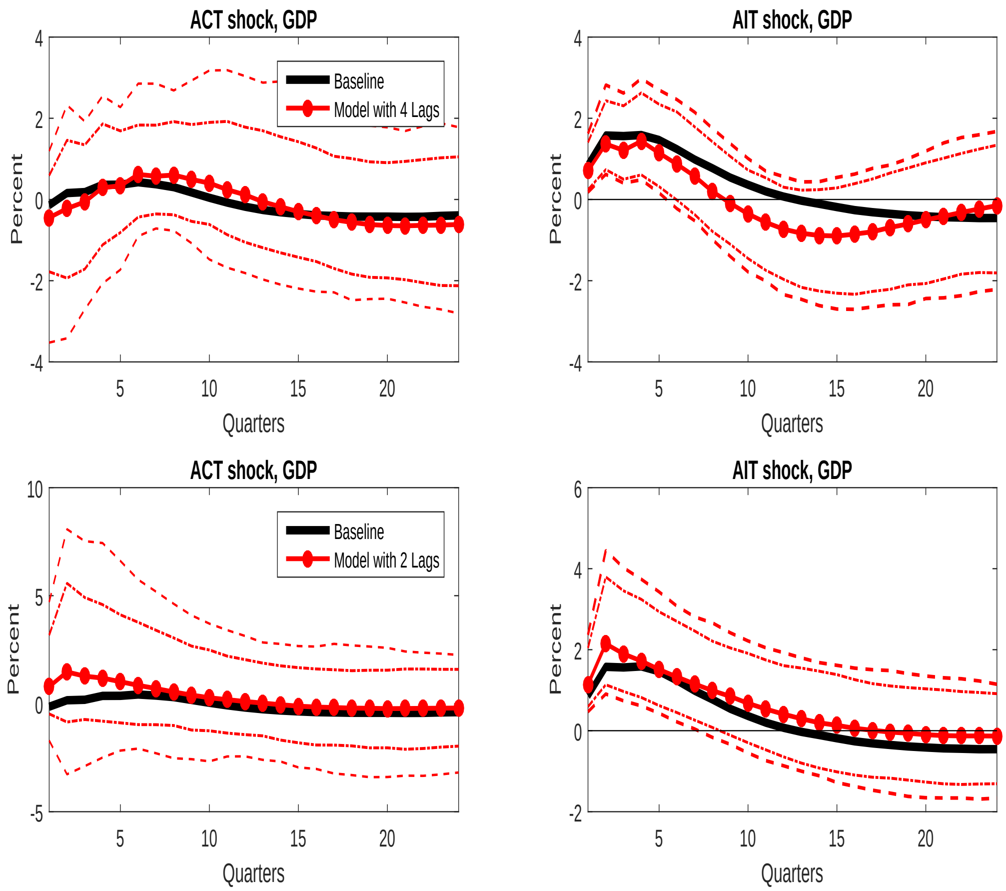
Figure C.4 – Benchmark specification with growth rates and stationary data. Broken lines represent the 90 and 95 percent confidence intervals, respectively.



C.3 Different Lag Structures

We experiment with different lag structures of our baseline model (11) and report the results in Figure C.5. The main results of the paper are robust to different lag structures, although the confidence bounds on ACT shocks get wider. Here we report the results on GDP of imposing 4 and 2 lags, respectively. For the model with 4 lags, we find that the eigenvalues of Ξ are now [0.30, 0.59]. For the model with 2 lags, we find that the eigenvalues of Ξ are [0.21 0.61].

