Taxes and Growth: New Narrative Evidence from Interwar Britain*

James Cloyne Nicholas Dimsdale Natacha Postel-Vinay

August 2023

Abstract

The impact of fiscal policy on economic activity is still a matter of great debate. And, ever since Keynes first commented on it, interwar Britain, 1918-1939, has remained a particularly interesting and contentious case — not least because of its high debt environment and turbulent business cycle. This debate has often focused on the effects of government spending, but little is known about the effects of tax changes. In fact, a number of tax reforms in the period focused on long-term and social objectives, often reflecting the personality of British Chancellors. Based on extensive historiographical research, we apply a narrative approach to the interwar period in Britain and isolate a new series of exogenous tax changes. We find that tax changes have a sizable effect on GDP, with multipliers exceeding 2 within two years. Our estimates provide new evidence on the effects of tax changes, contribute to the historical debate about fiscal policy in the interwar period and are also consistent with the sizable tax multipliers found after WWII.

JEL classification: E23, E32, E62, H2, H30, N1, N44

Keywords: Macroeconomic Policy, Fiscal Policy, Taxation, Public Finance, Fiscal History, Multiplier, Narrative Approach

*Address for correspondence: James Cloyne (University of California Davis, NBER and CEPR); Nicholas Dimsdale (University of Oxford); Natacha Postel-Vinay (London School of Economics and CEPR). We thank Hanguo Huang for excellent research assistance. We are grateful for comments and advice from Benjamin Born, Brad De Long, Barry Eichengreen, Jason Lennard, Chris Meissner, Eric Monnet, Emi Nakamura, Albrecht Ritschl, Christina Romer, David Romer, Jon Steinsson, Alan Taylor, Ryland Thomas, Jim Tomlinson, Harald Uhlig, the editor and the referees. We also thank numerous seminar and conference participants for helpful comments and suggestions. We would also like to thank Jason Lennard for kindly sharing his new fiscal data. All errors remain our own.

1 Introduction

What is the impact of changes in fiscal policy on economic activity? Will tax cuts stimulate the economy? Will austerity hurt? Despite nearly 100 years of debate, the efficacy of fiscal policy remains a particularly contentious issue and there is still much uncertainty about the empirical effects. We use the events and circumstances of Britain between the two world wars to shed light on an important aspect of this debate: how do tax changes affect the macroeconomy? This remains a first-order policy question. Many countries have cut taxes significantly in response to recent events. And, with public debt currently at unprecedented levels, others are, once again, debating whether or when to raise taxes.

One reason why the macroeconomic effects of fiscal policy are hotly debated is that natural experiments in macroeconomics are rare and causal inference from observational data is challenging. While fiscal policy may affect the economy, the reverse is also true, making it hard to interpret any correlation between the two. A careful examination of specific historical episodes can, however, provide us with interesting changes in policy that are better suited for this task. We show that interwar Britain is one such episode.

Interwar Britain provides an interesting laboratory for studying the effects of tax changes. In fact, this period has long stood out as a particularly contested case in the debate about the effects of fiscal changes on the economy. In the 1920s and 1930s, Britain stuck to a policy of austerity when other countries did not. Furthermore, with its high overhanging debt and sluggish economic performance, interwar Britain resembles many economies today. This period therefore remains highly relevant.

Drawing on dozens of British Budget documents and political speeches, we construct a new dataset of tax changes based on 22 budgets, covering over 200 individual tax adjustments from 1918 to 1938. We show that interwar Britain is particularly well-suited to this task. As is still the case today, tax changes were clearly announced in the government's spring Budget each year. Thus, by using each Budget's *Financial Statement* and related official publications, we are able to construct an extensive dataset of individual tax policy changes during the 1920s and 1930s.

We apply a narrative approach to isolate "exogenous" movements in taxes. Specifically, we follow Romer and Romer (2010) (who examine the US after WWII) and use historical ev-

idence in order to examine policymakers' motivations behind each Budget. The econometric objective is to isolate policy reforms that were directly influenced by current or prospective fluctuations in the economy from those that can be seen as more independent, and use the latter to estimate the impact of fiscal policy on the economy. The reasons for these decisions were detailed in the annual Chancellor's Budget speech, which provides rich information about the nature of the policy changes as well as the Chancellor's motives. Fiscal policy over this period looks quite different to the post-WWII era and, in addition to the Budget documents and speeches, we also make extensive use of the historiography surrounding Budget decisions to understand the circumstances surrounding each Budget and the priorities and personality of individual Chancellors.[1]

Based on our detailed historical analysis, we show that British fiscal policy was distinctly pre-Keynesian. Macroeconomic stabilisation policy was largely absent and fiscal instruments were generally not used for such purposes. Tax changes tended to flow from longer-term political (following, for instance, an ideal of fairness) or economic goals (e.g. supply-side measures aimed at increasing productivity), and often reflected the personalities of particular politicians. And despite the generally austere character of fiscal policy at the time, there were surprisingly numerous tax rises and cuts we can use for identification. As evidence of the success of our econometric strategy, we show that our "exogenous" tax changes are not Granger-caused by economic conditions, but the "endogenous" group are indeed highly predictable. A background companion paper (Cloyne et al., 2023) provides the detailed year-by-year history of interwar U.K. tax policy, the narrative analysis and the new data themselves. We hope this rich historical analysis is a further unique contribution of this paper that should have broad interest for economists and historians alike.

Using our narrative identified tax changes, our main empirical result is that tax "multipliers" can be sizable. Broadly speaking, the multiplier tells us how many additional units of GDP are gained by a one unit fiscal stimulus.[2] We find that tax multipliers are around 2-3 after one year. These effects are consistent with narrative-based estimates for the post-WWII period, which also tend to find tax multipliers larger than 1. Our findings are also remarkably

[1] (Alesina, 1988, Daunton, 2007, Matthew, 1986, Middleton, 1985, 1996, Moggridge, 1972, Peden, 1987, Pollard, 1970, Short, 1985, Tomlinson, 1990, for instance).

[2] We will define this more precisely later. Although often attributed to Keynes, Henderson and Keynes (1929) did not explicitly refer to a "multiplier." The concept was worked out for employment by Kahn (1931), and was first taken up by Keynes (1933) and Keynes (1936).

stable across a range of specifications and robustness checks. In addition, we show that tax cuts reduce unemployment and lead to an increase in interest rates. The evidence for an effect on prices is relatively weak. Although the implied tax multiplier estimates are sometimes lower than Keynes suggested in his *Means to Prosperity* (Keynes (1933)) (for expenditure), they still suggest that greater stimuli to the economy may have been achievable by cutting taxes.

What historical aspects of U.K. fiscal policy are particularly useful for identifying these effects? The fact that, as a rule, neither spending nor taxes were used for stabilisation purposes is a very useful point of departure. Although some Budgets were clearly the product of current economic developments, as noted above many policy reforms were related to longer-term goals. For example, by documenting Chancellors' motives, we are able to show, for instance, that Chancellor Philip Snowden's 1924 tax cuts were partly aimed at reducing the tax burden on the poor, following his long-standing will to make society less unequal. Likewise, Imperial Preference was a trade-related principle favoured generation after generation in the Chamberlain family, and thus put forward by both Austen Chamberlain in the 1920s and Neville Chamberlain in the 1930s.

While policymakers did not generally use fiscal policy counter-cyclically and many reforms clearly followed from long-run objectives, much attention has been given to the high overhanging debt over this period. The repayment of the debt and a focus on tight fiscal policy therefore still features heavily throughout the interwar Budgets. As a result, one might be worried that periods of poor growth and high fiscal stress might have systematically triggered further fiscal tightening. Tax changes might also have been related to government spending. But, again, institutional aspects of the fiscal policy regime over this period are particularly useful for our exercise. The strength of the narrative method is that we can therefore carefully examine and address these additional hypotheses.

We find that debt repayment would not generally elicit large new changes in taxation. One institutional feature of this period is that Chancellors sought to make regular — indeed almost mechanical — payments into a sinking fund. Such sinking funds tended to represent long-term commitments that did not respond to short-term fluctuations in the economy. Year-to-year tax policy was generally separated from this consolidation plan. This steady approach to fiscal consolidation was deeply rooted in the budgetary orthodoxy followed by

the British Treasury since the Victorian era. Still, the strong emphasis on strict budget policy meant that policymakers were particularly sensitive to periods where poor economic conditions led to alarming, contemporaneous movements in the deficit. In such cases, taxes were usually increased and these are good examples of tax changes that we do classify as endogenous actions. On the spending side, we use our narrative evidence to exclude tax changes directly associated with changes in government spending.[3]

Our paper relates to at least two sets of existing work. First, we connect to the macroeconomics literature studying the effects of fiscal policy. The narrative methodology itself has been applied, utilized or extended by a number of authors for the post-WWII era to examine the effects of changes in taxes (for example Cloyne (2013), Mertens and Ravn (2013), Guajardo et al. (2014), Hayo and Uhl (2014), Cloyne and Surico (2017), Gunter et al. (2021), Barro and Redlick (2011), Mertens and Ravn (2012, 2013, 2014), Nguyen et al. (2021) and Hussain and Liu (2018)). These papers tend to find large effects, but the literature on the macroeconomic effects of tax changes still remains relatively small and limited to post-WWII episodes.[4] There also remains much disagreement about the macroeconomic effects of tax changes. New narrative evidence is therefore important but also challenging to generate because it requires historical periods with sufficient variation in tax policy together with extensive historical analysis to isolate reforms that can be regarded as "exogenous". Our paper hopefully therefore provides an important addition to this literature, and provides an extensive new dataset — for a period where official data already tend to be limited — that can be used by other researchers.

Second, we contribute to the lively debate about the efficacy of fiscal policy in the interwar years. The fiscal stimulus debate originates from this period and the economic conditions of the time have a number of similarities to circumstances today. Much of the existing literature on the interwar period has, however, focused on the role of government spending, and defense

[3]Furthermore, while all empirical fiscal multipliers are conditional on the response of other fiscal policy instruments, we show that the subsequent response of government spending to our identified tax policy shocks is more muted overall and the response of tax revenues is large (as one would expect). Our multiplier estimates are therefore closer to the concept of a deficit-financed tax multiplier than a balanced-budget multiplier.

[4]The larger effects from narrative-based studies contrasts with the smaller effects typically found using older structural VAR approaches such as Blanchard and Perotti (2002). Mountford and Uhlig (2009) propose an alternative SVAR approach using sign-restrictions and find larger multipliers. Some papers have specifically focused on fiscal consolidations and argue that fiscal consolidations might have *positive* effects, although this is less likely for those based on tax increases (Alesina and Ardagna (2010), Alesina and Giavazzi (2012)).

expenditure in particular. Prominent recent contributions are Crafts and Mills (2013, 2015), who employ narrative methods for identifying variation in defense expenditure (see important contributions for the United States such as Ramey (2011) and Ramey and Zubairy (2018)). This newer literature has found that interwar spending multipliers in the United Kingdom are well below Keynes (1933)'s own calculations (who famously advocated a fiscal stimulus to combat rising unemployment during the 1930s).[5] In the postwar years, of course, changes in tax rates became an important part of demand management policy (Dow, 1964, Cairncross and Watts, 1989), despite the fact that Keynes himself was not a supporter of stimulus via taxes.[6] Whether the economy could have been stimulated using tax policy has therefore been largely overlooked in the economic history literature.[7] We hope our results, together with our extensive historiographical analysis over this period, will provide a novel contribution to the debate about this important and remarkably relevant period.

The remainder of the paper is structured as follows. In Section 2, we provide more detail on the empirical approach, the data sources and our narrative strategy for interwar Britain. In Section 3 we provide an overview of the tax changes identified using our narrative approach and explore the properties of our new series of exogenous tax changes. Section 4 presents our baseline results. Section 5 presents a range of further results and robustness checks. Section 6 concludes.

2 Empirical Approach and The New Dataset

2.1 Approach and Identification

Isolating the causal effect of tax changes on macroeconomic activity is notoriously difficult. Policymakers respond to economic conditions and, in turn, tax policy may have economic consequences. The co-movement of taxes and macroeconomic outcomes observed in the data has no clear causal interpretation and disentangling cause and effect is one of the most important challenges to address. In addition, studying the effects of tax changes in the

[5] On the other hand, papers supportive of a sizable spending multiplier include Thomas (1981, 1983), Broadberry (1986), Hatton (1987), Dimsdale and Horsewood (1995), Almunia et al. (2010).

[6] When James Meade suggested to him that changes in taxes would be more appropriate, Keynes firmly disagreed (Dimsdale, 1987, p. 225).

[7] One exception is Romer and Romer (2014) who examine the incentive effects of changes in marginal tax rates on taxable income during the interwar period in the United States.

interwar period poses a further challenge not usually faced by those using post-WWII data. To our knowledge, official fiscal data are not available at a quarterly frequency pre-WWII.

Using detailed historical documents, we first collect an extensive new dataset of all the legislated tax changes in Britain from 1918 to 1938. There were 22 Budgets where new policy changes were announced and legislated over these twenty years: 21 regular annual budgets plus one emergency Budget in September 1931. In early 1932 there was also the emergency Import Duties Act. Overall, we find that over 200 individual taxes were adjusted over this period and this provides a fascinating degree of variation in fiscal policy.

To tackle the issue of causality, we employ a narrative approach following Romer and Romer (2010) for the United States and Cloyne (2013) for the United Kingdom for the period after WWII. The idea behind this strategy is simple: tax changes are not always motivated by changing economic conditions. Although we see clear evidence of tax policy sometimes responding to macroeconomic conditions, many tax decisions in interwar Britain were taken for other reasons, and often reflected the priorities of particular politicians.

To establish a causal relationship, we therefore need to isolate the variation in tax policy that was not responding to economic fluctuations. In other words, we need some "exogenous" variation in tax policy to achieve identification. In isolating what we will refer to as "exogenous tax changes", we follow Romer and Romer (2010) and look for tax reforms that were not responding to changes in current or future expected economic conditions, a form of weak exogeneity. Even if the reform had been influenced by past economic conditions, we can still identify the effects of tax changes if we control for historical movements in macroeconomic variables. In principle, of course, it would be useful for the tax changes to be exogenous with respect to the entire history (past, current and future) of economic shocks (strictly exogenous) and, wherever possible, our goal is to look for reforms that might satisfy this stronger condition as well.[8] Following Romer and Romer (2010), we isolate these exogenous tax changes by examining policymakers' motivations for each tax reform. As we discuss below, interwar Britain provides a range of interesting quasi-natural experiments. Not only is there a considerable number of tax reforms; we believe that a sizable proportion can be regarded as exogenous. As an ex-post validity check, we will show that our "exogenous" shocks are not predictable from past economic conditions, while the "endogenous" changes

[8] Our regressions will, however, include lagged macroeconomic controls as "insurance".

are highly predictable.

2.2 Method and Sources

The main event of the U.K. fiscal year is the annual Budget. Two key elements of the Budget are the annual Financial Statement — which outlines the state of the public finances and announces any new changes in taxation — and the speech of the Chancellor of the Exchequer (the U.K.'s Finance Minister) to the U.K. Parliament. The Budget speech outlines the fiscal situation, the government's fiscal priorities for the coming year and then runs through all the new tax announcements. To construct our new dataset of tax policy changes, our primary source is therefore the U.K. government's *Financial Statements* (Stationary Office, 1918-1938). Not only does the Financial Statement outline all the tax changes, it (usually) reports the precise implementation date and provides an estimate of the projected impact on revenue of the changes.[9] The revenue estimate is computed for each tax change and is based on the assumption of an unchanged tax base. This is useful because the projection can then be seen as the effect of the tax holding all else constant. The revenue estimate is also given for a "Full Year" which is an estimate of the on-going change in tax liabilities associated with the reform. These features of the data side-step common problems with ex-post tax revenue data, which are a function the tax reform, the macroeconomic effect of the tax reform and the other economic fluctuations.

Table 1 provides an example of the tax data available from the April 1920 Budget. There are 17 individual tax changes in this Budget. The reforms are distributed across a range of tax categories and with a range of implementation dates. By collecting this information for all Budgets between 1918 and 1938 we therefore construct a new dataset of all interwar tax reforms.

[9]Sometimes the precise implementation date is given in the Budget speech or the U.K. Finance Act which enacts the Budget's measures. Occasionally we therefore have to cross-reference the Financial Statement with the Budget speech or the relevant Act of Parliament.

Table 1: Tax changes announced by Chancellor Chamberlain, April 1920 Budget

Tax item	Implementation date	Cost/Yield in a Full Year (£m)
Postage	04.08.1920	9.5
Motor spirit	01.01.1921	-3.2
Motor car	01.01.1921	-1.2
Spirits (Customs)	20.04.1920	6.4
Spirits (Excise)	20.04.1920	18.1
Beer (Customs)	20.04.1920	0.02
Beer (Excise)	20.04.1920	29.98
Wine	20.04.1920	4.1
Tobacco	20.04.1920	0.53
Total stamps	04.08.1920	6.3
Income tax (Graduation, differentiation)	06.04.1920	-29.2
Abolition of temporary war reliefs	06.04.1920	3.9
Relief for double income tax	06.04.1920	-2
Super tax	06.04.1920	11
Excess profits duty	01.01.1920	100
Corporations profits	01.01.1920	35
Motor Vehicle Duties	01.01.1921	9

Source: Stationary Office (1918-1938). The final columns shows the projected revenue effect of each policy change in £m. To put these numbers in context, nominal GDP in 1920 was £6,119m. The sum of column three is £198m, around 3% of GDP. A positive value corresponds to a tax increase and a negative value to a tax cut.

