UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Cambridge Working Papers in Economics

Cambridge Working Papers in Economics: 2008

FRANCHISE EXTENSION AND FISCAL STRUCTURE IN THE UNITED KINGDOM 1820-1913: A NEW TEST OF THE REDISTRIBUTION HYPOTHESIS

Toke S Aidt Stanley L. Winer Peng Zhang

9 February 2020

We study the effect of franchise extension on the fiscal structure of central and local governments in the United Kingdom between 1820 and 1913 to revisit the Redistribution Hypothesis - the prediction that franchise extension causes an increase in state-sponsored redistribution. We adopt a novel method of uncovering causality from non-experimental data proposed by Hoover (2001). This method is based on tests for structural breaks in the marginal and conditional distributions of the franchise and fiscal structure time series preceded by a detailed historical narrative analysis. We do not find any compelling evidence that supports the Redistribution Hypothesis.

Franchise extension and fiscal structure in the United Kingdom 1820-1913: A new test of the Redistribution Hypothesis*

Toke Aidt
University of Cambridge *and* CESifo, Munich†
Stanley L. Winer
Carleton University, Ottawa-Carleton Graduate School of Economics
and CESifo, Munich‡
Peng Zhang
Simon Fraser University§

February 9, 2020

Abstract

We study the effect of franchise extension on the fiscal structure of central and local governments in the United Kingdom between 1820 and 1913 to revisit the Redistribution Hypothesis - the prediction that franchise extension causes an increase in state-sponsored redistribution. We adopt a novel method of uncovering causality from non-experimental data proposed by Hoover (2001). This method is based on tests for structural breaks in the marginal and conditional distributions of the franchise and fiscal structure time series preceded by a detailed historical narrative analysis. We do not find any compelling evidence that supports the Redistribution Hypothesis.

Keywords: Franchise extension; redistribution; democratization; causality; structural breaks; local government, central government, historical narrative.
JEL codes: D72

†Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, CB3 9DD Cambridge, UK; email: tsa23@econ.cam.ac.uk.

‡School of Public Policy and Department of Economics, Carleton University, Ottawa, Canada; email: stanley.winer@carleton.ca.

§Beedie School of Business, Simon Fraser University, Vancouver, BC Canada; email: peng_ zhang_ 7@sfu.ca.

*This research was made possible by a grant from the Keynes Fund (Cambridge) (JHOA). Constructive comments from Kevin Hoover, Tom Groll, and Jean Lacroix and from participants in the European Public Choice Society meeting in Rome 2018, the EEA/ES meeting in Manchester 2019, the International Institute of Public Finance Congress 2019 in Glasgow and in seminars at University of York and Cambridge are greatly appreciated.

1 Introduction

Almost all political economy models of voting and public policy contain within them the Redistribution Hypothesis - the prediction that enlargement of the voting franchise to poorer citizens will lead to demands for more state-sponsored redistribution which are at least partly satisfied by politicians elected on the basis of the now broader suffrage. The Redistribution Hypothesis is most clearly expressed in the work of Meltzer and Richard (1981, 1983).[1] However, it has proved difficult to establish empirically a causal, positive relationship between extension of voting rights and the extent of redistribution through the public sector. In this paper, we adopt a method proposed by Hoover (2001) to provide new evidence about this fundamental issue in empirical political economy. This method is based on a combination of detailed historical narrative analysis and study of the pattern of structural breaks in the conditional *and* marginal distributions of the time series representing the franchise and fiscal structure. We apply the method to study central and local government actions in the United Kingdom over the long nineteenth century from 1820 to 1913.

The dominant approach in the empirical literature is to investigate cross country panel data and to test if the conditional mean of the relevant fiscal policy variables - total spending or taxation relative to GDP, measures of tax or expenditure structure, and marginal tax rates - react in the predicted way to episodes of democratization.[2] Acemoglu et al. (2015, p. 1885) provide a comprehensive survey of this literature and conclude that "democracy does not lead to a uniform decline in post-tax inequality, but can result in changes in fiscal redistribution and economic structure that have ambiguous effects on inequality." Some of these fiscal adjustments are consistent with the Redistribution Hypothesis (e.g., Lindert (1994) or Aidt et al. (2006)) but others are not (e.g., Aidt and Jensen (2009b) or Profeta et al. (2013)).

The main weakness of the empirical literature on the Redistribution Hypothesis is that it struggles to isolate causal effects. While interesting and suggestive correlations have been uncovered, they can only be given a causal interpretation if the assumption of conditional independence is tenable. The fact that the relationship between democratization and the evolution of the fiscal system is complex and interwoven with structural transformation, industrialization, urbanization, and economic growth inevitably casts doubt

[1] Meltzer and Richard (1981) build on the median voter model which is not ideal for thinking about complex fiscal systems with many policy dimensions. Hettich and Winer (1999) and Tridimas and Winer (2005), however, show that the redistribution hypothesis holds within the context of the more appropriate probabilistic voting model. There are, of course, many limits to redistribution (e.g., Corneo and Gruner 2000; Harms and Zink 2003; Seghezza and Morelli 2019; Piketty 1995). Winer (2019) provides further discussion and an upto date survey of related literature.