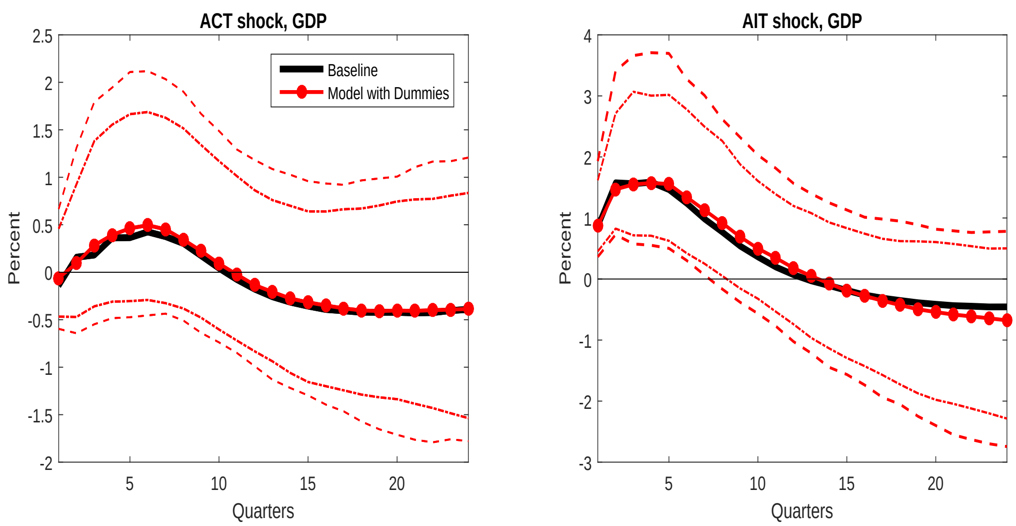
Figure C.5 – Benchmark specification with four lags (Top Panel) and two lags (Bottom Panel). Broken lines represent the 90 and 95 percent confidence intervals, respectively.



C.4 Dummy Variables

In 1979, the UK economy experienced a period of macroeconomic turbulence and large variation in policy actions. The variance of quarter-on-quarter (QoQ) GDP growth during this year was the largest of the sample. For example, the QoQ GDP growth was more than 4 percent in 1979:II, followed by a substantial fall of 2 percent in 1979:III. For this reason, a number of studies, e.g Engle and Hendry (1993) and Hendry and Mizon (1993) among others, suggest to include dummies for 1979:II and 1979:III as regressors. As a robustness check we do the same, and expand the benchmark (11) with two dummies for these two quarters. We find that both the BIC and the AIC tests favour the model with dummies. Furthermore the Wald test rejects the null hypothesis that the coefficients of these dummies are zero at 1 percent significance level. With this new specification, the reliability of our proxies increase to [0.26, 0.63]. Figure C.6 present the responses of output to income and consumption tax shocks in the model with dummies and the benchmark for comparison. The two responses are virtually identical, thus confirming the main results of the paper.

Figure C.6 – Benchmark specification with dummy variables in 1979 and 1988. Broken lines represent the 90 and 95 percent confidence intervals, respectively.