The next step is to establish whether the motivations given for the policy changes are "exogenous" or "endogenous". To do this, we first carefully study the Chancellor's speech to the U.K. Parliament (recorded in the Official Parliamentary Record, UK Parliament (1918-1938)). Individual Chancellors in the U.K. have traditionally enjoyed considerable influence over the country's budget policy (Peden, 1987, p. 10), and because of the U.K.'s Parliamentary system, all the measures in the Budget are almost always implemented. Given the importance of the Chancellor in the U.K. political process, their annual or semi-annual announcement of new measures has usually been accompanied by a detailed exposition and defense to a captive audience. The speech is therefore an excellent source for examining the objectives and motivation for the Budget and the rationale for the policy changes. This provides a very useful point of departure and a sizable proportion of our narrative evidence is drawn directly from the Budget speech. However, it is hard to interpret the speeches, and to assign motives, without a fuller understanding of the period and the different influences that might have affected the Chancellor's decisions. We therefore use a range of historical sources to understand both the economic orthodoxy of the time, the key economic developments during the period and the various influences that might have affected each Chancellor. All this research informs the precise categorization system we will explain in more detail below.

Although each Budget usually contained many individual tax changes, it is useful to

regard the Budget as a package of measures. Many Budgets tended to have an overarching goal, and we first seek to establish what those overall objectives were. Historiographical information along with the Budget speech itself help us ascertain those motives. Individual tax changes in each Budget will have different implementation dates and may differ slightly in their specific or secondary objectives, but in general we will classify the Budget as a whole as either "exogenous" or "endogenous". All in all, 19 out of the 22 Budgets easily fit this "package" description. We also treat the Import Duties Act 1932 as a package.

There are, however, three Budgets where we need to treat a small number of tax changes differently from the rest of the package. The first situation arises where the Budget is deemed to be exogenous overall, but there are one or two tax changes that were specifically designed to fund a spending commitment. Although there are some Budgets that are correlated with major spending policies (such as in 1918 or in the late 1930s during rearmament), spending and tax policies were not usually adjusted together in the Budget, which focused more on setting taxes for the coming year. Spending-driven tax changes in our data set occur where a tax adjustment is hypothecated for a specific spending purpose.

Even if the Budget seems exogenous overall, we still classify these tax changes as correlated with spending. Two Budgets — 1923 and 1929 — have some spending-driven changes like this. The second situation arises where there is no clear overarching motive and the Budget is made up of a number of idiosyncratic and sometimes technical reforms. Many of these changes fit the bill as exogenous, although a small number might have been designed to target conditions in parts of the economy. In practice there are only two cases like this: in 1929 (already discussed above) and 1933. We treat these as Budgets with two separate packages of reforms. We will discuss the specific Budgets in more detail below.[10]

2.3 Key Aspects of Interwar Fiscal Policy

Before discussing the precise categorizations used to isolate exogenous tax policy changes, it is helpful to ask: Which features of fiscal policy in the 1920s and 1930s are particularly useful for our purposes and which aspects require special attention? On the one hand, macroeconomic policy was "pre-Keynesian" and the idea that aggregate demand could be steered using tax

[10] As a robustness check, in Appendix A4.8, we classify these three Budgets more crudely according to an overall motive and show that this does not significantly change our results.

policy was not the prevailing orthodoxy. As a result, many Budgets focused on more long-term and social objectives, often reflecting the personality of each Chancellor. These aspects of the policy environment make the interwar period particularly well-suited for a narrative analysis. Many decisions were made for ideological reasons. For instance, many tax reforms reflected support or disdain for Imperial Preference, a protectionist policy.[11] Likewise, some measures were passed with a view of making society less unequal. Others had to do with long-run productivity performance of the economy as a whole, or of certain sectors.

On the other hand, although countercyclical fiscal policy was not the norm — indeed there are almost no countercyclical tax changes in a standard Keynesian sense in our dataset — this is not a sufficient condition for identification. For example, policymakers were very concerned with fiscal deficits and the national debt. In fact, this was a key part of the economic orthodoxy of the period and raises a number of issues to consider. For example, attempts to offset a deterioration in the budget in response to economic fluctuations are still likely to be endogenous for our purposes. A deeper understanding of the prevailing attitudes to fiscal policy is therefore important for understanding the motivations of each Chancellor and the potential sources of endogeneity in this period (Appendix A2 provides the full list of Chancellors and their affiliation).

A key aspect of the interwar fiscal framework is what has often been called Budgetary orthodoxy. In general, a balanced *and* small budget was regarded as desirable in most circumstances except war time. The literature is unanimous in describing Budgetary orthodoxy as a strong, pervasive, and long-standing principle of fiscal policymaking in Britain. Indeed, it is widely agreed that it underlay the British Treasury's fiscal stance for decades since the Victorian era, at least until the rearmament phase in the late 1930s. As a commitment to this policy, the U.K. Treasury, and successive Chancellors, sought to steadily reduce the national debt — which was sizable following the First World War — by isolating the repayments of debt from fluctuations in the economy by, for instance, committing to regular (indeed almost automatic) payments into a sinking fund. Sinking fund payments were regarded as *expenditure* for calculating the budget deficit and, in practice, successive Chancellors therefore consistently engineered a surplus on conventional metrics. Moreover, any realized surplus

[11] Those supporting it favoured protecting British production by raising tariffs on imports from the rest of the world, with some exceptions made for the British Empire. Members of the Chamberlain family (Joseph, Austin and Neville) were particularly strong supporters (Self, 2006).

during the year was required to be put towards further debt redemption. At the same time, governments aimed to keep expenditure and taxes low and fiscal policy was not to be used for countercyclical purposes. Middleton deems this general principle so pervasive in British budget-making history that he refers to it as the minimal balanced budget rule (MBBR) (Middleton, 1996, p. 181). An appreciation of the historical setting in which this principle evolved is therefore essential in understanding its importance over the period and its acceptance across the political spectrum; its "strength and diuturnity" as (Middleton, 1985, p. 84) calls it. The term "rule" here may, however, be somewhat of a misnomer — while it illustrates its ideological power, the MBBR reflects more a long-term objective than anything set into law. In fact, Chancellors did not always seek to balance the budget exactly every year and the low-tax ideal sometimes seemed to conflict with the desirability of budget balance (see Tomlinson (1990, p. 67), Middleton (1985, pp. 83-5) and Middleton (1996, p. 184)). As we will see below, some Chancellors seemed to give relatively more priority to debt repayment, whereas others gave priority to the low tax ideal. Below we examine the reasons for support for each of these priorities in turn.

Low taxation

What was the rationale for lower taxation? The British defense of minimal taxes and expenditure has its origins in the *laissez-faire* view of the economy (starting with Locke (1689) and Hume (1987)) and a critique of eighteenth-century mercantilism, with the idea that government can too easily yield to specific interest groups (Middleton, 1996, p. 53). Adam Smith himself had a deep mistrust of government: "always, and without any exception, the greatest spendthrifts in the society" (Smith, 1884[1976], p.142). Tied to this suspicion of big government was the idea that government spending and taxes crowded out private investment (Middleton, 1996, p. 181). This principle gathered strength, until it came to form the basis of economic policymaking in nineteenth-century Victorian Britain, especially under the Chancellorship of William Gladstone, famous for his judgement that "money should be best left to fructify in the pocket of the people" (Peden, 1987, p. 27) (see also Buchanan (1985)). As Hicks (1953) emphasised: "Gladstonian budgeting is inextricably bound up with the theory of the ever-balanced (or even over-balanced) budget and with a perpetual desire for economy in public outlay."

Chancellor Gladstone, himself inspired by Scottish evangelical thinker Thomas Chalmers,

spread his minimal budget ethos to all areas of government, with long-lasting institutional effects (Matthew, 1986, pp. 57, 73, 112). Spending plans, for instance, could not be allowed to automatically run from year to year, which is when the annual Budget became the centrepiece of the political year (see Daunton (2007, p. 463) and Middleton (1985, p. 85)). In 1866, he passed the Exchequer and Audit Act, which asserted Treasury control over all other departments (Matthew, 1986, pp. 106), thereby imposing his minimal balanced budget to most areas of policymaking.

The First World War, however, led to society's acceptance of a higher threshold of taxation, shifting upwards the "minimal" part of the MBBR. Democratisation, together with the broadening of both taxable income and national insurance coverage, meant that both taxation and public sector expenditure moved to a permanently higher level after WWI, relative to the prewar period (see Peacock and Wiseman (1961) and Middleton (1996, pp. 92, 98)). As Middleton (2010) notes, by the 1920s and 1930s "the budget was no longer minimal". Nevertheless, the Treasury — and many Chancellors — still concentrated its planning on getting back to a so-called "normal year." The normal year was referred to by policymakers as a pre-war economic standard in which taxes and government expenditure were low. With a few exceptions (Horne and Churchill in particular), this ideal would permeate policymaking across the political spectrum throughout the interwar era (Short, 1985, pp. xvii, 27-28). This approach meant that tax policy was often set looking through current conditions and with a more long-term perspective. This explains why various tax reductions can be seen as based on "exogenous" ideological and/or long-run objectives.

The most striking example of the application of the normal year principle is Labour Chancellor Philip Snowden's sweeping tax cuts of 1924. By and large, Keynesian demand management was not adopted until much later and, in any case, economic conditions had been relatively stable in the run up to the 1924 Budget. Instead, despite his Labour affiliation, Snowden fully embraced the view that taxes should be low in the long run and set about sweeping reductions in the tax burden.

Debt redemption

What factors explain the widespread emphasis on debt redemption? There were a number of reasons for the desire to maintain a tight fiscal stance throughout this period. First, Victorian Britain sustained a deep aversion towards debt, and the Treasury recommended

that debt should be reduced at all costs and in most circumstances except war time. Interwar support for the MBBR was rooted in the idea that, even at low employment levels, any deficit spending would lower confidence in the government's ability to repay its debt, eventually harming the private sector (Middleton, 1996, p. 323). This "whole political culture of "safety first"" explains why we sometimes see tax increases in our dataset even when the deficit was not getting any worse (see Tomlinson (1990, p. 67) and Middleton (1985, pp. 83-5)). Taxes were sometimes increased even when there was a surplus. Britain, indeed, was more particular about its overhanging debt (137 percent of GDP in 1919) than any other war-ravaged country (Alesina, 1988). As Middleton (1996, p. 340) makes clear, "the higher taxation relative to France and Germany was a clear consequence of the earlier imposition of strict budgetary control in Britain in the immediate postwar years" (see also Alford (1972, p. 65), Alesina (1988, p. 64) and Ritschl (1996)). Edwin Montagu indeed asked Lloyd George in 1921: "Is it conceivable that England should prostitute itself to the level of France and budget for a deficit?" (Short, 1985, p. 177).

Second, support for the Gold Standard also arguably played in favour of a tight fiscal policy. In order to stay on the Gold Standard and resist inflationary tendencies, Britain had to repay its debt in an orderly manner. Staying on the Gold Standard was a long-term goal which mattered to Britain because of its importance for trade and the City, as reaffirmed by the Cunliffe Committee (1918) just after the war (see Brown (1929, p.63), Pollard (1970, p. 17), Moggridge (1972), Dimsdale (1981), Peden (1987, p.17) and Accominotti (2012)). Of course, adherence to the Gold Standard was not just an end in itself; it was also a way for the government to tie its own hands from a budgetary perspective.[12]

Finally, some have argued that adherence to a tight budgetary stance was also spurred by a will to support the rentier class. Alesina (1988, p. 66-7) suggests that the policy of tax increases, deflation and high interest rates mainly served debt holders to the detriment of taxpayers. Such intentions were famously denounced by both Keynes and Churchill.

Policymakers therefore sought to steadily reduce the national debt burden for one or several of the aforementioned reasons and this is useful for our purposes. While there were some

[12]Chancellor Winston Churchill hosted a dinner party in March 1925, attended by former Chancellor Reginald McKenna and John Maynard Keynes, in which "[John] Bradbury [former Joint Permanent Secretary of the Treasury] made a great point of the fact that the Gold Standard was knave-proof. It could not be rigged for political or even unworthy reasons" (as reported by James Grigg, Churchill's private secretary; cited in Moggridge (1969, p. 61)).

instances in which the fiscal situation seemed to deteriorate quickly and demand immediate action (such as in 1920 or 1931), in many other instances there was no such sense of emergency. This long-run interest in fiscal consolidation enabled the orderly payment of debt interest, which made up about half of government outlays every year, and regular payments into a sinking fund. The interwar view of the sinking fund was that of a "new" sinking fund whose amounts were to be pre-determined, scheduled in advance, and not responding to short-run economic fluctuations — usually around £50 million per year (Brittain, 1959, pp. 202-203).[13]

The fact that sinking fund payments were included as expenditures in of the Treasury's calculation for the budget surplus made any deficit look worse than it would today. This was unusual even among European countries at the time and reflects the importance of this pre-determined commitment in shaping Chancellors' views on the budget. While there were a few changes in the pre-committed amount, once the sinking fund commitments had reached a new plateau, such as in 1923 and 1933, any subsequent changes in the following decade were minor (see Appendix Figure A.5). Figure 1 illustrates the tight stance of fiscal policy over this period. Although the budget was rarely exactly in balance and, as noted above, different Chancellors put relatively more and less weight on debt redemption objectives, Figure 1 shows that the central government budget balance, with sinking fund payments included in expenditures, was usually less than 1% of GDP in either direction. This can be compared with deficits of 3 percent of GDP which would become routine in the postwar era (Middleton, 2010, p. 431). Excluding the sinking fund payments from expenditures, however, shows that governments were actually running clear surpluses for a number of years over this period.

To summarize, debt redemption policy was largely separated from year-to-year tax policy reforms. There were instances, however, when concerns about deficits and the debt did explicitly affect tax policy decisions. First, some tax increases were designed to accelerate the repayment of the inherited level of debt. Such Budgets fit well the Romer and Romer (2010) concept of deficit-driven policy changes unrelated to current or expected economic conditions and we classify these as exogenous. Second, there were sometimes emergency steps taken to deal with a sudden deterioration in the budget. These types of changes, in

[13]This is to be contrasted with Walpole's "old" sinking fund concept whereby only any arising surplus from the previous year could be used for such payments. The "old" sinking fund concept is closer to the post-WWII view of sinking funds Brittain (1959).

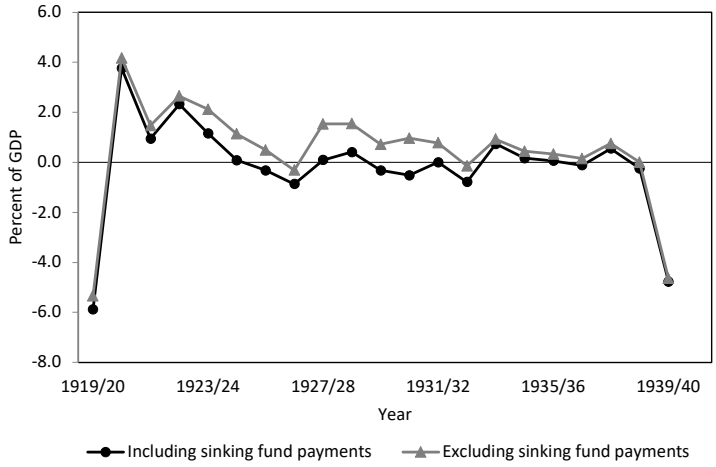


Figure 1: Central government budget balance (% GDP), 1919-1939

Notes: The darker line represents central government budget surplus figures as they were represented at the time, with spending estimates including sinking fund payments and excluding social insurance funds (the data is from Mitchell (1988)). Modern budget surplus definitions would exclude the sinking fund but include the social insurance funds, as Middleton (1985, pp. 78-81) showed. Even under this modern definition, though, the budget deficit would still look remarkably low relative to other countries and relative to today (see Table 6.1 in Middleton (1985, p. 96)). In the lighter-shaded line we calculate the budget balance excluding sinking fund payments from spending estimates. The sinking fund figures are from Mallet and George (1933a), Mallet and George (1933b), and Peden (2000). GDP figures are from Mitchell et al. (2012).

contrast, will be treated as endogenous.

One further issue to consider is whether some of the tax changes we study might imply further spending changes, and vice versa. Although empirical fiscal multipliers always implicitly depend on the response of other fiscal variables, given the emphasis on fiscal prudence during this period it is worth paying additional attention to this. First, as mentioned above, it is worth noting that in terms of new policy announcements, the Budget focused much more on setting taxes for the coming year. It was less common to see new spending announcements. But, by examining the narrative we are able to classify as endogenous those specific tax changes that were directly associated with specific spending changes. Still, one may be concerned that previous trends in spending might be influencing our results, or that spending might subsequently decline following the tax cuts we classify as exogenous. We can address this in two ways. First, we can control for lagged government spending and, second, we will estimate the response of government spending to our identified tax reforms, which allows us examine the importance of subsequent movements in government spending for the

interpretation of our multiplier estimates.[14]

2.4 Summary of Classifications

Having outlined the general objectives of fiscal policy during this period and the key econometric issues, we can now describe the categories we use to assign motives to policy changes. The primary distinction is whether the policy changes were endogenous — responding to current or prospective fluctuations in economic conditions such as GDP or the deficit — or whether the tax change was taken for other reasons, possibly related to ideological objectives. Table 2 summarises our classification system. The most important classification is between endogenous (N) and exogenous (X) Budgets. But, to aid the classification of motives, we also provide a number of sub-categories. Within each broad category, there are different types of exogenous and endogenous changes. Like Romer and Romer (2010) we therefore provide more information on the rationale for each change. Note that, within the endogenous and exogenous groups, the subcategories are *not* intended to be uncorrelated with each other. Rather they provide more detail on the nature of the motivation for the exogenous or endogenous classification.

We start by discussing the nature of the tax changes we regard as endogenous. First, given the focus on debt reduction over this period, there are occasions where Chancellors raised taxes when contemporaneous economic conditions deteriorated and generated fiscal stress. In the most severe downturns, taxes were actually increased (e.g. in 1920 and 1931). We refer to these changes as urgent deficit reduction measures (DR). The focus on debt redemption generated somewhat of an asymmetry in the sense that realized surpluses were automatically applied to the redemption of debt, but concerns about a realized deficit tended to motivate tax increases.