[2] Redistribution can also take place via manipulation of non-fiscal policy instruments such as labour market and industry regulation (see, e.g., Acemoglu 2006). We do not consider such avenues.

on this assumption. While suffrage reforms may cause changes in the fiscal system, as hypothesized by Meltzer and Richard (1981) and integrated into the theory of franchise extension developed by Acemoglu and Robinson (2000), it is also possible that demands on the fiscal system - e.g., a desire to tap into tax sources that require a high degree of voluntary compliance which can be ensured by sharing voting rights - cause suffrage reforms. It is, of course, also possible that the two processes are caused by the same underlying forces, such as enlightenment and the spread of ideas about the value of a liberal political system and the benefit of state-sponsored social insurance, social welfare and progressive taxation.

In this paper, we consider the issue of the causal order from a new perspective. In a sequence of papers (Hoover 1991; Hoover and Sheffrin 1992; Hoover and Siegler 2000) and in a subsequent book (Hoover 2001), Kevin Hoover and his co-authors develop the approach that we adopt.[3] This approach to uncovering causality combines a detailed historical narrative with structural breaks econometrics in order to determine the causal order. The main innovation in the Hoover approach is to study the evolution over time of *both* the marginal and the conditional distributions involved, and to look for patterns of breaks in the underlying time series that then establish the causal order. The historical narrative provides an important cross-check on the tests for structural breaks - a break identified statistically when no relevant intervention can be identified historically may indicate statistical misspecification. The method can uncover the direction of causality, but it cannot identify the size of any causal effect.

To get a sense of the underlying logic, consider the following example. We want to know if a widening of the suffrage causes a change in fiscal structure in a manner likely to be associated with more redistribution, or if the two are jointly determined, independent, or if it is, perhaps, the other way around, namely that changes in fiscal structure are what drive franchise extension. Let e and f represent two time series of measures of the suffrage and of fiscal redistribution, respectively. Denote the joint, marginal and conditional distributions of the two series by $D(f, e)$, $D(e)$, $D(f)$, $D(f \mid e)$ and $D(e \mid f)$, respectively. Let us assume that it is, in fact, e that causes f. Now, suppose that we have prior (non-statistical) historical knowledge that there was an intervention that affected the franchise series but not the fiscal structure series in some year. It could, for example, be a major reform of the franchise. This naturally leads to a break in $D(e)$ and in $D(e \mid f)$. It also leads to a break in $D(f)$, but *not* under the maintained assumption in $D(f \mid e)$. This combination of structural breaks joined with prior historical knowledge that is consistent with them enables us to conclude that in that particular year, the extension

[3] Hoover (1991) studies the causal relationship between inflation and money growth, while Hoover and Sheffrin (1992) and Hoover and Siegler (2000) study the relationship between taxes and public spending.

of suffrage must have caused the observed change in the fiscal system. Alternatively, we might for some other year have historical knowledge of an intervention in the fiscal structure series. It could, for example, be a major tax collection innovation. In this case, we expect a break in $D(f)$ and in $D(f|e)$. The intervention would also lead to a break in $D(e|f)$, but *not* in $D(e)$ because of the maintained assumption that it is e that causes f. If we were to find such a pattern, we can again conclude, for that particular year, that it is e that causes f and not the other way around. It is clear from this that evidence of a causal relationship consistent with the Redistribution Hypothesis can come from two sources: if in a year where there is historical evidence of an intervention in the franchise (say, a reform) all distributions but $D(f|e)$ exhibit a statistical structural break *or* if in a year where there is historical evidence of an intervention in the fiscal system (say, a tax collection intervention) all distributions but $D(e)$ exhibit a statistical structural break, then e causes f. Importantly, however, compelling evidence for this (and other causal orders) requires that *both* of these "tests" are passed for some year, i.e., the evidence must come from interventions in both the suffrage and in the fiscal series. The novelty of the approach is that the causal order is derived from the patterns revealed by *both* conditional and marginal distributions of the time series that describe the fiscal and the suffrage processes.

With the Hoover method, the threat to causal inference comes from omitted factors unrelated to the Redistribution Hypothesis that cause breaks in e or f. However, as we shall see, to be confused with evidence of the Redistribution Hypothesis, these omitted factors need to generate breaks either in all distributions but $D(f|e)$ or in all distributions but $D(e)$. To minimize the risk of this, prior to any statistical testing the data are filtered for factors that could induce breaks in e and/or f for reasons unrelated to the Redistribution Hypothesis. The method, then, requires that after pre-filtering there are no remaining omitted factors that cause breaks in e or f. This is a different (and arguably weaker) assumption than the conditional independence assumption required for causal interpretation of a standard OLS regression. The Hoover method, however, does not simply rest on this assumption, as we have noted. The aim of the historical narrative analysis that accompanies statistical testing is to guard against erroneous conclusions based on structural breaks that are unrelated to historically documented interventions in e or in f.