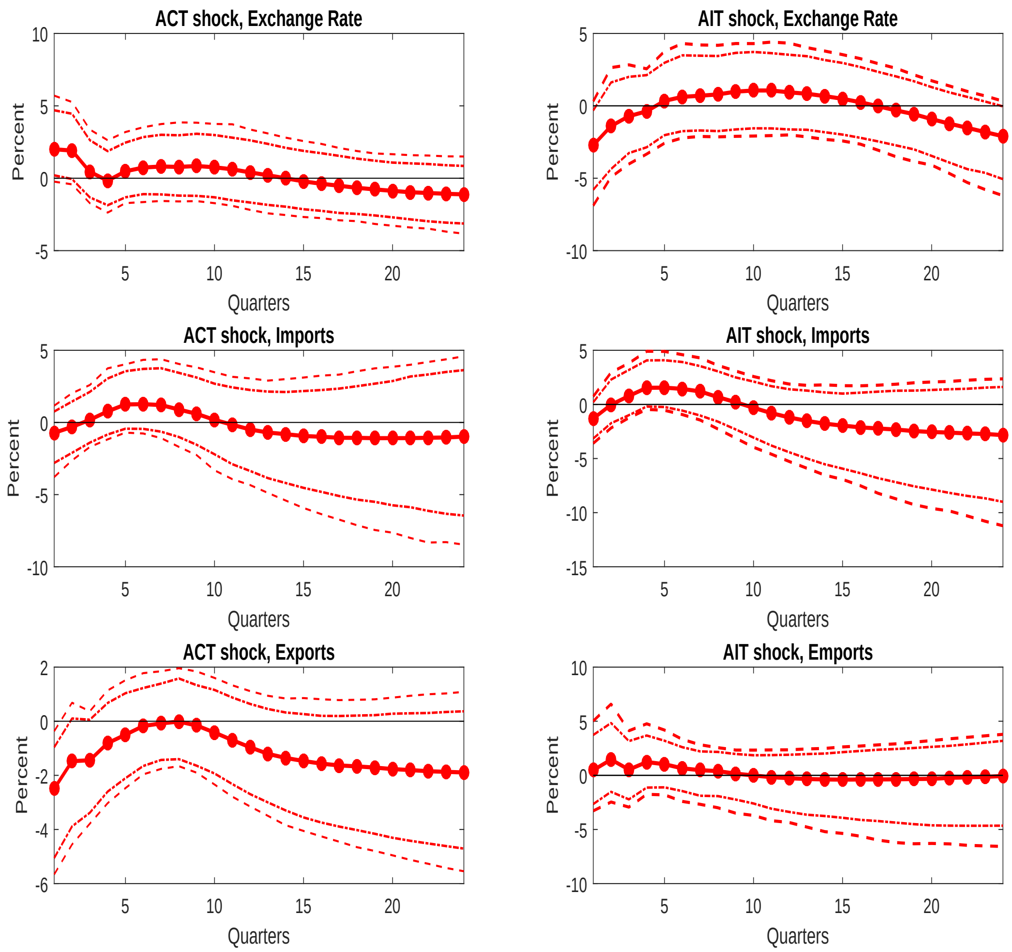


C.5 The Open Economy Dimension of Tax Changes

Here we provide evidence over the international dimension of fiscal policy. This is important as the UK can be considered a small open economy. In particular we add the effective nominal exchange rate, imports and exports to the set of observables in (11). We add these variables one at a time in order to preserve as many degrees of freedom as possible. Results from these expanded econometric models are reported in Figure C.7. A few things are worth noticing. First, the effects of both tax shocks on the exchange rate is weak. We register an appreciation on impact for an ACT shock at 90 percent significance and a depreciation on impact for an AIT shock, always at 90 percent significance. Secondly, neither of the tax shocks has a significant effect on imports, although the response of this variable to an AIT shock appears larger. Thirdly, we find a short but significant at 95 percent, negative effect on impact of an ACT shock on export, while this variable seems to not respond to AIT shocks. Not surprisingly, expanding the number of observables in our SVAR increases the reliability of the proxies on the latent tax shocks.[19]

[19] The eigenvalues of the Ξ statistic are: [0.3 0.59] for model with NER, [0.28 0.64] for the model with imports and [0.29 0.60] for the model with exports.

Figure C.7 – Adding Nominal Exchange Rate (Top Panel), Imports (Middle Panel) and Exports (Bottom Panel) to the Baseline. Broken lines represent the 90 and 95 percent confidence intervals, respectively.



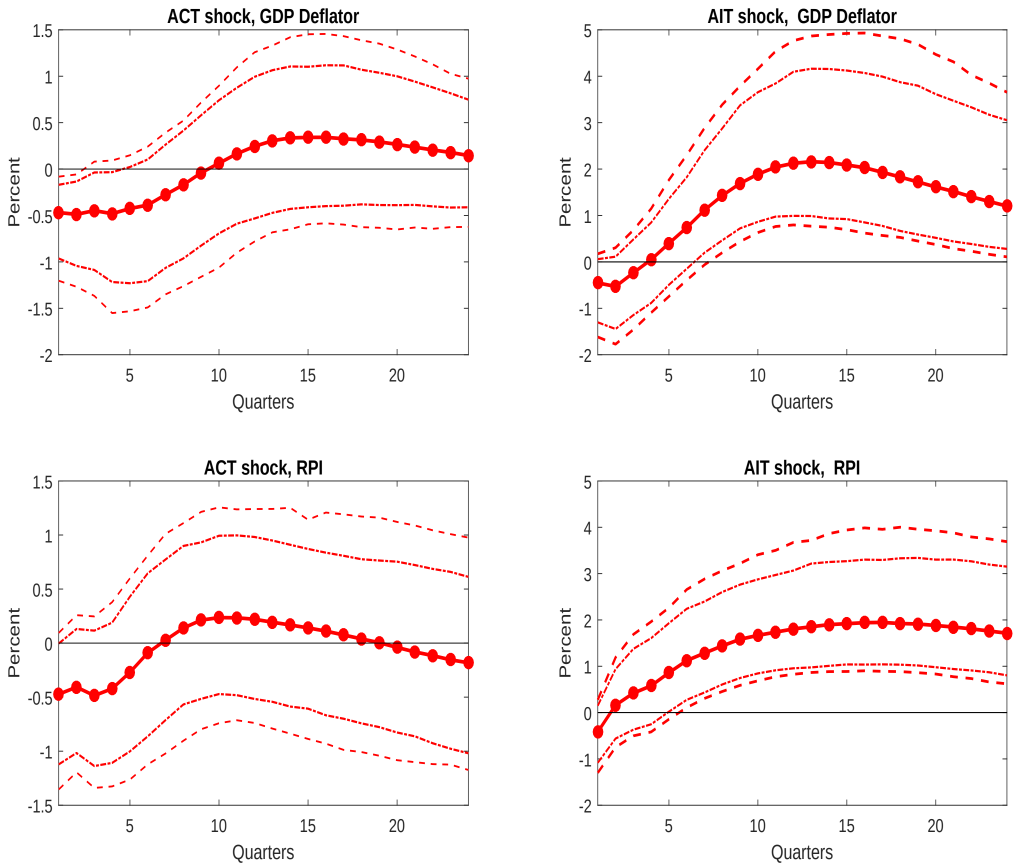
C.6 Different Definitions of Price Level