Second, although policy was pre-Keynesian, there were Budgets that sought to provide some form of stimulus given changing economic conditions. These actions do not exactly resemble countercyclical policy actions seen in the decades after World War II. Stimulus policies tended to be targeted interventions and were often measures designed to encourage

[14] In particular, we will show that the response of government spending is in line with Budgetary orthodoxy, but the magnitudes are actually relatively muted. It is, of course, worth emphasizing that all empirical studies on the tax multiplier typically have government spending implications from tax shocks. We discuss these fiscal financing issues in more detail in Section 5.

production rather than boost aggregate demand. In short, these were not deficit-financed tax cuts, but were still motivated by current economic fluctuations. Policies that were designed to impart a stimulus to production we will label "supply stimulus" (SS) measures, and those to encourage private expenditure we will (loosely) call demand management (DM) measures. The former types of changes are much more common than the latter, reflecting the economic orthodoxy of the time. There was, therefore, some variation across Chancellors in the weight given to debt redemption, although, as noted above, stimulus actions tended to be targeted rather than Keynesian style boosts to aggregate demand. They also sometimes occurred when the budget was already in surplus.

Third, some policies may have been specifically tied to a spending change. Because these are correlated with an (omitted) spending variable, we refer to these tax reforms as endogenous spending-driven (SD) changes and these are therefore excluded from our analysis.

Turning to the exogenous tax changes, these have a closer connection to the categories often used in more recent narrative exercises (e.g. Romer and Romer (2010) and Cloyne (2013)), although the precise goals are still a reflection of the priorities of the time. For example, various Budgets targeted social objectives, reflected a particularly ideology (for example Imperial Preference) or reflected personal priorities of the Chancellor (for example Churchill's Betting Duty – see Cloyne et al. (2023)). These types of changes we refer to as socially or ideologically motivated (IL). The fiscal orthodoxy of the period also generates two further categories. First, there was often a desire to keep taxes as low as possible to stimulate long-run growth. Following Romer and Romer (2010), we refer to these as long-run tax reforms (LR). Second, given the general fiscal consolidation philosophy over this period, successive governments engaged in a steady process of trying to lower the overall level of government debt over time. On occasion, commitments to this plan were renewed or reinforced. As noted above, these decisions are motivated by a long-run desire to reduce inherited debt levels, rather than from year-to-year movements in the deficit. Following Romer and Romer (2010), we regard these as exogenous and refer to them as long-run fiscal consolidation measures (FC).[15]

[15]Earlier we noted that it is important to consider possible unobserved relationships between taxes and spending in this period. This may be particularly relevant for *X-FC* changes. When we can detect a clear contemporaneous correlation between new spending announcements and tax changes, these would be classified as endogenous. Still, spending might have been falling prior to the Budget, or might be adjusted afterwards. As mentioned above, we handle this by controlling for lagged government spending (which makes

Table 2: Categories of tax changes

Group	Sub-category
Endogenous (N)	1. Stimulus measures: either for supply (SS) or demand (DM) 2. Urgent deficit reduction measures (DR) 3. Spending-driven changes (SD)
Exogenous (X)	1. Social/ideological objectives (IL) 2. Long-run performance (LR) 3. Long-run fiscal consolidation (FC)

3 Interwar Tax Changes

3.1 An Overview of the Narrative Evidence

What were the tax changes over this period and how were they motivated? In what follows we provide a high-level overview of our narrative account and the main sources of variation we consider. We also try to give a flavour of how we weigh-up the different sources of evidence and compare aggregate policy motivations with more specific ones. The background companion paper Cloyne et al. (2023) provides much greater detail on the tax changes, the motivations and justifications given in the Chancellor's Budget speech. This companion paper also provides information on all the precise classifications and contains the data themselves.

Postwar Policy: 1919-1924

From 1919 to 1924 Britain first went into postwar recession but soon emerged into a more stable period. This era was marked by both endogenous and exogenous Budgets. The most striking ones are Austen Chamberlain's urgent deficit reduction tax increases in the 1920 Budget and Snowden's exogenous ideological and long-run tax cuts in the 1924 Budget.

Budgetary orthodoxy was an important pillar of Austen Chamberlain's policymaking (Short, 1985, p. xvii) and this was followed just coming out of the war. There was little concern about the deficit in early 1919 and the Chancellor did not seem to feel any strong

little difference) and by examining the subsequent response of government spending to our tax shocks (see Section 5.1). Finally, in the end there is only one Budget in our sample period classified as *X-FC*. Excluding this 1930Q2 shock entirely also does not materially affect our conclusions, as shown in Appendix A4.9.

sense of fiscal emergency in his April 1919 Budget. In part this may have been due to Chamberlain having grossly underestimated expenditure, which he assumed to be close to its pre-war basis (Short, 1985, p. 11). In general, Chamberlain's 1919 Budget instead focused on two longer-run objectives: to fortify the long-run fiscal position of the country while at the same time applying as far as possible his protectionist principles. This Budget therefore fits the description of an exogenous (X) budget. The rise in estate duties, for instance, is easily seen as a longer-term measure to fortify revenue, and hence these are related to a *fiscal consolidation* (FC) motive, all the more so that "Death Duties are not a suitable instrument for meeting a temporary emergency" (UK Parliament, 1918-1938, HC Deb 30 April 1919 vol 115 cc206-7). A similar interpretation can be offered regarding the increase in revenue from spirits. On the other hand, Imperial Preference drove numerous cuts in customs duties, and take on a *ideological* (IL) character.

After the underevaluation of expenditure became evident in the summer of 1919, Austen Chamberlain was pressed by the Treasury and even *The Times* to pay more attention to the size of the national debt (Short, 1985, pp. 15-7).[16] In a clear contrast with the preceding one, a strong sense of emergency therefore emanates from the 1920 Budget and the Chancellor faced a very large deficit. Given these contemporaneous developments he felt compelled to act. The tax increases were very sizable, as much as 3% of GDP. This Budget is best seen as a package of measures designed to urgently tackle the fiscal situation and we therefore classify this as an endogenous (N) Budget for *deficit reduction* (DR). In a move that put Britain on a much stronger footing for years to come, Chamberlain increased revenue from the Excess Profits Duty to £100 million and diverted it away from war expenditure, while at the same time raising the corporations profits duty to £35million. Despite this, for reasons of fairness Chamberlain still sought to substantially reduce income tax on the lower brackets, following the recommendations of the Royal Commission on Income Tax which had deliberated in 1919-20. These remissions reveal some longer-term preferences, but it is doubtful whether income tax could have been reduced to such an extent (£29 million) without a commensurate increase in taxes elsewhere.

Horne was the only other Chancellor along with Winston Churchill who did not strictly

[16]Harold Harmsworth, also known as Lord Rothermere, who was a leading newspaper proprietor of the time, actively campaigned against government spending (the "Anti-Waste" campaign) and used *The Times* as his main medium of expression.

adhere to the prevailing orthodoxy (see Peden (1987, p. 147-9) and Short (1985, p. 152, 166)). Britain entered a full-blown recession in 1921 but he preferred to lower taxes, giving the appearance of a balanced budget by temporarily suspending sinking fund payments (Short, 1985, p. 180). Horne makes various comments that suggest his goal was to provide a stimulus "in a very trying, anxious and critical time for the commerce and industry of this country" (UK Parliament, 1918-1938, HC Deb 01 May 1922 vol 153 cc1036-7). The tax changes tended to be described as imparting a stimulus to production, rather than for supporting demand. We therefore classify this budget as endogenous (N), being designed to impart a *stimulus to supply* (SS).

On the other hand, the normal year concept was fully embraced by Baldwin and Snowden. This period saw relative economic stability from year to year and both Chancellors seemed influenced by the long-run aspiration of a low tax environment and set about significantly reducing the long-term burden of taxation. Perhaps surprisingly given his support for the Labour Party, Snowden deeply believed in the importance of a limited state and low taxes for Britain (Snowden, 1920). As Boothby noted:

> To every outworn shibboleth of nineteenth century economics he clung with fanatical tenacity. Economy, Free Trade, Gold – these were the key-notes of his political philosophy; and deflation the path he trod with almost ghoulish enthusiasm. (...) To every plea for expansion – and many were made from both sides of the House of Commons – he remained totally deaf" (cited in Middleton (1996, p. 320)).[17]

Snowden's 1924 Budget is therefore classified as exogenous (X), following several *long-run* (LR) and *ideological* (IL) motives. He differed from his Conservative counterparts only in his insistence that tax policy should be used to increase fairness and redistribution.

Churchill and the Gold Standard: 1925-1930

With Britain back on gold in 1925 under Winston Churchill, one might have expected further policies to keep her safely on this path. However, neither low taxes for everyone nor a focus on balanced budgets were really Churchill's cup of tea.[18] Churchill even implemented a

[17] Baldwin's philosophy, which was similar to Snowden's, emanated from a fear of big government in the face of influences from communism and fascism (Middleton, 1996, p. 317).

[18] Although it was Churchill who led Britain back to gold, he eventually thought it had been "the greatest mistake of [his] life" (Capie and Wood, 2012, p.187).

stimulus to supply following the General Strike in the 1928 Budget. To preserve an appearance of integrity in such situations, he resorted to what he himself called "my adventitious resources" by, for example, raiding the Road Fund (see Hicks (1938, p. 7) and Hancock (1970)). Indeed Churchill often sought to flow against the stream, fuelled in particular by lengthy conversations with Keynes (Pollard, 1970). Deficit reduction was only resorted to when he perceived a dangerously widening gap in the state's finances – for example, in the 1926 and 1927 Budgets. As a result, Churchill's 1926, 1927 and 1928 Budgets are classified as endogenous (N), either due to their desire to stimulate supply or to address the urgent budgetary situation.

Churchill also shared with Keynes the view that "gigantic taxation" and deflation mainly served rentiers at the expense of the average taxpayer, and sought to restore the balance by relieving the middle class while unnerving the elite (see Daunton (2002, p. 124) and Short (1985, p. 211, 223)).[19] Tempted by the capital levy as examined by the Colwyn Committee on the National Debt and Taxation, he eventually backed down. But his 1925 Budget greatly reduced the standard rate of income tax while substantially raising the estate duty. This Budget is classified as exogenous (X) as these tax changes were motivated by his ideological preference for the re-organization of tax policy, rather than in response to any changes in the economic outlook. The 1929 Budget, by contrast, lacks a clear overall objective and is made up of lots of small, idiosyncratic changes each motivated in a specific way. We classify this Budget as exogenous (X), following a number of *long-run* (LR) and *ideological* (IL) objectives, although, as discussed in our companion background paper, there is one minor *spending-driven* (SD) item that we separate from the rest of the Budget measures.

Despite relatively high levels of unemployment at the end of the 1920s, the Liberal party's pamphlet *We Can Conquer Unemployment* did not manage to garner sufficient political support. Instead, the Labour Party returned to power in the 1929 General Election. With Snowden back at the Exchequer, Churchill's "relative profligacy" (Tomlinson, 1990, p. 77) could only be met with contempt. In 1930 Snowden reasserted his will to restore an austere budgetary approach and raised all major tax items substantially. The general tone of the Budget is not one of emergency.[20] As Middleton (1996, p. 321) points out, Snowden's return

[19] He said: "There is more to the life of a nation than the development of an immense rentier class quartered in perpetuity upon the struggling producer of new wealth" (Daunton, 2002, p. 123).

[20] Churchill eventually accused Snowden and the Treasury to be like-minded spirits who "embraced them-

to the Treasury "was welcome by officials who had been only too conscious of expenditure growth during the second Baldwin administration and Churchill's propensity to be distracted from the path of strict orthodoxy." Snowden thus arrived at the Treasury with a desire to tackle the inherited debt burden more seriously. His 1930 Budget can be seen as a package of measures entirely for this purpose and we classify it as exogenous (X), *fiscal consolidation* (FC).

Crisis and Recovery: 1931-1935

Snowden had remained quite hopeful throughout most of 1930, and his 1930 Budget had promised that no new taxes would be imposed in the next Budget. However, Britain's economy substantially deteriorated over the winter of 1930-1931. Although up to 1931 the Unemployment Insurance Fund had remained broadly outside the central government budget (Peden, 2000, p. 238), unemployment rose to 12% and financing the fund through the usual channels (employer and employee contributions) became very difficult.[21] The tax increases in the April 1931 Budget were modest, although clearly driven by concerns about the deterioration in the fiscal position. But Snowden also appointed the May Committee to look for ways to economise on expenditure going forward. Although a prospective deficit of £120 million, or 3.1% of 1931/2 GDP, would become common in the postwar era, it was widely viewed as alarmingly high (Middleton, 2010, p. 431). Labour could not, however, agree on cuts to unemployment benefits, leading to the formation of a new National coalition government in August, which precipitated the fall in the pound (Alford, 1972, Tomlinson, 1990, Capie and Wood, 2012).

The tone of the supplementary Budget speech delivered in September 1931 by Snowden, who had remained Chancellor in the new National government, is thus one of emergency. In a significant departure from his habitual policy goals, Snowden not only substantially increased taxes but did not even refrain from placing some of the burden on the middle class and the poor (Short, 1985, p.293) (see also Daunton (2002, p. 159)). Given the serious economic conditions, the supplementary Budget of 1931 is therefore a clear example of an urgent, and

selves with all the fervour of two long-separated kindred lizards" (Daunton, 2002, p. 144).

[21]Note that before 1931 social insurance payments did not come entirely from employer and employee contributions. Some Exchequer payments went into those funds. For example, in the early 1920s Exchequer contributions to the Unemployment Insurance Fund accounted for a fifth to a quarter the Fund's expenses, but were not officially included in the central government Budget (Stationary Office, 1997, Table 80, p. 106), and thus not discussed by Chancellors in their Budgets until September 1931.

endogenous (N), increase in taxes for *deficit reduction* (DR) purposes.

In the following year the debt continued to be threateningly high. In October 1931 the National government was reelected, this time with Neville Chamberlain as Chancellor, who, in 1932, seized the opportunity to realise his long-standing wish of protection through tariffs.[22] Although it is possible to view the policy changes related to tariffs in February 1932 and the April 1932 Budget as ideological, the seriousness of the situation still does not seem to warrant an exogenous treatment. Given all the comments in the 1932 Budget speech, it seems more likely that the economic situation motivated large measures to tackle the deficit. And these circumstances provided Chamberlain a rationale to then increase tariffs. The 1932 Budget (including the earlier measures in February) is therefore classified as endogenous (N), and for *deficit reduction* (DR).

With Britain off gold in 1931, the budget under control and a successful conversion operation in the summer of 1932, interest rates could be brought down from 5 to 3.5%. The economy was now operating more normally and Neville Chamberlain's 1933 Budget is a potpourri of measures for which it was hard to find an overarching goal. This is the second and last "idiosyncratic" Budget after Churchill's 1929 one. The majority of reforms were technical in nature and relatively small. All these suggest the Budget was largely exogenous in nature, although there were some targeted stimulus measures designed to encourage production which we treat as a separate mini-package of endogenous actions. We explain this decision in much more detail in the background companion paper. In his 1934 Budget, Chamberlain implemented a substantial income tax cut (£24 million). The motivation for this could, potentially, be partly ideological but overall Chamberlain seemed to be influenced by recent economic performance, which he hoped to speed up. We therefore regard this as an endogenous (N) Budget for stimulus purposes, mainly for demand management (DM). By 1935, the tax reforms start to take on a longer-term nature and we therefore classify this Budget as exogenous (X). The 1935 Budget is seen through the lens of *long-run* (LR) and *ideological* (IL) reforms.

Rearmament: 1936-1938

Preparation for war against Germany dominated the rest of the period. Budgets between 1936 and 1938 all refer to rearmament as the main priority, and are therefore entirely

[22]Neville, like his father Joseph, had always worked against free trade (Self, 2006).

endogenous (N) and *spending-driven* (SD).

3.2 The new tax shock series

Across the 23 fiscal events in our sample, there are over 200 individual tax changes each with their own implementation date. We assign tax changes to the implementation dates given in the Budget documents and aggregate the tax changes into a quarterly time series for economic analysis. The tax changes are grouped into exogenous and endogenous categories following the classifications discussed above. Following Romer and Romer (2010), a tax change implemented in the final half of any quarter is assigned to the next quarter but, in Appendix A4.6, we show that this timing choice makes virtually no difference to our results. The choice to assign a tax change to the implementation date (rather than the announcement date) also raises the issue of anticipation effects. If a tax change is implemented several quarters after it is announced, the economic effects could be realized before the measure is implemented. In Section 5.3 we show that most tax changes are implemented without a lag and show that our baseline results are very similar when we restrict attention to tax changes that were implemented in the same quarter that they were announced. Much more detail on the classification itself and some of these technical choices can be found in the companion background paper.

We therefore take our new set of tax changes, assign them to quarters, aggregate and scale by nominal GDP.[23] Figure 2 shows the resulting aggregate series for exogenous and endogenous tax changes.

There is considerable variation throughout the sample, and many of the tax changes are sizable. This reinforces the suitability of this period of U.K. history for our analysis. Some of the key reforms outlined in Section 3.1 are also clearly visible. In 1919, the black upward spike corresponds to Chamberlain's 1919 tax increases. However, the large endogenous increases in the early 1920s come from his emergency measures to tackle the large deficit following the economic downturn. Horne's stimulus in 1922 is easily recognizable as the sizable downward

[23] In the baseline specification we use nominal GDP in the previous quarter, but the results are not sensitive to this. Quarterly nominal GDP is also not directly available so we construct a nominal GDP series by multiplying the real GDP series from Thomas and Dimsdale (2017) by a price deflator. Our price deflator is a weighted average of consumer and export prices. The weights are chosen by matching our new series with the annual GDP deflator that is available for this period. The results are similar using lagged annual nominal GDP data to scale the tax changes.