We study central and local government in the United Kingdom between 1820 and 1913 for two particular reasons. First, in the United Kingdom, the franchise governing elections to the House of Commons was sequentially extended through three important and distinct reforms in 1832, 1867 and 1884, with the third Reform being the most

substantial increasing the fraction of adult males with the right to vote by 78 percent.[4] For local government - the Municipal Boroughs - the franchise was also extended in a sequence of reforms in 1869, 1878, 1882, 1888 and 1894. Unlike many other countries in Europe where democratic rights were extended to (almost) all males in one go, these sequential reforms makes the UK an ideal testing ground for the Redistribution Hypothesis. Second, the Hoover approach requires long and reliable time series on both the franchise and on fiscal variables related to redistribution. For the central government in the United Kingdom such data are available from secondary sources for the entire period 1820 to 1913. For local government, we have collected, from primary sources in the British Parliamentary Papers, a new data set that tracks the fiscal system of the Municipal Boroughs and the number of voters with the right to vote in local elections for the period 1867 to 1912. These data enable us to test the Redistribution Hypothesis for both central and local levels of government.

For the nexus between the fiscal system of the central government and the suffrage for the House of Commons, we find one instance in which the historical narrative of interventions to the franchise - specifically, the 3rd Reform Act 1884 - and the pattern of structural breaks are consistent with the Redistribution Hypothesis. However, this evidence concerns the effect of franchise reform on the expenditure size of government only. It is not found in the tax size and, moreover, it is not robust to variation in the factors used to pre-filter the f and e series to control for structural breaks that are unrelated to the Redistribution Hypothesis. We also find instances in which interventions to the fiscal system - in particular the fiscal rules laid down by William Gladstone in the 1850s and dismantled in the 1890s - and the pattern of structural breaks are consistent with the Redistribution Hypothesis, in the sense that these structural shifts in fiscal policy did not affect the franchise process, which was evidently independent of them. The result related to Gladstone's fiscal rules is robust to the choice of variables used to pre-filter the series. For local government, we find evidence that the Municipal Franchise Act of 1869 may have had a causal impact on the tax structure - specifically, the share of property tax revenues - of the Municipal Boroughs. This causal order cannot, however, be confirmed by evidence from historically plausible interventions in the local fiscal system. Overall, then, we do not find compelling results, which must include evidence that franchise extension causes fiscal innovation, from either level of government in support of the Redistribution Hypothesis. Indeed, our most robust finding is that innovations in the fiscal policy process at the central government level (i.e., the beginning and end of Gladstone's fiscal rules) were independent of the suffrage process.

[4]We do not include the 4th reform act in 1918 because it coincided with the end of World War I which makes it impossible to disentangle the effect of war on the fiscal system on the one hand and the extension of suffrage on the other.

The rest of the paper is organized as follows. In Section 2, we review the empirical literature on the Redistribution Hypothesis. In Section 3, we explain in more detail how the Hoover approach works. In Section 4, we present the historical narrative analysis. This plays an important role as it forces discipline on the econometric analysis and combines historical with statistical information. In Section 5, we introduce the data, specifying exactly how the franchise and elements of fiscal structure are measured. In Section 6, we discuss the pre-filtering of time series to remove breaks unrelated to the Redistribution Hypothesis and how we conduct the tests for structural breaks. In Section 7, we present our results. Finally, in Section 8, we offer some concluding remarks concerning the difficulties of testing the Redistribution Hypothesis.

2 Literature review

The Redistribution Hypothesis claims that franchise extensions that enable poorer citizens to vote cause an increase in state-sponsored redistribution through the fiscal system. Demands for redistribution can manifest themselves either through additional spending on goods and services that benefit the newly enfranchised voters and are financed through general taxation or through a more progressive tax system, or both.

Many cross national studies have tested the hypothesis on samples of countries that today are established democracies, often using data from the 19th and early 20th century when these countries relaxed restrictions on the right to vote. On the expenditure side, Lindert (1994) and Kim (2007) find a strong correlation between central government social spending and the fraction of the population that can vote in a sample of mainly European countries prior to World War II.[5] Aidt et al. (2006) study the composition of central government spending in a similar panel and find evidence that franchise extension was associated with a shift towards more spending on publicly provided private goods. This is broadly consistent with the Redistribution Hypothesis.

The historical evidence related to the reaction of the revenue side to democratization is more mixed. In Western Europe prior to World War II, the relationship between franchise extension and central government taxation is complicated by a number of mitigating factors. For example, the share of direct taxes (including the personal income tax) was positively affected by the franchise extension, but only if tax collection costs were below a given threshold (Aidt and Jensen 2009a) and franchise extension delayed the adoption of the income tax possibly because the adoption required support from landowners (see

[5]Lindert (2004a,b) is the authoritative study of the historical evolution of the fiscal state. Dincecco (2011) studies links between political transformations more broadly and the central government's public finances in Europe between 1650 and 1913.

Aidt and Jensen 2009b; Mares and Queralt 2015). Outside Europe, the limited historical evidence is more encouraging for the Redistribution Hypothesis. Aidt and Eterovic (2011) study Latin America during the 20th century and find that higher levels of political participation was associated with an increase in taxation (relative to GDP) in general and with an increase in income taxation in particular.