We estimate the effects of AIT and ACT on the price level when using different measures of the price level in (11). Ideally one would like to use the Consumer Price Index (CPI), cleared by consumption and other indirect taxes. However an official measure of CPI is not available for the entire sample and the only price level where indirect taxes are excluded, the RPIY, is available only from 1987. For this reason we adopt the final consumption expenditure by households and NPISH deflator as benchmark, whose cyclical component correlates the most with CPI and is consistent with the measure used in Mertens and Ravn (2013). Here we test the transmission mechanism of tax shocks to different definition of prices. In particular we use the GDP deflator and the Retail Price Index (RPI). Results are reported in Figure C.8.

The short run effects of tax shocks on GDP Deflator, a broader measure of inflation, is very similar to the response of Consumer Price Deflator index reported in the main paper. A consumption tax shock causes a decrease of the price level on impact. This is mainly due to accounting issues, as the GDP deflator contains indirect taxes. An income tax shock leads to an increase of prices that is both qualitatively and quantitatively very similar to the benchmark. The eigenvalues of the Ξ statistics are sensibly smaller than the benchmark, i.e. [0.21, 0.51], thus revealing a lower reliability of the Proxy.

Finally, we find that the decrease in prices following a consumption tax cut is no longer significant. As in the benchmark, we still find that an income tax shock generates a persistent increase in the price level. Of the three measure of prices, the RPI is the one that leads to the lowest reliability of the proxy, i.e. the eigenvalues of Ξ reduces to [0.20, 0.49].

Figure C.8 – Baseline with GDP Deflator (Top Panel) and Retail Price Index (Bottom Panel) as price index. Broken lines represent the 90 and 95 percent confidence intervals, respectively.



C.7 Alternative Identification

Here we provide the estimation of income and consumption tax shocks under an alternative econometric specification, where we use an 'univariate model', that is popular in the literature on empirical fiscal policy. In particular, we adopt a single equation approach similar to Mertens (2015), where we regress the first differences of output $\Delta ln(Y_{t+h})$ on the changes of each tax category, i.e.

$$\Delta ln(Y_{t+h}) = \beta_c \Delta ln(ACT_t) + \epsilon_{c,t+h}, \tag{19}$$

for consumption tax changes and

$$\Delta ln(Y_{t+h}) = \beta_i \Delta ln(AIT_t) + \epsilon_{i,t+h}, \tag{20}$$

for income tax changes. In order to deal with the endogeneity of tax rates, we use the tax shocks identified in the benchmark 'Proxy-SVAR' as instruments. The first stage F-statistics for $\Delta ln(ACT_t)$ and $\Delta ln(AIT_t)$ are 30.78 and 20.22, respectively. These values suggest that we are using good instruments and indirectly test the relevance of our narrative proxies in the baseline SVAR. To derive the impulse response functions, we implement the local projection method proposed by Jordà (2005) and present the results in Figure C.9. As it can be seen, the magnitude and dynamics are very similar to the baseline.

Figure C.9 – Local projection method. Broken lines represent the 90 and 95 percent confidence intervals, respectively.