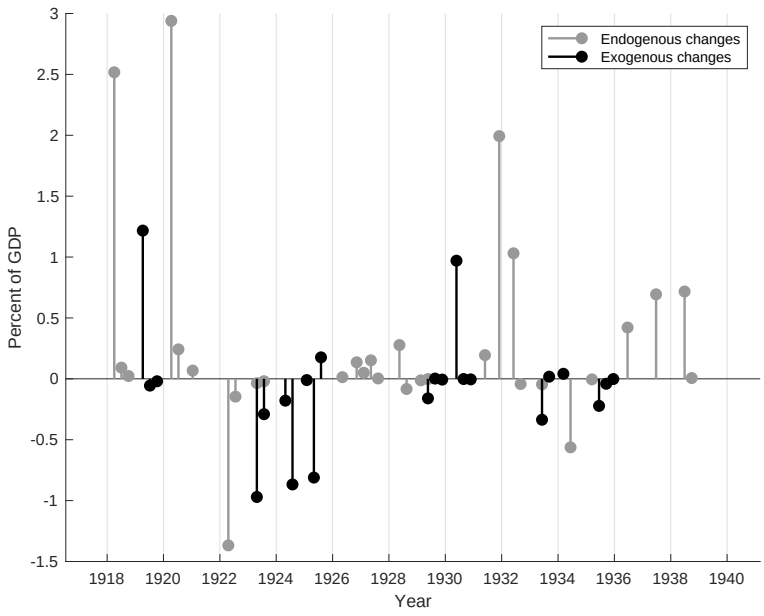


Figure 2: Tax shocks in interwar Britain

Sources: Stationary Office (1918-1938) and authors' calculations

grey spike. The sizable exogenous tax cuts in the mid-1920s reflect ideological beliefs held by both Baldwin and Snowden that the tax burden should be made permanently lower. Over the second half of the 1920s, Churchill's complex character resulted in many little changes, but few sizable ones.

The year 1930 sees Snowden back to power and the effect of his long-standing desire to restore fiscal order following Churchill's long-term inaction, which is illustrated by the black spikes in the early 1930s. But, as discussed above, the unfolding crisis in 1931 led him to raise taxes significantly in response to contemporary economic troubles, which is captured in the sizable grey spike in 1931. The following years see a number of smaller exogenous and endogenous changes. The fiscal consequences of rearmament become clear from 1936 onwards.

Table 3: Granger causality tests

Series	Test statistic	p-value
Exogenous series		
GDP	0.08	0.99
Unemployment	0.37	0.83
Bank rate	0.91	0.47
Consumer prices	0.53	0.71
Government spending	0.95	0.44
Endogenous series		
GDP	2.72	0.04
Unemployment	2.85	0.03
Bank rate	6.51	0.00
Consumer prices	2.07	0.10
Government spending	0.52	0.72

Note: This table shows the test statistics and p-values associated with the Granger causality tests. A high p-value implies that it is not possible to reject the hypothesis that each variable does not predict the tax shock series. Each row shows the results from each bivariate VAR with our tax shocks (exogenous and endogenous) and GDP growth, the change in the unemployment rate, the log change in prices, the log change in government spending and the change in the Bank of England policy rate (Bank Rate). Four lags are used. Similar results are obtained using a different number of lags. Results are also similar when we directly regress the tax shock on lags of the variable of interest (rather than using the more traditional VAR Granger causality setup).

The narrative approach isolates exogenous tax changes based on the motivations given by policymakers. These are identification assumptions and we cannot test the contemporaneous exogeneity of our tax changes. We can, however, provide some statistical reassurance that our method has been successful by checking whether our exogenous changes are predictable based on past macroeconomic variables. Table 3 shows results of Granger causality tests. We attempt to predict our new series of tax changes using information on past quarterly GDP growth, the change in the unemployment rate, consumer prices and Bank rate, government spending and the Bank of England's policy interest rate. The results are striking: the endogenous tax series is highly predictable but the exogenous series is not. The null hypothesis is that lags of these variables do not Granger cause the exogenous series and this hypothesis is clearly not rejected, with very high p-values.

4 The Macroeconomic Effects of Tax Changes

Using our new series of exogenous tax changes we can now explore the link between tax reforms and economic outcomes. Tax changes are likely to have contemporaneous and dynamic effects: policy changes may affect the economy gradually over time and changes in tax rates may also persist for a number of years. We are therefore interested in the dynamic causal effects of changes in taxation. To do so, we start by estimating impulse response functions using the local projection technique of Jordà (2005). Specifically, we estimate the following sequence of regressions:

$$y_{t+h} - y_{t-1} = \alpha^h + \beta^h \Delta\tau_t + \Gamma^h(L)\mathbf{X}_{t-1} + \Theta^h(L)\Delta\tau_{t-1} + u_{t+h} \tag{1}$$

where y_{t+h} is the variable of interest, for example log real GDP, at horizon h and $\Delta\tau_t$ is our new series of identified exogenous tax changes. β^h is the impulse response function at horizon h. Identification of the causal effect requires that the tax changes $\Delta\tau_t$ are uncorrelated with the other macroeconomic disturbances contained in the error term u_t, conditional on controls. Our narrative identification strategy attempts to isolate policy changes that are unrelated to the shocks driving u.[24]. $\mathbf{X}$ is a vector that will contain lagged macroeconomic controls. Although our narrative approach primarily looks for tax changes that are not responding to current or expected future economic fluctuations, if our shocks are closer to being strictly exogenous then adding controls should make little difference.[25] The number of lags for the shocks is denoted Q (lag polynomial Θ) and the lag length for the other controls is P (lag polynomial Γ). The residuals may be serially correlated so we compute standard errors using the approach of Newey and West (1987).[26]

First we examine the percentage response of real GDP to our narrative identified tax shocks. y is log real GDP and $\mathbf{X}$ includes the change in: log real GDP, Bank rate, unemployment rate and the log consumer price index. These historical data are available from Thomas and Dimsdale (2017) and more information is given in Appendix A3. Quarterly real

[24] Stock and Watson (2018) call this as "lead-lag" exogeneity, at least conditional on $\mathbf{X}$ and lags of $\Delta\tau$

[25] In Appendix A4.1 we show that the choice of controls does not materially affect our results.

[26] Since our local projections are "lag augmented", Huber-White standard errors can potentially be used as an alternative to the common Newey-West approach, as shown by Montiel Olea and Plagborg-Møller (2021), although implementation still requires the researcher to choose the appropriate lag length for the controls. That said, conclusions are very similar using Huber-White standard errors, as mentioned below when we present the main results.

GDP data are available from 1920Q1 to 1938Q4 so, to maximize the sample period, we set $P = 1$. We also include four lags of the shocks.[27] Appendix A4.2 shows that the main results are robust to considering other P, Q pairs.[28]

Figure 3 shows the percentage response of real GDP h periods after a cut in taxes of 1% of GDP.[29] The tax cut leads to a sizable and persistent effect on real GDP. The effect builds up over time with the peak effect felt around the start of the second year. The effect starts close to zero in the initial quarter but quickly rises to over 3% in the second year. The figure also plots 95% and 68% confidence intervals, which are approximately equal to two and one standard deviation error bands. The effect is statistically significant between horizons 1 and 4 using two standard deviations and is significant for most of the IRF period using a one standard deviation threshold.[30]

Given that $\Delta\tau_t$ is measured as the projected change in tax revenues scaled by GDP, the impulse response function in Figure 3 can already be interpreted as a particular measure of the so-called fiscal "multiplier": the £ change in GDP h periods ahead for a £1 cut in taxes. While informative, this way of reporting the multiplier does not capture the persistence of the movements in taxes over time. The tax shock, $\Delta\tau_t$, is in differences and the simulation in Figure 3 can be thought of as capturing a shock that potentially lowers the tax-to-GDP ratio persistently. As a result, another way to examine the GDP response is to convert the impulse response function for GDP into a *cumulative multiplier* (see, for example, Mountford and Uhlig (2009), Uhlig (2010), Ramey (2016), Ramey (2019)). This statistic is the *total* £ change in GDP up to period h, divided by the total tax remission, in £, over the same period. The cumulative GDP tax multiplier seeks to compare the total gain in GDP at some

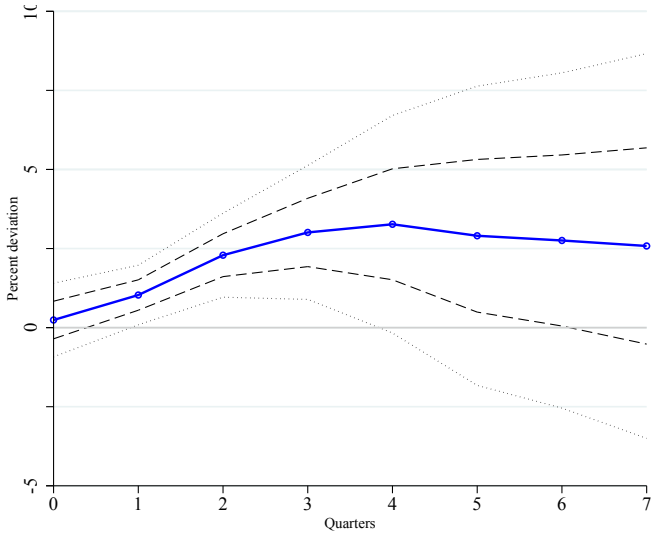
[27] Note that, as discussed in Section 3.1, we have identified shocks for 1918 and 1919 so it is possible to have a longer value for Q than P. Appendix A5.3 also shows results based on an extended sample period where we use GDP data back to 1918 by linearly interpolating real GDP from 1920 to 1918.

[28] Given the difference specification on the left hand side of equation 1 and that we include one lag of the controls ($P = 1$), our regression sample starts in 1920Q3. As usual with local projections, the regression sample is reduced by $h = 8$ observations prior to 1938Q4. To be consistent for each h, we use the same sample for every period. An alternative strategy is to allow the sample size to vary with h but this makes little difference to our results.

[29] A growing recent literature has noted that the flexibility of the local projections framework can come at the cost of more volatile IRFs see, e.g., Barnichon and Brownlees (2019). For presentation, the IRF is therefore smoothed using the 3 quarter moving average approach in Tenreyro and Thwaites (2016). The raw estimates are also presented in Appendix A4.3.

[30] A very similar conclusion emerges using Huber-White robust standard errors for Figure 3 and again for the LP-IV multiplier estimates shown later in Figure 4.

Figure 3: Response of GDP (%)



Notes: The figure shows the percentage response of real GDP h periods after a 1% of GDP cut in taxes. Dashed and dotted lines represent one and two standard error bands computed using the Newey-West approach. The impulse response function has been smoothed using a 3-quarter moving average following Tenreyro and Thwaites (2016). The raw IRF is included in Appendix A4.3.

horizon h, relative to the total lost tax revenue over the same period. More specifically:

$$\frac{\sum_{j=0}^{h} \Delta^j GDP_{t+j}}{\sum_{j=0}^{h} \Delta^j T_{t+j}} \tag{2}$$

where $\Delta^j GDP_{t+j}$ and $\Delta^j T_{t+j}$ are the £ change in GDP and tax revenue from the initial period $t-1$ to $t+j$ (relative to the case where taxes are not adjusted).[31]

Ramey (2016) shows how the cumulative fiscal multiplier at horizon h can be computed directly using the narrative shock as an instrument for the change in actual tax revenues over time.[32] The approach of using the narrative shocks as instruments also addresses the point raised by Mertens and Ravn (2013) that the narrative shock might only be a proxy for true structural shock to taxes as a share of GDP. To estimate the cumulative fiscal multiplier using a local projections instrumental variables (LP-IV) approach we use the following specification:

[31] It is also possible to include discounting, although this should make little difference.

[32] Jordà and Taylor (2015) also use narrative shocks as instruments in a local projections (LP) framework. For more detail on using instruments in time series applications, including the differences between VAR and LP frameworks see, e.g., Stock and Watson (2018) and Plagborg-Moller and Wolf (2021).

$$\sum_{j=0}^{h}(\ln GDP_{t+j} - \ln GDP_{t-1}) = \alpha^h + B^h \sum_{j=0}^{h}\left(\frac{T_{t+j} - T_{t-1}}{GDP_{t-1}}\right) + \Gamma^h(L)\mathbf{X}_{t-1} + \Theta^h(L)\mathbf{\Delta}\tau_{t-1} + u_{t+h} \tag{3}$$

where GDP is real GDP and T is real tax revenues (deflated using the GDP deflator).[33] $\sum_{j=0}^{h}\left(\frac{T_{t+j}-T_{t-1}}{Y_{t-1}}\right)$ is then instrumented with the narrative identified shocks. The object B^h is then the cumulative fiscal multiplier at a particular horizon h.

The challenge we face is that the U.K. National Accounts do not provide tax revenue data prior to 1946 and there is only an annual series available in Thomas and Dimsdale (2017). Instead, we make use of a new series of government receipts collected by Lennard (2020) from *The Economist*. That said, there are a number of limitations in our context. Receipts are somewhat volatile over this period and tend to exhibit very large spikes depending on the collection dates of different taxes. Receipts can be sensitive to the precise timing of collection, which may be more problematic in this historical period before Pay as You Earn income tax systems were available. The series also generates some large outliers in $\frac{T_{t+j}-T_{t-1}}{Y_{t-1}}$ in the first half of 1920. To tackle these issues, we seasonally adjust these data (see Appendix A3 for data sources and definitions) and we also start our sample in 1920Q4 to avoid these outliers.

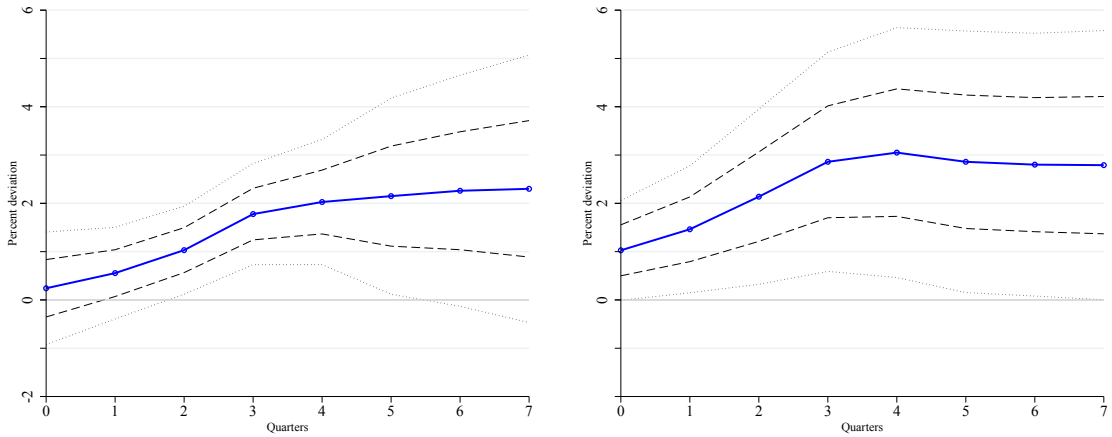
The narrative shocks are a measure of the full year projected change in tax revenues. As such, they can be seen as a proxy for the actual change in taxes over the coming year. It is therefore natural to consider using both the current period value and its lags as potential instruments. In our baseline specification we use the current period and first lag of the narrative shock. For the lag structure we continue to use $P = 1$ and $Q = 4$. $\mathbf{X}$ contains lags of the dependent variable, the change in tax receipts as a share of GDP and we also include further lags of the narrative shock.[34]

Given the discussion above, Figure 4 presents two versions of the cumulative multiplier. Panel (a) assumes that the narrative shocks do indeed lower taxes to GDP by 1% for the whole period of the impulse response function. This is a reformulation of the impulse response function in Figure 3 where we cumulate the IRF for real GDP and divide by the assumed cumulative reduction in taxes. We refer to this as the baseline "static" approach because

[33] Also, to help compare equations 2 and 3 note that the log difference is approximately $(GDP_{t+j} - GDP_{t-1})/GDP_{t-1}$.

[34] Appendix A4.4 contains further robustness checks regarding the LP-IV specification.

Figure 4: Cumulative GDP tax multiplier



(a) Baseline "Static" Approach (b) IV Approach

Notes: The figure shows the cumulative GDP multiplier, $\sum_{j=0}^{h} \Delta GDP_{t+j} / \sum_{j=0}^{h} \Delta T_{t+j}$, h periods after the shock for 8 quarters. Panel (a) assumes taxes are reduced as a share of GDP by 1% and is therefore a reformulation of impulse response function in Figure 3. Results in Panel (b) are computed using the local projection IV strategy using tax receipts data as outlined in the text. $P = 1$ and $Q = 4$. In panel (b) we use the current and lagged value of the shock as instruments. Dashed and dotted lines represent one and two standard error bands using the approach of Newey and West (1987).

it does not account for the potential dynamic effects on actual tax revenues in constructing the denominator. This is an interesting object in its own right as it provides an estimate of the effect on GDP for a tax change that policymakers *thought* would lower taxes by 1% of GDP. In other words, this is an ex-ante concept of the multiplier for a 1% of GDP *intended* impulse. Panel (b) presents the result obtained by directly estimating the LP-IV specification in equation 3, which naturally takes into account the actual effect on tax receipts.