Another branch the literature on the Redistribution Hypothesis studies the period after World War II. This has the advantage that better data are available for more countries. Yet, the literature struggles to separate the effect of franchise extension from other aspects of democracy because democratic reforms during this period almost always entail a package (including voting rights to all adult citizens, the secret ballot, etc.). The evidence from this literature is mixed. Kenny and Winer (2006) study a large sample of 100 democratic and non-democratic countries at differing stages of development, in which mature democracies in fact rely substantially more on income taxation, and find that this remains so after allowing for the roles of the level of development and of economic structure in shaping tax systems. Their results indicate that reliance on income taxation is related positively to the degree of democracy, and possibly to the voluntary compliance, forthcoming in democracies, that is required for a modern income tax. Mueller and Stratmann (2003) find a positive relationship between election turnout and the size of government. Husted and Kenny (1997) also find evidence that is consistent with the Meltzer-Richard hypothesis by exploring the effects on social spending of the abolition of literacy and poll tax restrictions on the right to vote in the 1960s in a panel of US states. In contrast, Profeta et al. (2013) find, for a large sample of developing countries between 1990 and 2005, that neither total tax revenue nor the composition of taxes are, in general, significantly correlated with the strength of democratic institutions or with the protection of civil liberties. Others have argued that the effect of democratization is conditional. Boix (2003) shows, for example, that the positive (Meltzer-Richard) effect is present only above a certain GDP per capita threshold. Plumper and Martin (2003) find a non-linear (U-shaped) effect of democracy on the overall size of government spending, suggesting that democratization is initially associated with retrenchment rather than fiscal expansion and redistribution.

Aidt and Jensen (2013) is one of the few studies that makes an explicit attempt to isolate plausibly exogenous variation in the allocation of voting rights. They build on the theory of Acemoglu and Robinson (2000) and extended and tested in Aidt and Jensen (2014), and develop an instrumental variables strategy where suffrage reform is instrumented by a measure of the threat of revolution. They find some evidence that suffrage reform caused fiscal expansion, but the effect, economically speaking, is small and in some cases, suffrage reform appears to be associated with retrenchment.

A promising approach to causal inference concerning the Redistribution Hypothesis is to study the relationship between local government public finances and the extension of franchise in local elections. This franchise governing local elections are often decided by the parliament or legislature of the country. This makes it possible to exploit natural experiments induced by nation-wide suffrage reforms which are plausibly exogenous from the perspective of the local government units that they affect. Aidt et al. (2010) and Chapman (2018), for example, investigate the extension of the franchise in local elections in England and Wales in the second half of the 19th century and find, contrary to the Redistribution Hypothesis, that franchise extension can be associated with a reduction in local government spending on sanitation and other local amenities. The logic behind this retrenchment effect is that, unlike in Meltzer and Richard (1981), the tax price is higher than the benefit of more public services for the new voters and they, therefore, demand less spending and lower taxes (see also Husted 1989).[6]

To summarize, the literature contains a wide variety of results, and there is no consensus as to whether the Redistribution Hypothesis is correct, wrong or perhaps just incomplete. In the next section, we discuss an approach to uncovering causality between the franchise and the fiscal structure that has not previously been applied to investigate the Redistribution Hypothesis.

3 The Hoover approach

The Hoover approach to causality consists of two steps (Hoover 2001) and can be adopted to study the causal relationship between the extension of the franchise and aspects of fiscal structure. The first step is to construct a narrative of possible interventions in the suffrage and in the fiscal structure from the historical record. This is done prior to any statistical testing. The second step is to use statistical tests to find patterns of structural breaks which are then compared to the historical narrative to draw causal inferences. In this section, we present the basic logic behind the second step of the Hoover approach. We return to the historical narrative in Section 4.

To begin, suppose we have two time series that measure the extension of the franchise e and some aspect of the the fiscal system f. With regard to the causal relationship between the two processes underlying the observed time series, there are four possibilities

$$e \leftarrow f;\ e \rightarrow f;\ f \leftrightarrow e;\ f \perp e,$$

[6]Retrenchment may not be strictly at odds with the Redistribution Hypothesis, in the sense that it stems from suffrage extension, in a manner that is consistent with the interests of new voters (who are better off with lower taxes). It does suggest, however, that the standard formulation of the hypothesis is, at least, incomplete.

where the right and left arrows indicate direction of causality, $\leftrightarrow$ indicates mutual dependency, and $\perp$ indicates independence. Now, consider the marginal and conditional distributions that make up the joint distribution $D(e, f)$ of the two series. This joint distribution can be written as the product of a conditional and a marginal distribution in two equivalent ways:

$$D(f, e) = D(e)D(f \mid e) = D(f)D(e \mid f). \tag{1}$$

While the two factorizations of the joint distribution are observationally equivalent, Hoover (2001) shows that breaks in the process governing the mean of f will lead to one pattern of breaks in (or shifts in the mean of) these conditional and marginal distributions if the fiscal structure causes the franchise, and to a different pattern of breaks or shifts if the redistribution hypothesis is true and a break in the process for e causes a break in the process governing *f*. By identifying structural breaks in the time series for e and *f*, and then observing the pattern of breaks in the associated marginal and conditional distributions, Hoover shows that it is possible to determine the direction of causality between them.

An example, following Hoover (2001, 192-201), illustrates how causality may be determined. Since this argument may be unfamiliar to many readers, we summarize the example at length using the notation and context of our paper. The premise of the example is that e causes f in the sense that the value of f depends on the value of e, but not vice-versa (Hoover 2001, 59). For simplicity, we assume that the time series for e and f are normally, independently distributed processes.[7] Combining these two assumptions, we can write the data generating processes as

$$f = \alpha \cdot e + \epsilon; \text{ with } \epsilon \sim N(0, \sigma_\epsilon^2) \tag{2}$$

$$e = \beta + \eta; \text{ with } \eta \sim N(0, \sigma_\eta^2), \tag{3}$$

where $cov(\epsilon, \eta) = 0$, $E(\epsilon_t, \epsilon_s) = 0$, $E(\eta_t, \eta_s) = 0$, and α and β are parameters. The reduced forms of equations (2) and (3) which, along with the distributions of ϵ and η, describe the joint probability distribution *D(e,f)* are:

$$f = \alpha \cdot \beta + \alpha \cdot \eta + \epsilon \tag{4}$$

$$e = \beta + \eta. \tag{5}$$

[7] We omit time subscripts on the variables for simplicity, but it is understood that the data consist of long time series.