Figure 4 shows that the cumulative tax multiplier for GDP is large and persistent. Using the LP-IV approach, the cumulative multiplier for GDP rises to around 3 by the second year after the policy change. On the "static" tax base approach, the cumulative multiplier rises to around 2 in the second year. The figure also shows the 68 and 95 percent confidence intervals. These effects are statistically significant. For example, using two standard error bands the LP-IV cumulative multiplier is significant for almost the entire period. In terms of the IV first-stage, Appendix A5.1 also reports the usual weak instrument statistics. The Kleibergen-Paap statistics are larger than 10 at almost all horizons but the thresholds for these statistics are only valid for serially uncorrelated errors. We therefore also report the

stricter Montiel Olea and Pflueger (2013) F-statistics. The figure also shows the Montiel Olea and Pflueger (2013) thresholds for 10% and 20% of worst-case bias. For the peak multiplier, the F-statistic is well above these thresholds, and this is also the case at most other horizons.[35] [36]

The difference between the two panels in Figure 4 reflects several factors. First, given our narrative shocks are estimates of the projected tax revenue in a full year, the very short-term response of actual receipts may be muted.[37] Furthermore, the actual response of receipts may depend on the timing of when receipts were received and the revenue projections in the Budget may also be imperfect estimates of the actual changes in the tax yield (in the language of Mertens and Ravn (2013), the narrative shocks may be "proxies" rather than the true shocks). Second, the tax cut may have a positive effect on the tax base. This, in turn, may dampen any observed decline in tax revenues. The LP-IV approach therefore includes these dynamic feedback effects on the tax base and the LP-IV multiplier estimates should be larger than the static approach, which is indeed the case in Figure 4.[38] For all these reasons, we see both panels as informative. Panel (a) provides an estimate of the multiplier for an ex-ante intended change in taxes as a share of GDP. Panel (b) provides an estimate incorporating feedback effects.

Before ending this section several things are worth mentioning about the magnitudes. First, unlike the spending multipliers found by Crafts and Mills (2013, 2015) for interwar Britain, these GDP tax multipliers are well above 1. Of course, this does not necessarily imply that spending multipliers are always lower than tax multipliers as these papers focus on rearmament and study the effects using a different econometric approach. Second, our result that the interwar tax multiplier is larger than 1 is consistent with GDP tax multipliers found

[35] One exception is the impact effect where the Montiel Olea and Pflueger (2013) F-statistic is above the 30%, but below the 20% threshold. This is partly to be expected as our shock measures the expected change in tax liabilities over the coming year.

[36] The IRF for tax receipts as a share of GDP can also be seen as the response of the endogenous tax variable to the instrument, i.e. the first stage of an IV regression. This is shown later in Figure 5, where it can be seen that taxes as a share of GDP do indeed persistently decline over the IRF period.

[37] In particular, the response of GDP to a tax cut may occur immediately, but the rise in tax revenues may take time to fully materialize in the data. This would make the relationship between the narrative shock and revenues weaker at shorter horizons. This is another reason why it is interesting to present both the static and IV approaches together.

[38] The size of the "tax base" effect is likely to depend on the level of tax rates in the economy. The tax burden was considerably higher after World War II, which suggests these "tax base" effects could be even more relevant in more recent years. As Ramey (2019) notes, this can actually complicate computing the multiplier, especially if the tax revenue response ends up close to zero.

in post-WWII narrative evidence, for example by Romer and Romer (2010), Cloyne (2013) and Mertens and Ravn (2013). This suggests that tax cuts do indeed have sizable effects on the economy. Some narrative papers for the post-war period report impulse response functions rather than cumulative multipliers (Romer and Romer (2010), Cloyne (2013) and Mertens and Ravn (2013)). The larger magnitudes in Figure 3 are comparable to the post-WWII results in Romer and Romer (2010) (where the peak is just above 3) and Cloyne (2013) (where the peak is around 2.5). If anything, however, our results suggest slightly larger multipliers in interwar Britain. Formally comparing these periods seems a potentially interesting avenue for future work.

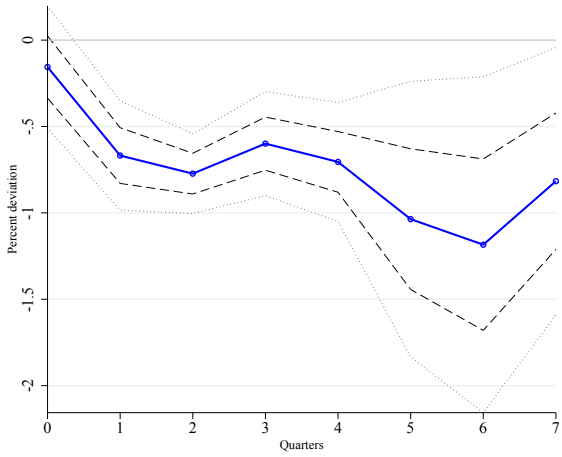
5 Further Results and Robustness

In this section we consider a number of extensions and robustness checks. Below we focus on the most important aspects, with further detail and exercises shown in the Appendix. In this section we consider: the interaction of taxes, spending and the deficit; the contribution of our tax shocks to the interwar business cycle; whether some of the tax reforms were anticipated; how to handle tax reforms with retroactive elements; the sensitivity of our results to particular events and to the sample period; the sensitivity of our results to different control variables; and, the response of other variables.

5.1 Taxes, government spending and the deficit

Fiscal multipliers are typically not structural parameters. Consequently, when interpreting multiplier estimates, an important question is: What's happening to the other aspects of fiscal policy? A tax cut that eventually leads to a fall in government spending should have different effects to a tax cut that is purely deficit-financed. Of course, this is an issue faced by the empirical literature in general but in our context it is particularly interesting given the apparent aversion to running large budget deficits during this period. Although we have excluded tax changes that are directly associated with spending changes, this does not rule out the possibility that government spending responds after our tax shocks occur. It is worth noting that, even if taxes are cut by the same amount as spending, this is still an interesting object but it is closer to a balanced budget multiplier than a purely deficit financed cut in taxes. In this section, we therefore investigate the relationship between taxes, spending and

Figure 5: Response of tax receipts relative to GDP to a -1% of GDP tax shock



Notes: The figure shows the effect of a -1% shock to our narrative tax series on actual tax receipts at horizon $t+h$ relative to GDP in $t-1$ for 8 quarters. This impulse response function is estimated using equation 1 in the text but where the dependent variable is $(T_{t+h} - T_{t-1})/Y_{t-1}$. Dashed and dotted lines represent one and two standard error bands computed using the Newey-West approach.

the deficit in more detail using the receipts and expenditure data from Lennard (2020).

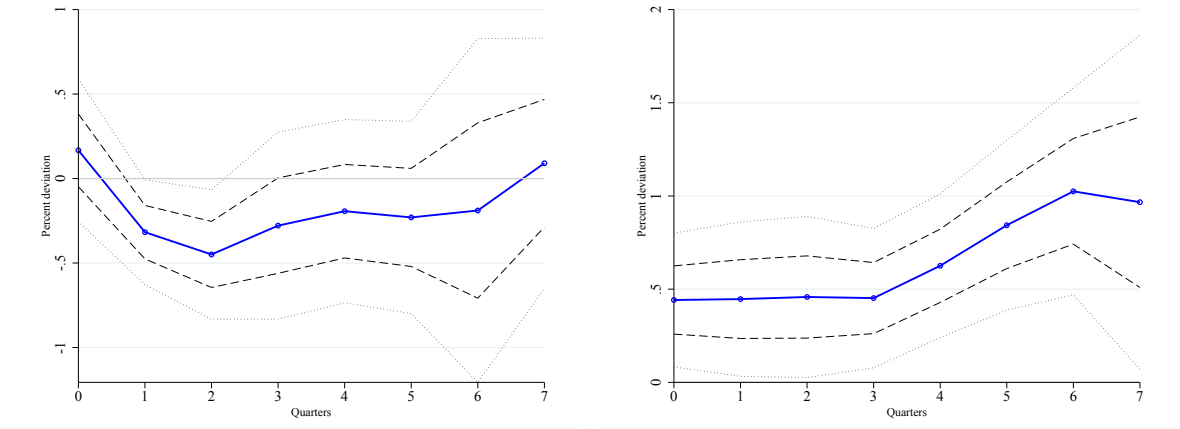
First we examine the response of tax revenues to our tax shock. This figure is produced by estimating our baseline local projection specification, equation 1, but where the dependent variable is $(T_{t+h} - T_{t-1})/Y_{t-1}$.[39] Figure 5 shows that our tax shock does indeed reduce tax receipts significantly.[40] As discussed in Section 4, the precise response of revenues depends on a number of factors. In particular, the precise response will depend on the timing of when tax receipts are collected, on whether the tax changes did indeed cost what was originally expected, whether other taxes were adjusted later, and on the dynamic feedbacks due to tax base effects. The tax base effects might be particularly strong when tax rates are high, although the tax burden was lower in this period relative to post-WWII.

Next, we analyze how government spending subsequently responds to a -1% narrative tax shock. This is shown in Figure 6a. Interestingly, while government spending as a share

[39] Consistent with the LP-IV discussion from earlier, we only use tax receipts data after 1920Q3 to exclude the initial outliers. This restricts the sample for these regressions by a few quarters. This makes little difference to the point estimates but improves the standard errors considerably at shorter horizons.

[40] The impulse response function for tax receipts as a share of GDP can also be seen as the response of the endogenous tax variable to the instrument, i.e. the first stage regression. As discussed in Ramey (2016) and Ramey and Zubairy (2018), with one instrument, the approach in equation 3 is equivalent to integrating the IRF for GDP and dividing by the response of revenues (as a share of GDP). See the earlier working paper version of this paper for results using a two-step approach.

Figure 6: Response of government spending and the fiscal deficit as a share GDP



(a) Government Spending (b) Deficit to GDP Ratio

Notes: The figure shows the effect of our tax shock on government spending as a share of initial GDP ($(G_{t+h} - G_{t-1})/Y_{t-1}$) and the deficit to GDP ratio for 8 quarters. Dashed and dotted lines represent one and two standard error bands computed using the Newey-West approach.

of GDP does decline in the first few quarters, this effect is quickly reversed. The effect over the whole period is more muted and much less significant than for taxes.[41] We could even be more agnostic and think of our shocks as an instrument for a more general fiscal shock. Figures 5 and 6a therefore tell us whether this is a largely tax-led or spending-led fiscal intervention. In short, this shock indeed looks like a tax cut. Although government spending does subsequently decline (which accords with the narrative account of the fiscal policy regime during the interwar years), these tax changes end up being more deficit financed. Figure 6b provides more direct evidence on this: following the tax cut the fiscal deficit rises significantly.

Figure 6a does, however, show some decline in government spending. Given this, our tax multiplier is not exactly a deficit financed multiplier and, as noted by Mountford and Uhlig (2009), this affects how we should interpret multiplier estimates. On the other hand, if we also had an estimate of the deficit financed government spending multiplier, we could get a sense of how our baseline results for the effects of taxes would change if government spending did not fall as in Figure 6a. Crafts and Mills (2013, 2015) provide estimates of the present value government spending multiplier in interwar Britain. One of their results

[41] For example, the average effect in Figure 6a is around -0.15%, relative to -0.75% for taxes in Figure 5.

is that this object is well below 1. We now conduct a back-of-the-envelope calculation to assess how the interpretation of our results might change if government spending were held fixed. To be conservative and make the calculations straightforward, assume that the deficit financed cumulative spending multiplier is always 1. Note that this is "conservative" because the fall in government spending offsets the positive effect of the tax cut on GDP. Given the path in Figure 6a, and these assumptions, it is straightforward to adjust Figure 4 and compute an approximate deficit financed tax multiplier. The full results of this exercise are shown in Appendix A4.5 for both the static approach and the IV approach. Given that the fall in government spending is relatively muted this adjustment does not materially affect the main conclusions above. The multiplier estimates are larger, as would be expected, but the adjusted estimates generally lie with the one standard error bands around the baseline results.

As a further robustness check, we also make use the Lennard (2020) data on government spending and tax receipts and add these as additional controls in our baseline regressions. As noted earlier, one concern might be that our shocks are correlated with past values of the deficit or government spending. The results are shown in Appendix A4.1. The first chart shows the effect of including total log (real) government spending. The second chart adds the deficit to GDP ratio. If our results were biased by the exclusion of other fiscal variables, this exercise should produce different results. The multiplier effects are, however, very close to the baseline.

5.2 Contribution to the business cycle

With all the variation in tax policy over this period, a natural question is: How important were these tax changes for the volatility of GDP over this period? Constructing variance decompositions in a local projections and an instrumental variables framework is not as straightforward as in the conventional Vector Autoregression literature. Two important recent contributions are by Gorodnichenko and Lee (2020) and Plagborg-Møller and Wolf (2022). In this section we implement these procedures.

Gorodnichenko and Lee (2020) develop a procedure for conducting a variance decomposition in a local projection framework where the narrative shock is included directly as a regressor. This is exactly relevant to the approach in equation 1 and the results in Figure

3. We therefore apply the Gorodnichenko and Lee (2020) procedure using the econometric specification in equation 1. Panel (a) in Appendix A5.2 reports the resulting forecast error variance decomposition for our fiscal shocks. This panel reports the Gorodnichenko and Lee (2020) preferred FEVD R^2-based metric, including a bias-correction that the authors recommend for smaller sample sizes. The figure shows that our tax shocks account for around 15% of the volatility of GDP at the peak.

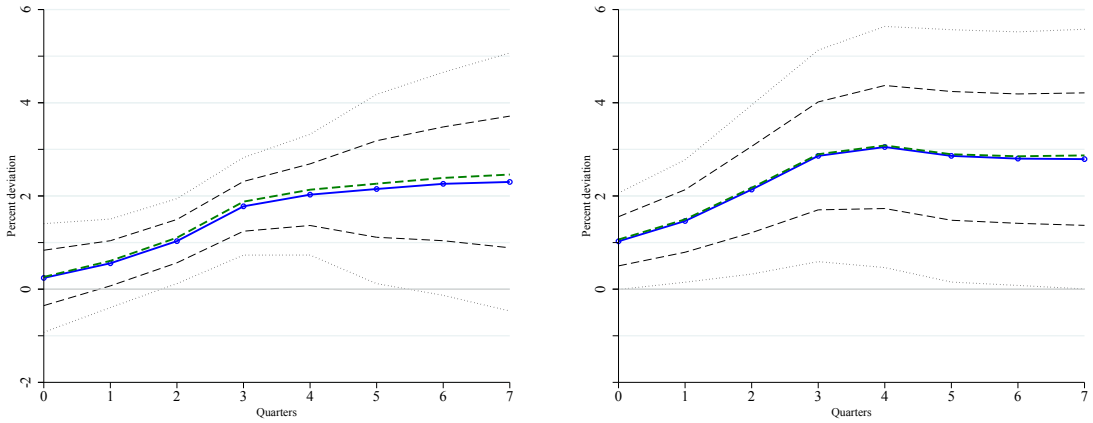
As in equation 1, the Gorodnichenko and Lee (2020) approach includes the narrative shock directly. An alternative approach is to consider the narrative shock as an instrument. Plagborg-Møller and Wolf (2022) show that when shocks are seen as instruments, the variance contribution is not point identified, although in theory the Gorodnichenko and Lee (2020) approach provides a lower bound. Plagborg-Møller and Wolf (2022) propose a Structural Vector Moving Average Instrumental Variable (SVMA-IV) approach for computing bounds for the variance decomposition. This involves estimating a reduced-form vector autoregression including a range of variables and the instrument itself. Bounds for the contribution of the shocks can then be computed from estimated moments. We now implement this approach with our shocks. We estimate a reduced form VAR including the usual variables in $\mathbf{X}$ and using the narrative shocks as an instrument for variation in tax receipts as a share of GDP. Panel (b) in Appendix A5.2 reports the two Plagborg-Møller and Wolf (2022) Forecast Variance Ratio bounds for the contribution of our tax shocks to GDP growth. The upper bound is around 20%, which is of a similar order of magnitude to the Gorodnichenko and Lee (2020) based results above.

5.3 Anticipation

Some tax changes were announced during the Budget speech but were implemented later in the fiscal year. One potential concern is that these types of tax changes were anticipated from the date of the speech. As a result, the impact of these particular reforms might have been felt before they were implemented. To tackle this possibility, we construct a new version of our narrative tax change series where "anticipated" tax changes are excluded. Following Mertens and Ravn (2012) we define an unanticipated tax reform as one that was implemented within 93 days of the announcement date.[42] In Appendix A6 we show the proportion of

[42]It is possible that the announcement itself might have been anticipated and GDP might then respond prior to that date. But, in this case the narrative shocks would then be predictable from past GDP realizations.

Figure 7: Cumulative GDP multiplier excluding anticipated tax changes



(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green/lighter dashed line overlays the point estimate from a robustness check where we exclude anticipated tax changes (defined as measures implemented more than 93 days after the announcement). As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy.

exogenous tax changes that were implemented within 93 days of their announcement date, both in terms of absolute number of changes and in terms of their value. On both definitions, the vast majority of exogenous tax reforms were actually "unanticipated". But, to explore the robustness of our findings, we exclude anticipated changes from our baseline series. The effect of this exercise is shown in Figure 7. The figure repeats the baseline results from above, but also includes the updated point estimate using the new series (the green/lighter dashed line). The green/lighter dashed line is extremely close to the baseline results, suggesting anticipation is not a major concern in this context.

5.4 Sensitivity to particular events and the sample period

Our narrative data set contains a number of sizable changes in taxes, as well as various smaller reforms. A natural concern is whether our results are all driven by a particular outlier. Appendix A4.9 examines this concern by re-estimating our baseline results for the cumulative multiplier excluding each of the largest shocks one at a time. We drop the shocks that occur in 1923Q2, 1924Q3, 1925Q2 and 1930Q2 in turn (see Figure 2). Appendix A4.9

In Section 3.2, however, we showed that this is not the case.

shows that, although removing specific events has some effect on the point estimates, in each of these cases the estimates are within one standard deviation of the original baseline result. This is true for both the static and IV approaches.

In constructing our narrative tax change dataset we were able to go back to the end of World War I. For the econometric analysis above, however, we were constrained by the availability of quarterly real GDP data which is only available from 1920 onwards. Annual real GDP data are, however, available from Thomas and Dimsdale (2017) much further back in time. We therefore conduct a further robustness check where we extend our quarterly real GDP series back to 1918 using annual GDP data prior to 1920. We do this using simple linear interpolation.[43]

The results are shown in Appendix A5.3. Note, because the tax receipts data only start in 1920, we are not able to apply the LP-IV approach. Again, although there is some variation across these experiments, the point estimates from the extended sample also lie within one standard deviation of the baseline result and therefore tell a very similar story.