The joint distribution *D(e,f)* is a multivariate normal with well-known form (see, e.g., Mood and Graybill 1963, Chapter 9). Accordingly, we can use our knowledge of the functional form of the multivariate normal along with the decomposition in equation (1) to specify the following conditional and marginal distributions:

$$D(f|e) = N(\alpha \cdot e, \sigma_\epsilon^2) \tag{6}$$

$$D(e) = N(\beta,\ \sigma_\eta^2) \tag{7}$$

$$D(e|f) = N(\frac{\alpha\sigma_\eta^2 f + \beta\sigma_\epsilon^2}{\alpha^2\sigma_\eta^2 + \sigma_\epsilon^2}, \frac{\sigma_\eta^2\sigma_\epsilon^2}{\alpha^2\sigma_\eta^2 + \sigma_\epsilon^2}) \tag{8}$$

$$D(f) = N(\alpha\beta, \alpha^2\sigma_\eta^2 + \sigma_\epsilon^2). \tag{9}$$

Now, recall that the premise of the example is that e causes *f*. If there is a break in the mean (β) of the suffrage series in a given year, e.g., because a significant suffrage reform took place in that year, then the equations above indicate that all of *D(e)*, $D(e|f)$ and *D(f)* break. But conditional on the franchise, the distribution $D(f|e)$ remains stable. That is, the distribution that corresponds to the true causal structure remains stable in the year where all the other distributions exhibits a structural break. Intuitively, since e causes f by assumption, controlling for changes in the suffrage process leaves the fiscal process unaffected. Analogously, and still assuming that e causes *f*, the equations above indicate that a break in the mean (α) of the process governing f - as when, for example, there is a tax collection innovation that causes the size of the public sector to increase for reasons unrelated to e - leads to breaks in all of the distributions including $D(f|e)$, except that for the marginal distribution *D(e)*. The latter distribution remains stable in the year in which the other distributions break because, by assumption about the direction of causality, the marginal distribution of e does not depend on *f*. Again, we see that it is the distribution that reflects the true underlying causal structure that remains stable.

If in the presence of breaks in e, we observe that $D(f|e)$ remains stable, *and* if in the presence of breaks in *f*, we observe that *D(e)* remains stable, we may conclude that there is evidence that e causes *f*. Hence, the first decomposition of the joint distribution in equation (1), $D(f|e) \cdot D(e)$ is different from the second one, in that at least one of its two elements remain stable if e causes *f*. But neither of the elements in the second decomposition remain stable in the face of these shocks. This is summarized in Table 1. We observe that we can learn whether the voting franchise causes redistribution from interventions in either the suffrage series or in the fiscal services. However, convincing evidence in favour of the Redistribution Hypothesis requires both, that is, that we can find years where the historical narrative indicates an intervention in the suffrage series

(break in β) and the statistical tests indicate the pattern in column 1 *and* that we can find (other) years where the historical narrative indicates an intervention in the fiscal series (break in α) and the statistical tests indicate the pattern in column 2. The case

Table 1: Example: Using structural breaks to determine the causal order.

Distribution	break in β	break in α
$D(f\|e) = N(\alpha e, \sigma_\epsilon^2)$	stable	break
$D(e) = N(\beta, \sigma_\eta^2)$	break	stable
$D(e\|f) = N(\frac{\alpha\sigma_\eta^2 f + \beta\sigma_\epsilon^2}{\alpha^2\sigma_\eta^2+\sigma_\epsilon^2}, \cdot)$	break	break
$D(f) = N(\alpha\beta, \cdot)$	break	break

Notes: The Table is based on the assumption that e causes f and shows how an intervention is e (column 1) or f (column 2) is associated with a unique configuration of breaks.

in which f causes e can be assessed in analogous fashion, with the opposite conclusion that one element in the second decomposition in equation (1) will remain "stable". If the processes are simultaneous, in the face of the two shocks considered all of the elements in the decompositions in equation (1) will, under one or the other shock, be unstable.

Adding this up, we can apply the following algorithm to locate years where the constellation of breaks in the marginal and conditional distributions of e and f enables us to draw inferences about the causal relationship between them:

Test A: (i) check that $D(e|f)$ and $D(e)$ break in year t; (ii) use Table 1A to check if in year t, $D(f)$ and $D(f|e)$ are stable or not and draw the inference indicated.

Test B: (i) check that $D(f|e)$ and $D(f)$ break in year t; (ii) use Table 1B to check if in year t, $D(e)$ and $D(e|f)$ are stable or not and draw the inference indicated.