5.5 Retroactive changes, timing of the shocks and an alternative classification

One issue that faces post-WWII narrative approaches is that some tax changes include retroactive components. This means that the change in tax liabilities is back-dated, so that the effective implementation date is prior to the announcement date. In the baseline specification we followed Romer and Romer (2010) and exclude retroactive components. The announcement date is then used as the implementation date for changes with retroactive elements. In this section, we now remove retroactive tax changes entirely from our data. Appendix A4.7 shows the results of this exercise.[44] The baseline result is again reported, with the updated point estimate from the new series overlaid (the green dashed line). Conclusions are very similar, again suggesting that retroactive tax reforms are not biasing our results. An additional timing issue is how to assign implementation dates to tax reforms that occur late in a quarter. As discussed above, we follow Romer and Romer (2010) and assign tax changes

[43] In practice this means, for example, that the overall annual change in real GDP between 1918 and 1919 is divided by 4 and applied sequentially to each quarter within the year to fill in the missing quarterly observations.

[44] Appendix A6 also shows that the vast majority of exogenous tax changes, both in terms of number and also total value, are not implemented retroactively.

that occur in the second half of a quarter to the next calendar quarter. In Appendix A4.6 we show that using the current quarter as the implementation date for all tax changes produces almost identical results. Finally, as noted in Section 2, although most Budgets are classified as packages, we have a minority of cases with separate classifications. As a robustness check we now classify all changes according to an overall motive for the Budget. Appendix A4.8 shows that the GDP multiplier is very similar.

5.6 Additional controls

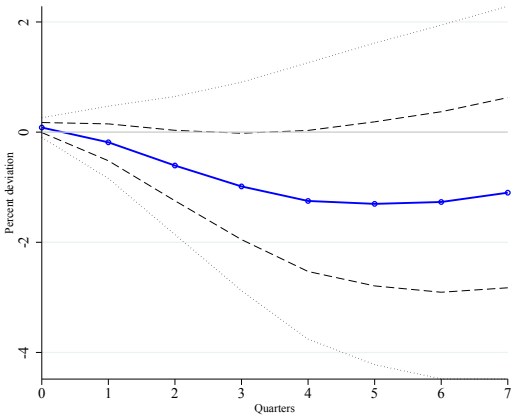
If our narrative identification strategy has been successful, changing the control set $\mathbf{X}$ in our regressions should not affect the baseline results. In smaller samples, there is, of course, still the possibility of chance correlations between variables, which is one reason for including the vector $\mathbf{X}$. In this section we explore the robustness of our findings to varying the variables included in $\mathbf{X}$. Appendix A4.1 considers a smaller specification with only the lags of GDP and tax shocks as controls, a medium sized specification with GDP, Bank rate and unemployment and a larger specification that includes wages. Although there is some variation across these experiments, the point estimates are largely within one standard deviation of the original point estimates. The size of $\mathbf{X}$ therefore does not alter our main conclusions of a sizable tax multiplier of 2-3 at the peak. Appendix A4.1 also reports the updated point estimates (green dashed lines) discussed above including lags of government spending and the fiscal deficit in $\mathbf{X}$.

5.7 Response of other variables

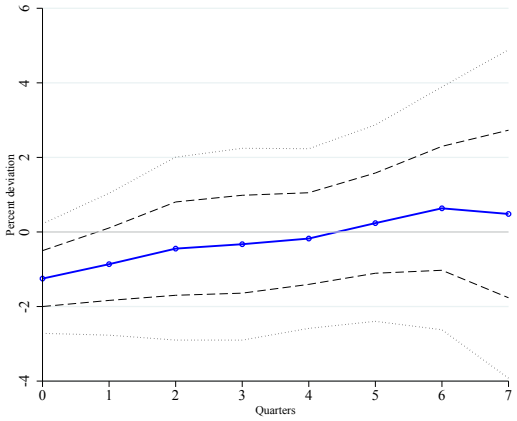
We now examine the impact of our exogenous tax changes on other variables. Figure 8 shows the counterpart of Figure 3 for the percentage point response of the unemployment rate and Bank Rate and the percentage response of the consumer price index. Unlike Figure 4, these figures show the impulse response functions as these variables are naturally expressed in percent, not £. Given the expansionary effects on GDP, it is unsurprising that the unemployment rate declines persistently following a tax cut, although the results for the unemployment rate are not statistically significant. Interestingly, the effect on prices also seems weak. Eventually prices rise, although the IRF is very imprecisely estimated. The

muted effect on prices might reflect the supply-side nature of many of the tax cuts.[45] The policy interest rate — Bank Rate — tends to increase when taxes are reduced.

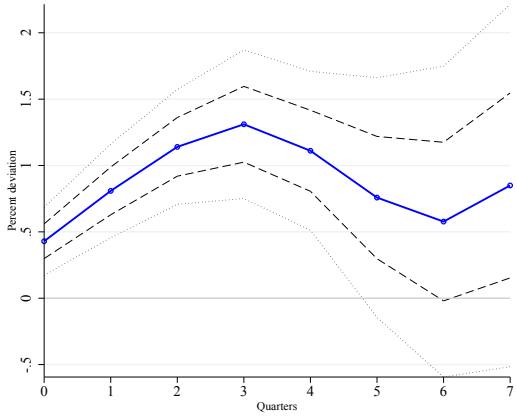
Figure 8: Response of the unemployment rate, the price level and Bank rate



(a) Unemployment rate (pp change)



(b) Price level (percentage change)



(c) Bank rate (pp change)

Notes: The figure shows the impulse response functions for the response of the unemployment rate (in pp), the percentage change in the price level, and the response of Bank Rate (in pp) for 8 quarters following a 1% cut in taxes as a share of GDP. Dashed and dotted lines represent one and two standard error bands.

6 Summary and conclusion

There has been much debate about whether changes in fiscal policy can affect the macroeconomy. Interwar Britain has always been a particularly contentious case and, given the high-debt, low interest rate environment (after 1931), it remains a particularly relevant case

[45] Some papers even find that tax cuts can lower prices, for example, see Mountford and Uhlig (2009).

today. Furthermore, despite the extensive use of tax policy for stabilization after WWII, estimates of the macroeconomic effects of tax changes are still somewhat scarce. We show that interwar Britain is an excellent environment for providing new evidence.

Keynes argued persistently from 1924 onwards in favour of fiscal expansion through increased public expenditure, most notably in *Can Lloyd George Do It?* a pamphlet written jointly with Hubert Henderson. His argument for fiscal expansion through a programme of public works was strengthened by Kahn (1931)'s development of the employment multiplier, which enabled the impact of public expenditure on employment to be quantified. As a result, much of the debate focused on the "spending multiplier". But, to our knowledge, evidence on the effects of tax changes in the interwar period is sparse. This is all the more remarkable given that tax policy formed a key part of the demand management toolkit after the Second World War.

Using extensive histographical research — an effort that we hope provides an interesting contribution in its own right — we construct a new dataset of tax changes for interwar Britain. Following the Romer and Romer (2010) approach, we identify a range of tax changes that can plausibly be regarded as exogenous and are strong candidates for evaluating the dynamic causal effect of tax changes on economic activity. In fact, with macroeconomic policy distinctly "pre-Keynesian" interwar Britain is particularly well-suited for this exercise.

Tax changes have large effects on GDP: a one percent of GDP cut in taxes raises GDP by 2-3 percent over 2 years. Although our estimates are sometimes a bit lower than Keynes's original multipliers, they are large relative to subsequent estimates of the expenditure multiplier for the interwar years. Our findings on the sizable effects of tax changes are, however, very consistent with narrative-based studies for the post-WWII period.

Our results suggest that tax changes had an important macroeconomic impact in interwar Britain, and that tax changes have the potential to generate sizable multipliers. Finally, we have provided a rich new dataset, and an extensive historical account of British interwar fiscal policy that should, we hope, provide a useful resource for future research.

Data Availability Statement

The data and code underlying this article is available on Zenodo at: https://doi.org/10.5281/zenodo.8097025.

References

Accominotti, Olivier. 2012. London Merchant Banks, the Central European Panic, and the Sterling Crisis of 1931. *The Journal of Economic History* 72(1): 1–43.

Alesina, A. 1988. The end of large public debts. In *High public debt : the Italian experience*, edited by Giavazzi, F., and L. Spaventa. Cambridge University Press.

Alesina, Alberto, and Silvia Ardagna. 2010. Large changes in fiscal policy: taxes versus spending. *Tax policy and the economy* 24(1): 35–68.

Alesina, Alberto, and Francesco Giavazzi. 2012. The austerity question: 'How' is as important as 'how much'. In *Austerity: Too Much of a Good Thing?*, edited by Corsetti, Giancarlo, 11–15.

Alford, B. W. E. 1972. *Depression and recovery ? : British economic growth, 1918-1939*. London: Macmillan.

Almunia, Miguel, Agustin Benetrix, Barry Eichengreen, Kevin H O'Rourke, and Gisela Rua. 2010. From Great Depression to Great Credit Crisis: Similarities, Differences and Lessons. *Economic policy* 25(62): 219–265.

Barnichon, Regis, and Christian Brownlees. 2019. Impulse Response Estimation by Smooth Local Projections. *The Review of Economics and Statistics* 101(3): 522–530.

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

Blanchard, Olivier, and Roberto Perotti. 2002. An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *Quarterly Journal of Economics* 117: 1329–1368.

Brittain, Sir H. 1959. *The British budgetary system*. London : Allen and Unwin.

Broadberry, Stephen. 1986. *The British Economy During the Wars*. Basil Blackwell.

Brown, W. 1929. The Conflict of Opinion and Economic Interest in England. In *The Gold Standard and Employment Policies Between the Wars*, edited by Pollard, Sydney. Methuen.

Buchanan, James M. 1985. The Moral Dimension of Debt Financing. *Economic Inquiry* 23(1): 1–6.

Cairncross, Alec, and Nita Watts. 1989. *The Economic Section : a study in economic advising*. London: Routledge.

Capie, F., and M. Collins. 1983. *The inter-war British economy: A statistical abstract*. Manchester University Press.

Capie, Forrest, and Geoffrey Wood. 2012. *Money over two centuries : selected topics in British monetary history*. Oxford University Press.

Cloyne, James, and Paolo Surico. 2017. Household Debt and the Dynamic Effects of Income Tax Changes. *Review of Economic Studies* 84(1): 45–81.

Cloyne, James S. 2013. Discretionary tax changes and the macroeconomy: new narrative evidence from the United Kingdom. *American Economic Review* 103(4): 1507–28.

Cloyne, James S., Nicholas Dimsdale, and Natacha Postel-Vinay. 2023. Tax Changes in the the United Kingdom, 1918-1938: A New Narrative Account and Dataset.

Crafts, Nicholas, and Terence C Mills. 2013. Rearmament to the rescue? New estimates of the impact of "Keynesian" policies in 1930s' Britain. *The Journal of Economic History* 73(4): 1077–1104.

Crafts, Nicholas, and Terence C Mills. 2015. Self-defeating austerity? Evidence from 1930s' Britain. *European Review of Economic History* 19(2): 109–127.

Cunliffe Committee. 1918. First interim report. *His Majesty's Stationary* .

Daunton, Martin. 2002. *Just taxes : the politics of taxation in Britain, 1914-1979*. Cambridge University Press.

Daunton, Martin. 2007. *Wealth and Welfare: An Economic and Social History of Britain, 1851-1951*. Oxford University Press.

Dimsdale, Nicholas H. 1981. British Monetary Policy and the Exchange Rate 1920-1938. *Oxford Economic Papers* 33: 306–349.

Dimsdale, Nicholas H. 1987. Keynes on British Budgetary Policy, 1914-1946. In *Private Saving and Public Debt*, edited by Boskin, Flemming John, Michael, and Stefano Gorini. Basil Blackwell.

Dimsdale, Nicholas H., and Nicholas Horsewood. 1995. Fiscal policy and employment in interwar Britain: Some evidence from a new model. *Oxford Economic Papers* 47: 369–396.

Dow, Christopher. 1964. *The management of the British economy 1945-60*. The Syndics of the Cambridge University Press.

Gorodnichenko, Yuriy, and Byoungchan Lee. 2020. Forecast Error Variance Decompositions with Local Projections. *Journal of Business and Economic Statistics* 38(4): 921–933.

Guajardo, Jaime, Daniel Leigh, and Andrea Pescatori. 2014. Expansionary austerity? International evidence. *Journal of the European Economic Association* 12(4): 949–968.

Gunter, Samara, Daniel Riera-Crichton, Carlos A. Vegh, and Guillermo Vuletin. 2021. Non-linear effects of tax changes on output: The role of the initial level of taxation. *Journal of International Economics* 131(C): 103450.

Hancock, K. 1970. The Reduction of Unemployment as a Problem of Public Policy, 1920-1929. In *The Gold Standard and Employment Policies Between the Wars*, edited by Pollard, Sydney. Methuen.

Hatton, Timothy J. 1987. The outlines of a Keynesian solution. In *The Road to Full Employment*. George Allen & Unwin London.

Hayo, Bernd, and Matthias Uhl. 2014. The macroeconomic effects of legislated tax changes in Germany. *Oxford Economic Papers* 66(2): 397–418.

Henderson, Hubert D, and JM Keynes. 1929. Can Lloyd George Do It. *The collected writings of John Maynard Keynes* 9: 86–125.

Hicks, Ursula Kathleen. 1938. *The finance of British government, 1920-1936*. London: Oxford University Press.

Hicks, Ursula Kathleen. 1953. The Budget as an Instrument of Policy, 1837-1953. *The Three Banks' Review* 16–34.

Hume, David. 1987. *Essays, moral, political, and literary.* Indianapolis: LibertyClassics, rev. ed. edition.

Hussain, Syed M., and Lin Liu. 2018. Comparing the effects of discretionary tax changes between the US and the UK. *The B.E. Journal of Macroeconomics* 18(1).

Jordà, Òscar. 2005. Estimation and inference of impulse responses by local projections. *American economic review* 95(1): 161–182.

Jordà, Òscar, and Alan M Taylor. 2015. The time for austerity: estimating the average treatment effect of fiscal policy. *The Economic Journal* 126(590): 219–255.

Kahn, Richard F. 1931. The relation of home investment to unemployment. *The Economic Journal* 41(162): 173–198.

Keynes, John Maynard. 1933. *The means to prosperity.* Macmillan.

Keynes, John Maynard. 1936. The Marginal Propensity to Consume and the Multiplier. In *The General Theory of Employment, Interest, and Money*, chapter 10. London: Macmillan.

Lennard, Jason. 2020. Uncertainty and the Great Slump. *The Economic History Review* 73(3): 844–867.

Locke, John. 1689. *Second Treatise of Government: An Essay Concerning the True Original, Extent and End of Civil Government.* John Wiley & Sons.

Mallet, Bernard, and Cecil George. 1933a. *British budgets : second series, 1913-1914 to 1920-21.* London: [s.n.].

Mallet, Bernard, and Cecil George. 1933b. *British budgets : third series, 1921-1922 to 1932-33.* London: [s.n.].

Matthew, H. C. G. 1986. *Gladstone 1809-1874.* Oxford: Clarendon.

Mertens, Karel, and Morten O. Ravn. 2012. Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated U.S. Tax Policy Shocks. *American Economic Journal: Economic Policy* 4(2): 145–181.

Mertens, Karel, and Morten 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, Karel, and Morten O. Ravn. 2014. A reconciliation of SVAR and narrative estimates of tax multipliers. *Journal of Monetary Economics* .

Middleton, Roger. 1985. *Towards the managed economy : Keynes, the Treasury and the fiscal policy debate of the 1930s.* Methuen.

Middleton, Roger. 1996. *Government Versus the Market: The Growth of the Public Sector, Economic Management and British Economic Performance, 1890-1979.* Edward Elgar.

Middleton, Roger. 2010. British Monetary and Fiscal Policy in the 1930s. *Oxford Review of Economic Policy* 26(3).

Mitchell, Brian R. 1988. *British Historical Statistics*. Cambridge: Cambridge University Press.

Mitchell, James, Solomos Solomou, and Martin Weale. 2012. Monthly GDP estimates for inter-war Britain. *Explorations in Economic History* 49(4): 543–556.

Moggridge, Donald Edward. 1969. *The return to gold, 1925: the formulation of economic policy and its critics*. Cambridge University Press.

Moggridge, Donald Edward. 1972. *British monetary policy, 1924-1931 : the Norman Conquest of ... 4.86 [dollars]*. Cambridge University Press.

Montiel Olea, José Luis, and Mikkel Plagborg-Møller. 2021. Local Projection Inference is Simpler and More Robust Than You Think. *Econometrica* 89(4): 1789–1823.

Montiel Olea, José Luis, and Carolin Pflueger. 2013. A Robust Test for Weak Instruments. *Journal of Business and Economic Statistics* 31(3): 358–369.

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

Newey, Whitney K., and Kenneth D. West. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55(3): 703–708.

Nguyen, Anh D. M., Luisanna Onnis, and Raffaele Rossi. 2021. The Macroeconomic Effects of Income and Consumption Tax Changes. *American Economic Journal: Economic Policy* 13(2): 439–66.

Peacock, Alan T., and Jack Wiseman. 1961. *The growth of public expenditure in the United Kingdom*. Princeton : London: Princeton University Press ; Oxford University Press.

Peden, G. C. 1987. *Keynes, the Treasury and British economic policy*. Basingstoke: Macmillan Education.

Peden, G. C. 2000. *The Treasury and British public policy, 1906-1959*. New York: Oxford University Press.

Plagborg-Moller, Mikkel, and Christian K. Wolf. 2021. Local Projections and VARs Estimate the Same Impulse Responses. *Econometrica* 89(2): 955–980.

Plagborg-Møller, Mikkel, and Christian K. Wolf. 2022. Instrumental Variable Identification of Dynamic Variance Decompositions. *Journal of Political Economy* 130(8): 2164–2202.

Pollard, Sydney. 1970. *The gold standard and employment policies between the Wars*. Methuen.