An integral feature of the Hoover approach is that the statistical analysis should be *preceded* by a detailed analysis of the historical narrative. This provides non-statistical information about possible structural breaks which is used to assess, confirm or reject the outcome of the formal statistical analysis. A prerequisite for drawing causal inferences from test A and test B is that the historical narrative clearly indicates that there was an intervention in the suffrage (test A) or in the fiscal structure (test B) in year t that is consistent with the historical facts. If not, no causal inference should be based on it.

Table 2: How to determine the causal order from patterns of structural breaks

Table A

	$D(f)$		
		Stable	Unstable
$D(f\|e)$	Stable	$f \perp e$	$e \rightarrow f$
	Unstable	$e \leftarrow f$	$f \leftrightarrow e$

Table B

	$D(e)$		
		Stable	Unstable
$D(e\|f)$	Stable	$f \perp e$	$e \leftarrow f$
	Unstable	$e \rightarrow f$	$f \leftrightarrow e$

Notes: Panel A or B are used to draw conclusions about the causal order conditional on test A(i) or test B(i) having been passed. The arrows indicate the direction of causality and $\perp$ means that the two series are independent.

4 The historical narrative

In this section, we review the historical record to identify plausible candidates for interventions (structural breaks) in the franchise rules and in the fiscal system.

4.1 Historical narrative: the central government, 1820-1913

Table 3 provides an overview of the major events that we view as plausible candidates for structural breaks in the rules governing election to the British House of Commons and in the fiscal system.

4.1.1 The suffrage rules for the House of Commons

Prior to 1832 the rules that governed elections to the House of Common had not been changed for almost two centuries. A small fraction of the population could vote and elections were often not contested. The first of three major suffrage reforms between 1820 and 1913 happened in 1832. The Great Reform Act of 1832 introduced a limited property-based suffrage, with different property value thresholds in the borough and county constituencies, and redistributed seats from small "rotten" boroughs to the industrial cities in the Midlands and the North of England (Brock 1973). The consequence was to increase the male electorate from 2-3 percent to 4-6 percent of the total English population of 13 million (Cannon 1973).

The next franchise extension happened in 1867 (Smith 1966). The Second Reform Act granted the vote to all householders in the borough constituencies as well as lodgers who paid rent of £10 a year or more, reduced the property threshold in the county constituencies and gave the vote to agricultural landowners and tenants with very small amounts of land. Prior to the reform less than one million of the seven million adult

Table 3: Central government: Historical narrative of "interventions" in the suffrage and in the fiscal system.

Year	**Fiscal system**	**Suffrage**	**Narrative intervals**
1802	War-time income tax introduced		1800-04
1816	War-time income tax repealed		1814-18
1832 (1832)		The Great Reform Act	1830-34
1842	Income tax and reduction in tariffs on industrial products		1840-44
1846	Repeal of the Corn Laws (tariffs on cereals)		1844-48
1853-57	Gladstone's Fiscal Constitution (GFC) established		1853-57
1867 (1868)		The Second Reform Act	1866-70
1872 (1874)		The Secret Ballot Act	1872-76
1883 (1885)		Corrupt Practices Act	1883-87
1884 (1885)		Third Reform Act	1883-87
1885 (1885)		Redistribution of Seats Act	1883-87
1894-1900	Gladstone's Fiscal Constitution (GFC) breaks down, beginning with Harcourt's graduated death duties		1894-00
1906	Asquith's budget introduces differentiated income tax and old age pensions		1904-08
1909	Lloyd George's "People's budget" with progressive income tax (super-tax)		1907-11
1911		The Parliamentary Act reduces the powers of the House of Lords	1909-13

Sources: Seligman (1911), Sabine (1966), Cannon (1973), Peters (1991), Daunton (2001), Evans (2000), and Keir (1953).
Note: The years in parenthesis in column one indicate the year of the first election where the reformed suffrage rules applied. This is the year relevant for the intervention. Column four (with the heading "narrative intervals") records the intervals which we use to judge if a statistical structural break in the relevant series is consistent with the historical narrative. For unique events, we use two years before and two years after the year of the intervention (or in the case of interventions in the suffrage series from the year of the first election under the new rules).

males in England and Wales could vote; the Act immediately doubled that number and it gave the skilled working class the majority in many urban constituencies (Evans 2000). The new rules were first applied in the general election in 1868.

The third step in the franchise extension process was William Gladstone's Representation of the People Act 1884 (the Third Reform Act) and the Redistribution Act 1885 (Glen 1885). Taken together, the two acts extended the same voting qualifications as existed in the borough constituencies to the counties and distributed seats from the countryside to the urban areas in particular to London (Blewett 1965). In essence, they established the modern one member constituency. The new rules were first applied in the general election in 1886. The size of the electorate expanded from about 36 percent of the adult male population to 64 percent (Flora et al. 1983).

Voting was open till the Ballot Act introduced secret ballot in 1872 (Fitzgerald 1876). The act was aimed at reducing vote buying (Aidt and Jensen 2017). The Corrupt Practices Act of 1883 was a continuation of this and criminalized attempts to bribe voters and standardized the cap on election expenses introduced in 1872. These two reforms did not have a direct effect on who could vote. However, they might have affected, by creating a more independent electorate, the identity of the *de facto* median voter and in that way have caused demands for redistribution. Likewise, the Parliamentary Act of 1911, which removed the right of the House of Lords to veto bills related to taxation and public spending (so-called money bills) completely, might have given more effective power to the median voter among the electors for the House of Commons (Keir 1953).

We stop the analysis in 1913, i.e., before the Representation of the People Act 1918. The rationale for this is that this act was introduced at the end of the war. The war and its aftermath is likely, as argued by Scheve and Stasavage (2016), to have affected both the fiscal system (need for war time spending) and the franchise. This also excludes the 1928 reform that granted women the right to vote.[8]

4.1.2 The fiscal system for the central government

Our sample period starts in 1820 after the Napoleonic Wars. The revenue demands of the Wars were met partly by debt and partly by the first British income tax. After the war in 1816, the wartime income tax was repealed As a consequence of the repeal, the British tax base in 1820 was very narrow and relied heavily on indirect taxes (customs and excise duties, in particular) with direct taxes, such as land and assessed taxes, contributing less than 15 percent of the total.

[8]Women's suffrage is likely to have a different logic unrelated to the redistribution hypothesis (Aidt and Dallal 2008; Bertocchi 2011; Hicks 2013).

Between Victory at Waterloo and World War I, the British fiscal system was subject to seven major fiscal interventions which could have induced structural breaks in the level and structure of central government spending and taxation (see Table 3). The first major fiscal intervention was Sir Robert Peel's re-introduction of the income tax in 1842, as a temporary measure to close a budget deficit and to fund the removal of tariffs on industrial goods (Seligman 1911, Chapter XX). The blueprint was the 1806 wartime income tax with a non-differentiated flat-rate tax of 2.9 percent (seven pence in the pound) with the exception threshold being about £150 per year. Babbage (1852) estimates that at most 150,000 of the electors under the 1832 franchise had an income above this threshold. It was renewed, first, in 1845 and then again in 1848 and 1853, and for the next 50 years repeated attempts were made to repeal it but without success. Peel's broader motivation for the new tax was that he wanted to create a tax system that was neutral between different social classes and to remove political tension from tax policy (Daunton 2001, p. 80).

The second major intervention was Sir Robert Peel's repeal of the Corn Laws in 1846 (Aydelotte 1967). The Corn Laws had been introduced in 1815 to protect the agricultural sector though tariffs on imported food and grain ("corn"). Its repeal, which happened despite significant opposition from within Peel's own party (the Tory party), was a victory for supporters of free trade (Schonhardt-Bailey 2006) and helped to entrench the income tax.

The third major intervention was William Gladstone's first budget as Chancellor of the Exchequer in 1853. It laid the foundation for the "Gladstonian Fiscal Constitution" which was firmly established by 1856-57. The constitution or social contract aimed at achieving fiscal retrenchment and at securing broad agreement on the principles underpinning taxation.[9] This innovation will prove to be of importance in the statistical analysis that follows below.

Daunton (2001, pp 66-68) describes the three pillars of the constitution:

1. **Budget unity:** All revenue should be unified and treated as a single pool of money which was separate from the purposes for which it was to be spent. This meant a stop to the practice of hypothecation of tax revenues and required new accounting practices.[10]

2. **No *virement* of funds and annual votes:** All expenditure items were minutely subdivided subject to annual votes in the House of Commons. Surplus funds from

[9]Baysinger and Tollison (1980) interpret Gladstonian finance in the light of constitutional tax rules (Brennan and Buchanan 1980), but the spending rules were at least as important (Leathers 1986).

[10]The system of double-entry book-keeping was extended to all government departments in 1857.

one budgetary head could not be moved to another and were allocated to a sinking fund to repay the (war) debt.

3. **The Goulburn principle:** The tax system as a whole should exhibit a balance between different types of taxes (direct and indirect) so that the various social classes all contributed a fair share to the total. This should be achieved by compensatory taxes rather than graduation of any particular tax.[11]

The Gladstonian Fiscal Constitution became fiscal orthodoxy for over four decades. Its limits, however, started to show in the mid 1890s and by 1900 the consensus behind the Constitution had all by broken down. We have identified four potential interventions in the fiscal system between 1894 and 1909 that contributed to the demise of Gladstonian public finance and could have cause a structural break in the fiscal system. The first major challenge came in 1894, when the Chancellor of the Exchequer, William Harcourt, achieved parliamentary approval for graduated taxation of estates at death. This prepared the way for the later reforms of the income tax by the Liberal politicians Herbert Henry Asquith in 1906 and Lloyd George in 1909. The next intervention happened as a consequence of the Boer war in 1899-1902. To fund the war, the income tax had to be raised to an unprecedented 6 percent and direct taxes for the first time raised more revenue than indirect taxes (Sabine 1966, Chapter 6). The third intervention came in 1906 when a Select Committee in the House of Commons concluded that graduation and differentiation of the income tax were both feasible. The report was quickly turned into legislation by the Liberal Chancellor Herbert Henry Asquith. He introduced in his 1906 budget differentiated taxes on earned and unearned income for the first time and also introduced the first old-age pension – the start of the welfare state (Macnicol 1998). Lloyd George's People's Budget from 1909 introduced the graduated super-tax into the income tax schedule, and this dealt the final blow to the Gladstonian Fiscal Constitution.

4.2 Historical narrative: Local government in England and Wales 1867 to 1912

We focus on the Municipal Boroughs outside London, both as Corporations and as acting Urban Sanitary Authorities.[12] Pubic finance data are available from 1867 to 1912, and we focus the historical narrative on this period.