Ramey, Valerie, and Sarah Zubairy. 2018. Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data. *Journal of Political Economy* 126(2): 850–901.

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

Ramey, Valerie A. 2016. Chapter 2 - Macroeconomic Shocks and Their Propagation. volume 2 of *Handbook of Macroeconomics*, 71–162. Elsevier.

Ramey, Valerie A. 2019. Ten Years after the Financial Crisis: What Have We Learned from the Renaissance in Fiscal Research? *Journal of Economic Perspectives* 33(2): 89–114.

Ritschl, A. 1996. Sustainability of high public debt: what the historical record shows. *Swedish Economic Policy Review* (3).

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

Romer, Christina D., and David H. Romer. 2014. The Incentive Effects of Marginal Tax Rates: Evidence from the Interwar Era. *American Economic Journal: Economic Policy* 6(3): 242–81.

Self, Robert C. 2006. *Neville Chamberlain : a biography*. Ashgate.

Short, M. E. 1985. *The politics of personal taxation: budget-making in Britain, 1917-31*. Ph.D. thesis, University of Cambridge, Cambridge.

Smith, Adam. 1884[1976]. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Nelson and Sons.

Snowden, Philip Snowden. 1920. *Labour and national finance,*. London: L. Parsons.

Stationary Office. 1918-1938. Financial Statement.

Stationary Office. 1997. *Annual abstract of statistics (Great Britain. Office for National Statistics : Online)*. London: Stationary Office.

Stock, James H., and Mark W. Watson. 2018. Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments. *The Economic Journal* 128(610): 917–948.

Tenreyro, Silvana, and Gregory Thwaites. 2016. Pushing on a String: US Monetary Policy Is Less Powerful in Recessions. *American Economic Journal: Macroeconomics* 8(4): 43–74.

Thomas, Mark. 1983. Rearmament and economic recovery in the late 1930s. *The Economic History Review* 36(4): 552–579.

Thomas, Ryland, and Nicholas Dimsdale. 2017. "A Millennium of UK Data", Bank of England OBRA dataset. https://www.bankofengland.co.uk/statistics/research-datasets.

Thomas, Terry. 1981. Aggregate demand in the United Kingdom 1918-45. In *The Economic History of Britain since 1700*, volume 2. Cambridge University Press.

Tomlinson, Jim. 1990. *Public policy and the economy since 1900*. Oxford: Oxford University Press.

Uhlig, Harald. 2010. Some Fiscal Calculus. *American Economic Review* 100(2): 30–34.

UK Parliament. 1918-1938. Hansard (Official Parliamentary Record).

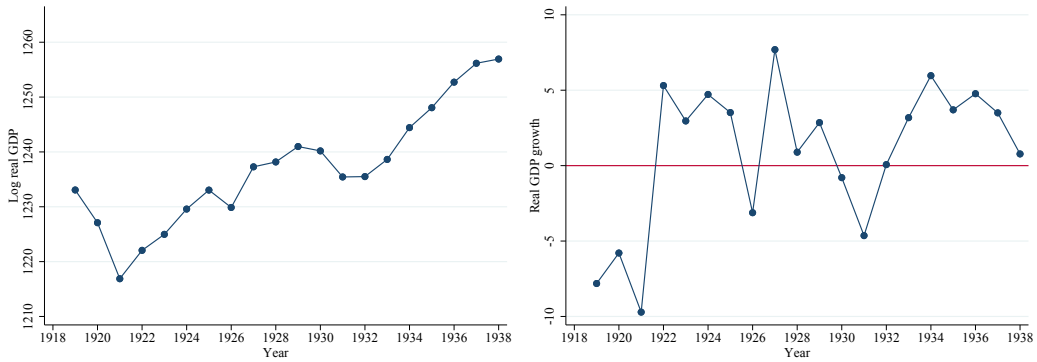
ONLINE APPENDIX

Taxes and Growth: New Evidence from Interwar Britain

James Cloyne, Nicholas Dimsdale and Natacha Postel-Vinay

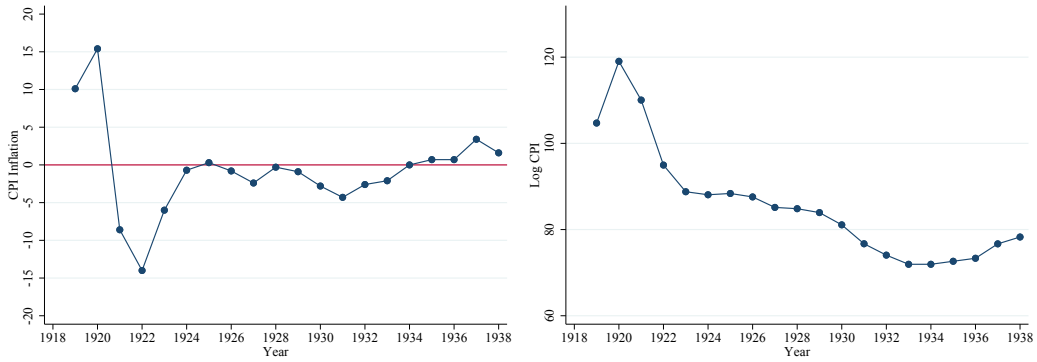
A1 Macro trends 1919-1940

Figure A.1: Gross Domestic Product



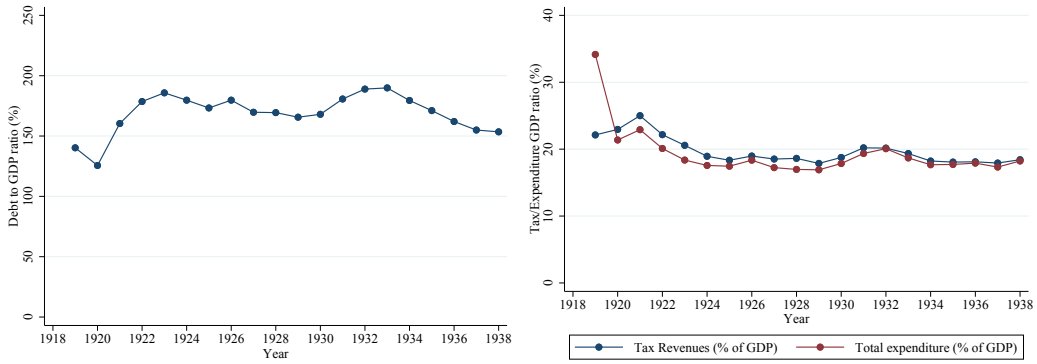
Notes: Panel A: log GDP, Panel B: GDP growth

Figure A.2: Consumer Prices Index



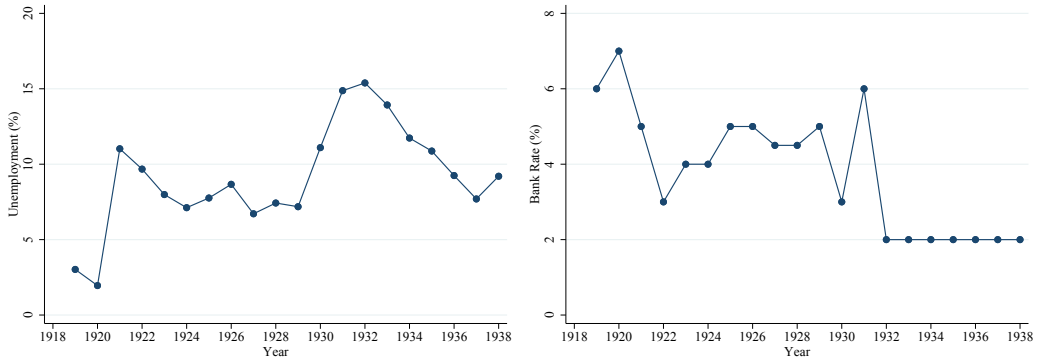
Notes: Panel A: CPI Inflation, Panel B: log CPI

Figure A.3: Debt, taxes and spending as a share of GDP



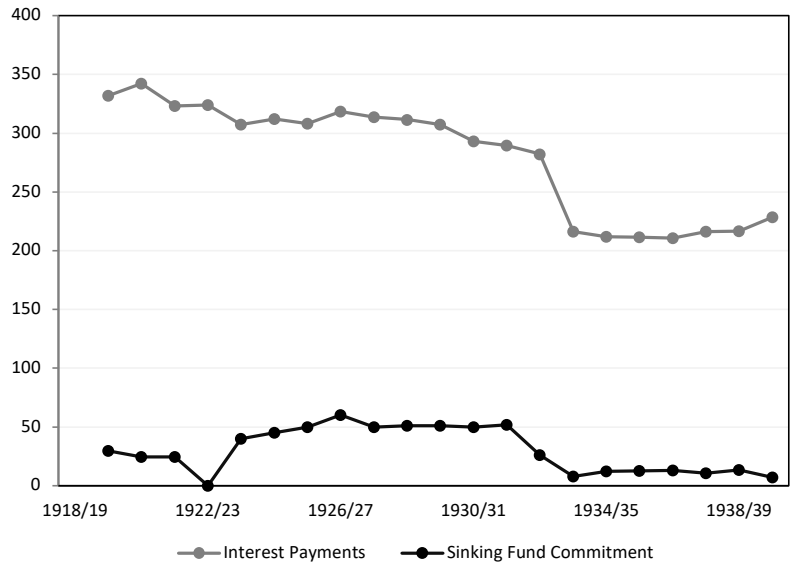
Notes: Panel A: Debt to GDP ratio, Panel B: Taxes and government spending as a share of GDP

Figure A.4: Unemployment Rate and Bank Rate (percent)



Notes: Panel A: Unemployment Rate, Panel B: Bank Rate

Figure A.5: Interest payments and sinking fund commitments (in millions of pounds), 1919-1939



Notes: Interest payments are from Mallet and George (1933a) and Mallet and George (1933b). Sinking fund commitments are from those same sources and Peden (2000). Actual sinking fund payments collected at the end of the financial year sometimes differed from the amounts committed, such as in 1922/23 and 1925/26. For our purposes only amounts committed matter (this is similar to the distinction between tax change and tax revenue.)

A2 British Chancellors of the Exchequer

Table A.1: British Chancellors of the Exchequer and their affiliation

Time period	Chancellor (Party)	Prime Minister (Party)
1918-1919	Bonar Law (Conservative)	David Lloyd George (Liberal-Conservative coalition)
1919-1921	Austen Chamberlain (Conservative)	David Lloyd George (Liberal-Conservative coalition)
1921-1923	Robert Horne (Conservative)	David Lloyd George (Liberal-Conservative coalition)
1923-1924	Stanley Baldwin (Conservative)	Andrew Bonar Law (Conservative)
1924-1925	Philip Snowden (Labour)	Ramsay MacDonald (Labour)
1925-1929	Winston Churchill (Conservative)	Stanley Baldwin (Conservative)
1930-1931	Philip Snowden (Labour)	Ramsey MacDonald (Labour)
1932-1937	Neville Chamberlain (Conservative)	Stanley Baldwin (Cross-party coalition)
1938-	John Simon (Liberal National)	Neville Chamberlain (Cross-party coalition)

Note: Cross-party coalition means that more than two-parties are in power.

A3 Macro data sources

Table A.2: Macro data sources

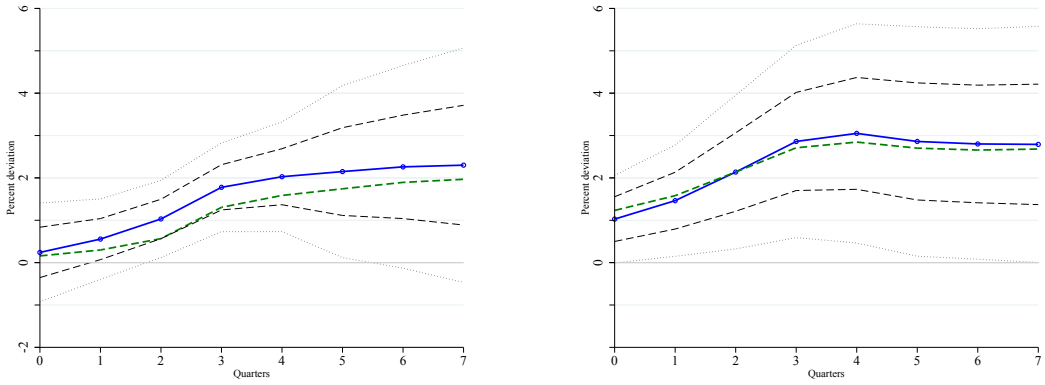
Variables	Source(s)
Quarterly Data	
Narrative tax changes	*Financial Statements* Stationary Office (1918-1938) *Hansard* UK Parliament (1918-1938) Finance Act(s) https://www.legislation.gov.uk/.
Real GDP	Thomas and Dimsdale (2017)
Bank Rate	Thomas and Dimsdale (2017)
Wages	Thomas and Dimsdale (2017), Avg. Weekly Earnings
Consumer prices	Thomas and Dimsdale (2017), Consumer Price Index seasonally adjusted using X-13ARIMA-SEATS.
Unemployment	Thomas and Dimsdale (2017), monthly unemployment rate, quarterly average.
Government expenditure and receipts	Lennard (2020). Seasonally adjusted using a 4-quarter moving average filter.
GDP deflator	Weighted average of the cost of living index and the export price deflator, Capie and Collins (1983)
Annual Data	
Real GDP	Thomas and Dimsdale (2017)
GDP deflator	Thomas and Dimsdale (2017)
Nominal GDP	Thomas and Dimsdale (2017)
Consumer prices	Thomas and Dimsdale (2017)
Debt to GDP ratio	Thomas and Dimsdale (2017)
Bank Rate	Thomas and Dimsdale (2017)
Unemployment Rate	Thomas and Dimsdale (2017)
Government spending to GDP ratio	Thomas and Dimsdale (2017)
Tax receipts to GDP ratio	Thomas and Dimsdale (2017)
Deficit as a share of GDP (Figure 1)	Thomas and Dimsdale (2017), Mallet and George (1933a), Peden (2000), Middleton (1985), Mitchell (1988), Mitchell et al. (2012)

A detailed list of all these sources, including precise data locations and codes, is contained in the replication package README file, available: https://doi.org/10.5281/zenodo.8097025.

A4 Robustness exercises

A4.1 Varying the controls

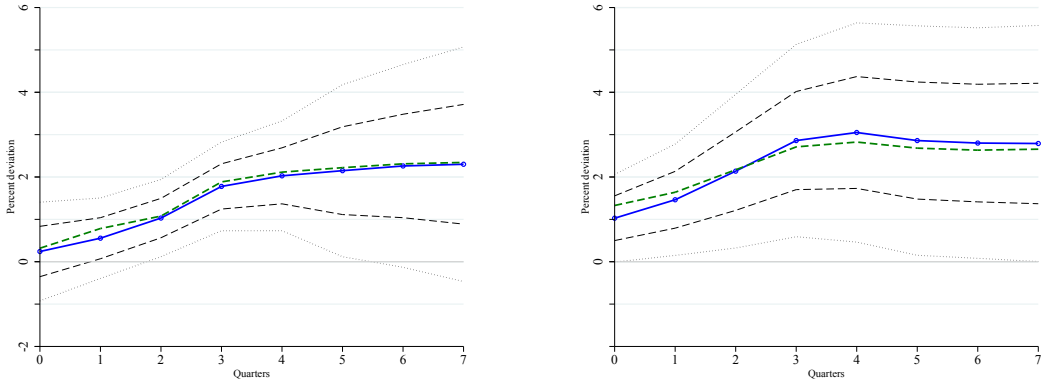
Figure A.6: Cumulative GDP multiplier: controlling for government spending



(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green dashed line overlays the point estimate from a robustness check where we also control for lagged government spending. As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy.

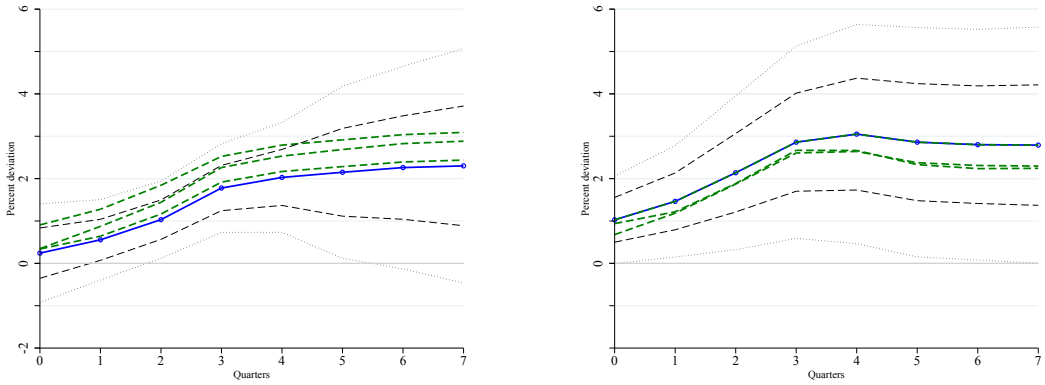
Figure A.7: Cumulative GDP multiplier: controlling for the deficit to GDP ratio



(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green dashed line overlays the point estimate from a robustness check where we also control for lagged deficit to GDP ratio. As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy.

Figure A.8: Cumulative GDP multiplier: varying the baseline controls

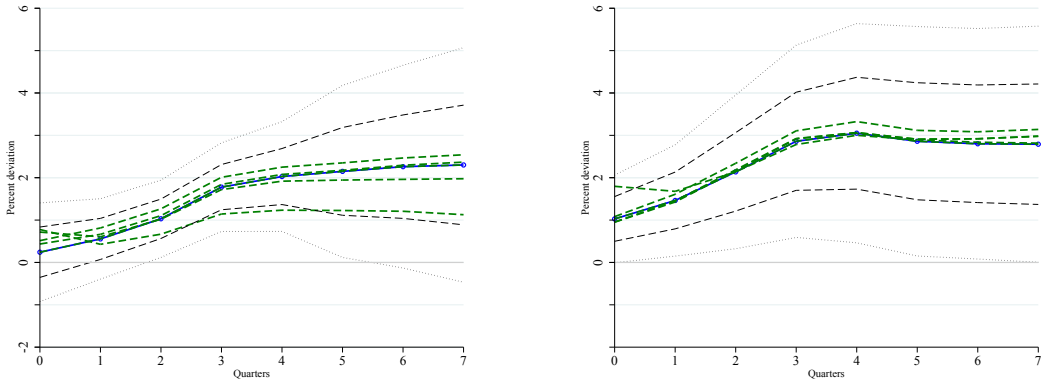


(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green dashed lines overlay the point estimates from robustness checks where the model is re-estimated (a) only with lagged real GDP and tax shocks; (b) including as controls: GDP, unemployment, and Bank Rate and c) including wages in addition to all other controls. As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy.

A4.2 Lag length sensitivity

Figure A.9: Sensitivity of the GDP response to different choices of lag length

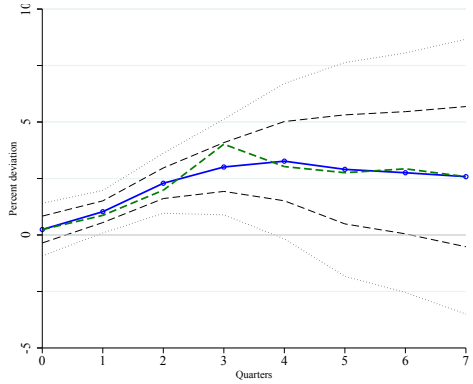


(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green dashed lines overlay the point estimate from a robustness check where we vary the different choices of the lag length P and Q. This chart shows the P, Q pairs $(1, 2)$, $(2, 2)$, $(1, 4)$, $(4, 4)$ (note that our baseline, in blue, is $1, 4$). As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy.

A4.3 Raw vs Smoothed Results

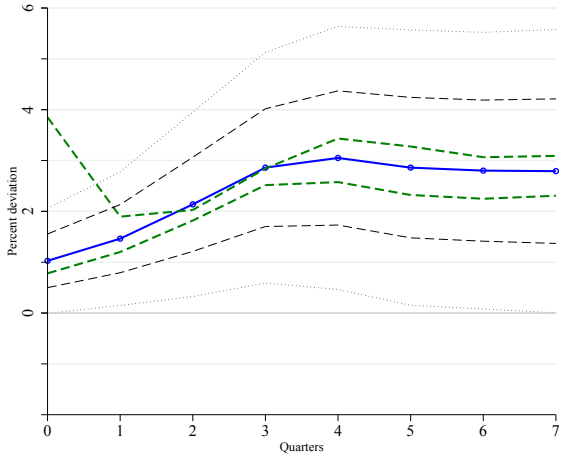
Figure A.10: Response of GDP (%): Smoothed vs. Raw IRF



Notes: This figure repeats the baseline impulse response function (in blue) and the associated standard error bands (black dashed/dotted lines) from Figure 3. The green dashed line shows the raw unsmoothed version of the IRF.

A4.4 LP-IV robustness

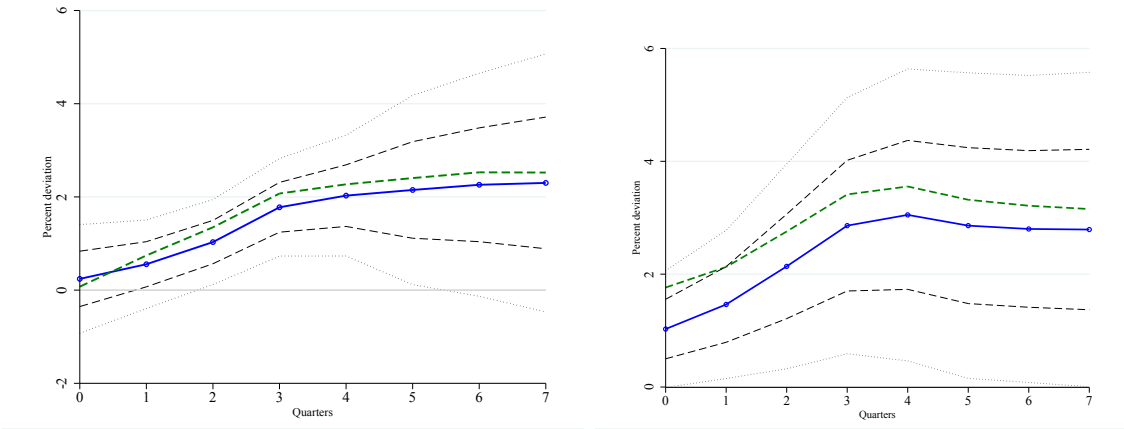
Figure A.11: Varying the number of instruments and the controls



Notes: The blue line and the dashed/dotted lines repeat the LP-IV cumulative multiplier and associated one and two standard error bands from Figure 4. The green lines show the point estimates from two robustness exercises: only using the contemporaneous value of the narrative shock as an instrument and including unemployment, Bank rate, log prices and government spending as additional controls in X.

A4.5 Adjusting for government spending

Figure A.12: Cumulative GDP multiplier adjusted for government spending

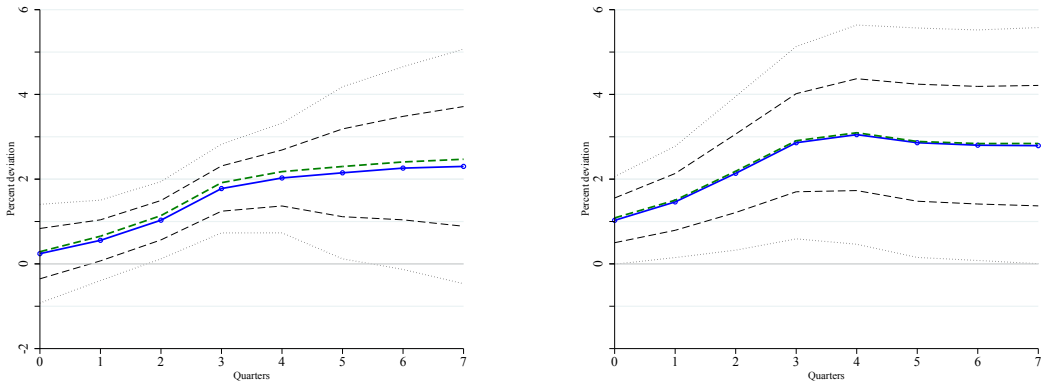


(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green dashed lines overlay the point estimate from a robustness check where we adjust the multiplier for the response of government spending in Figure 6a. See discussion in the text for more detail. As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy.

A4.6 Sensitivity to the timing of the shocks

Figure A.13: Cumulative GDP multiplier: different implementation dates

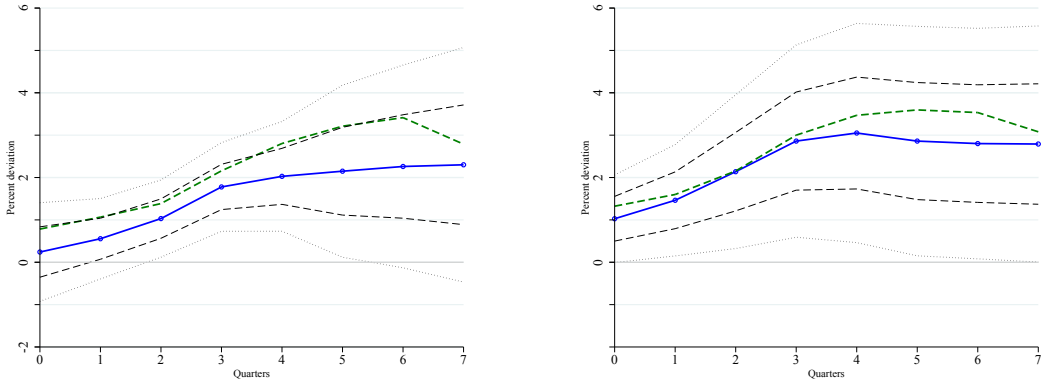


(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green dashed line overlays the point estimate from a robustness check where tax shocks are assigned to the calendar quarter based on their precise implementation date. As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy.

A4.7 Retroactive changes

Figure A.14: Cumulative GDP multiplier: excluding retroactive tax changes

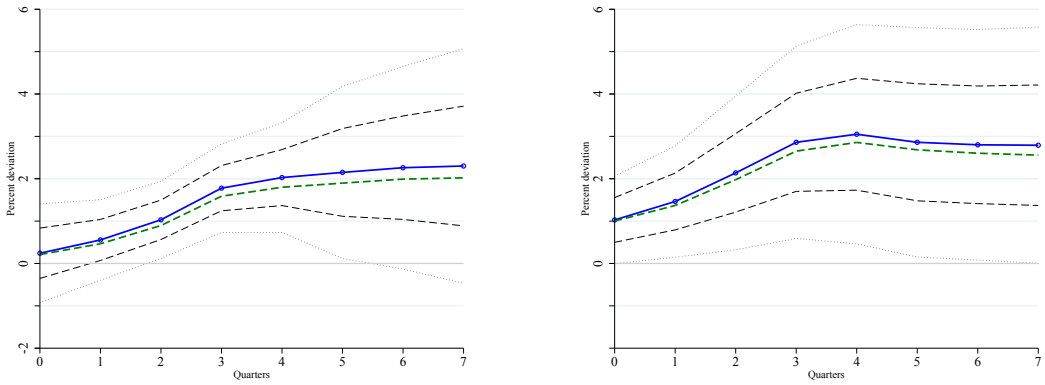


(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green dashed line overlays the point estimate from a robustness check where we completely exclude tax changes with a retroactive component. As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy.

A4.8 Alternative narrative classification

Figure A.15: Cumulative GDP multiplier using the alternative classification

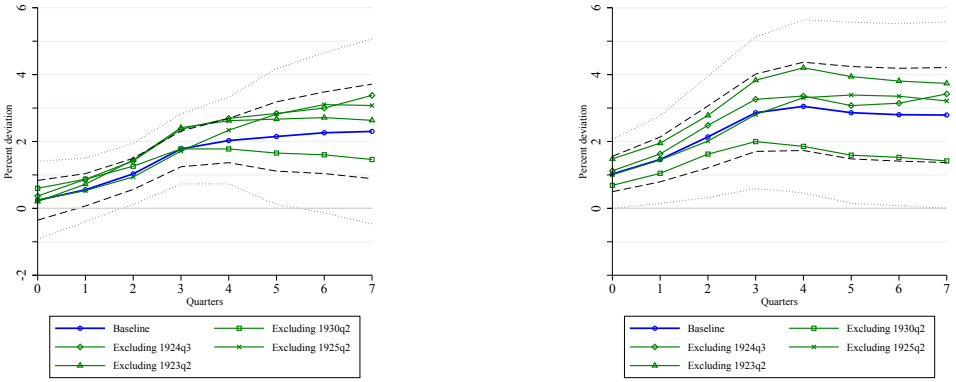


(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green dashed line overlays the point estimate from a robustness check where we classify all tax changes within a Budget together based on an overall motive. As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy.

A4.9 Sensitivity to major events

Figure A.16: Excluding Specific Shocks



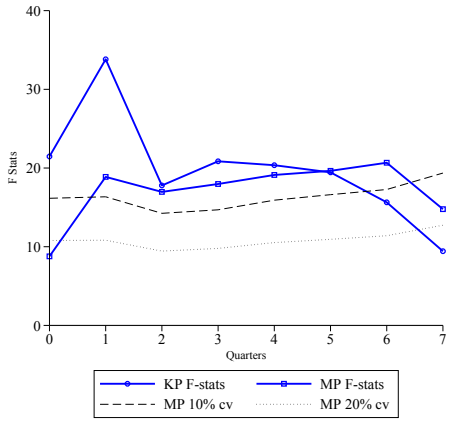
(a) Baseline "Static" Approach (b) IV Approach

Notes: The blue line and the black dashed/dotted lines repeat the baseline results for the cumulative fiscal multiplier and the associated one and two standard error bands from Figure 4. The green dashed lines show how the point estimates are affected by excluding the largest shocks in the sample one at a time. As in Figure 4, Panel (a) assumes taxes are reduced as a share of GDP by 1%. Panel (b) is computed using the LP-IV strategy

A5 Further results

A5.1 Weak instrument test results

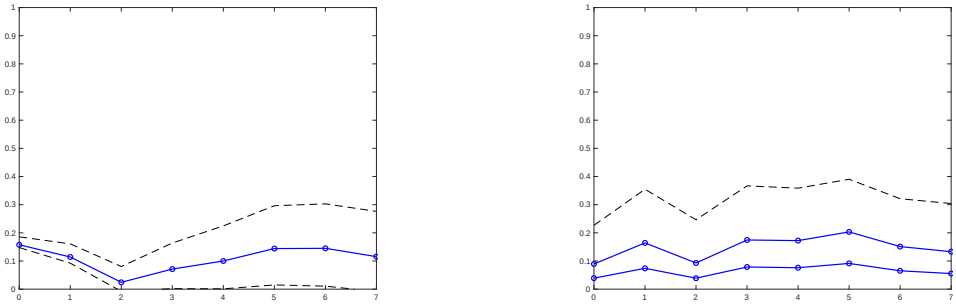
Figure A.17: Weak instrument test results



Notes: The figure shows the results of several weak instrument tests: the Kleibergen-Paap (KP) F statistics and those using the approach of Montiel Olea and Pflueger (2013) (MP). Dashed lines report the Montiel Olea and Pflueger (2013) (MP) thresholds for 10% and 20% of worst case bias at each horizon. Kleibergen-Paap F-stats are usually expected to be above 10, although assume serially uncorrelated errors.

A5.2 Contribution of the tax shocks

Figure A.18: Contribution of the Narrative Tax Shocks to GDP Volatility



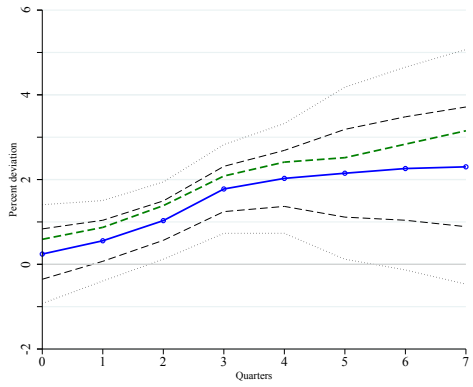
(a) Local Projection-based FEVD Gorodnichenko and Lee (2020)

(b) Forecast Variance Ratio SVMA-IV Plagborg-Møller and Wolf (2022)

Notes: The figure shows the contribution to GDP volatility of the narrative identified tax reforms. Panel (a) presents the results using the method of Gorodnichenko and Lee (2020) which is based on a regression specification similar to equation 1 in the text. Panel (b) presents the results using the method of Plagborg-Møller and Wolf (2022) where the narrative shock is used as an instrument. This method produces bounds, shown in blue. Dashed lines show 90% confidence bands. See text for details.

A5.3 Including data back to 1918

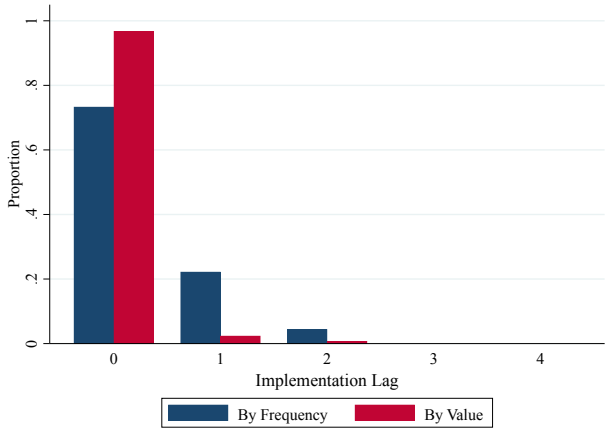
Figure A.19: Extending back to 1918



Notes: This figure repeats the main results for the cumulative fiscal multiplier from Figure 4 Panel (a) in blue together with the associated one and two standard error bands. The green dashed line reports how the point estimate changes when we extend the sample back to 1918. Annual GDP is interpolated to fill the quarterly series prior to 1920. The LP-IV version cannot be computed as we do not have tax receipts data prior to 1920.

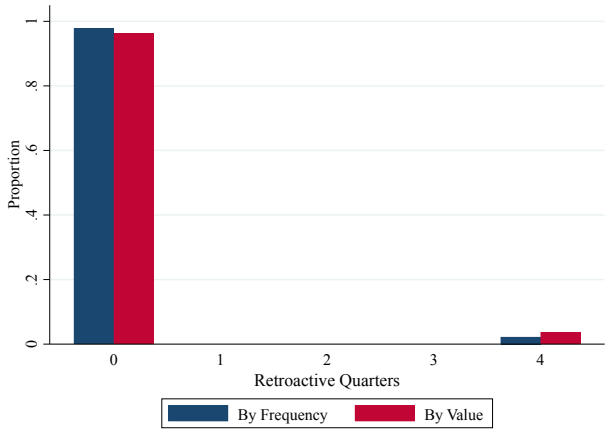
A6 Implementation lags and retroactive changes

Figure A.20: Proportion of exogenous tax changes by implementation lag



Notes: This figure shows the proportion of exogenous tax changes by their implementation lag in multiples of 93 days since their announcement date. The blue bars calculate this proportion using the frequency of individual tax changes. The red bars calculate the proportion based on the total revenue effect of these measures.

Figure A.21: Proportion of exogenous tax changes implemented retroactively



Notes: This figure shows the proportion of exogenous tax changes by their retroactive implementation quarter. Zero refers to a measure that was implemented within or after the announcement date. 1, ..., 4 means that the implementation date is 1, ..., 4 quarters earlier. The blue bars calculate this proportion using the frequency of individual tax changes. The red bars calculate the proportion based on the total revenue effect of these measures.