[11]The principle was articulated by Henry Goulburn, who was the Chancellor in 1848. He stressed that the way to "correct the inequality of one tax" was through "countervailing inequalities of another, so that the balance of the whole system of taxation should press equally and justly upon all classes" (Hansard 1848, col. 192).

[12]Local government in England and Wales consisted of a complex variety of local authorities with different powers to tax and spend, with different governance rules and covering different geographical units (see, e.g., Smellie 1946). The Municipal Boroughs and the Urban Sanitary Authorities were the

Table 4: Local government: Historical narrative of "interventions" in the municipal franchise and in the Municipal Boroughs' fiscal system, 1867-1912.

Year	**Fiscal system**	**Municipal suffrage**	**Narrative intervals**
1869		Municipal Franchise Act	1868-72
1875	The Public Health Act		1873-77
1878		The Parliamentary and Municipal Registration Act	1877-81
1882		Municipal Corporation Act	1881-85
1888		County electors act	1887-91
1888 (1889)	The Local Government Act		1887-91
1894		Local Government Act	1893-97
1902	The Education (Balfour) Act		1900-04

Sources: Vine (1878), Vine (1882), Keith-Lucas (1952), Keith-Lucas (1977), Daunton (2001), Davis (2000), Millward (2000), Doyle (2000), Smellie (1946), Redlich and Hirst (1958) and Eaglesham (1967). Note: Column four (with the heading "narrative intervals") records the intervals which we use to judge if a statistical structural break in the relevant series is consistent with the historical narrative. We use two year before and two years after the intervention to define these intervals. For interventions in the suffrage, we take into account that one one-third of the local government councils were elected each year and center the interval on the second election under the new rules (in the year after the adoption of the reform). For the interventions in the fiscal series, we center the interval on the fiscal year of the intervention. The years in parenthesis in column one indicate the year in which the legislation took effect if different from the year of adoption.

4.2.1 The municipal franchise

The Municipal Corporation Act 1835 established the legal framework for the yearly elections to the councils of the Municipal Boroughs. The municipal franchise was granted to male taxpayers who had paid the local property tax (called the rate), who had lived in the borough for two-and-a-half years, and who had registered to vote. In contrast to the 1832 Parliamentary franchise discussed above, which imposed a property value threshold on the right to vote, the franchise for the local elections applied no such threshold and upheld the principle of "one taxpayer one vote". However, as discussed by Keith-Lucas (1952, Chapter 3) this did not, in practice, give voting rights to "poor" householders. Firstly, it was common that properties of low value were exempt from paying the rates, which then automatically disenfranchised the occupiers. Secondly, the long residency requirement excluded many workers in the expanding industrial cities because they moved around a lot. Thirdly, many tenants were indirect ratepayers in the sense that their

most important of these local government authorities in provincial towns and the only authorities for which systematic quantitative data on the franchise can be found.

landlord paid the rate on the property they occupied and it was initially unclear if they qualified for the vote (Daunton 2001, Chapter 9).

The suffrage was gradually extended over the next 70 years through national legislation that reduced the residency requirement and clarified the rights of indirect ratepayers and tenants living in shared accommodation (lodgers). Table 4 lists five key interventions in the municipal franchise in the years after 1867. The Municipal Franchise Act of 1869 reduced the residency requirement from two-and-a-half years to one year, added "other buildings" to the list of qualifying properties and gave unmarried women who satisfied the other franchise requirements the right to vote (Vine 1878, p. 45-46). The Parliamentary and Municipal Registration Act of 1878 and the Municipal Corporation Act of 1882 settled the ambiguity related to indirect taxpayers and established that they did quality to the vote in local elections (Keith-Lucas 1952, Chapter 3). Two other bills, although not aimed specifically at the Municipal Boroughs had, nonetheless, implications for the municipal franchise. The County Electors Act of 1888, which introduced elected councils in the counties, gave those who had obtained the right to vote in parliamentary elections in 1885 the right to vote also in local elections and the Local Government Act of 1894 added lodgers to the electorate.

4.2.2 The local fiscal system

The fiscal responsibilities of the Municipal Boroughs were based on two fundamental principles (Waller 1983). Firstly, local expenditures were funded locally, either out of estate income or through the local property tax. Grant-in-aid from the central government were minimal and debt could only be issued for specific investment projects. Secondly, within the statutory duties laid down by law, the Municipal Boroughs had much flexibility to decide the scale of their activities and the quality and quantity of the services they wished to supply. The Municipal Boroughs were responsible for administration of justice, and the police and for public property. Importantly, where a borough acted as urban sanitary authority, it also was responsible for investments in clean water, sewage systems, garbage removal, maintenance and improvement of thoroughfares and paving of streets.[13]

Table 4 lists three potential interventions in the local fiscal system.[14] The first major intervention was the Public Health Act 1875.The act gave the boroughs that were or became urban sanitary authorities new fiscal powers, including the power to purchase,

[13]Vine (1878, p. 15) provides a complete list of the activities of the Municipal Boroughs that acted as urban sanitary authorities under the various Public Health Acts (from 1848, 1858 and 1875). Millward (2000) discuss their involved with the running of urban utilities (water, gas and electricity).

[14]The Public Health Act 1872 did not not have a direct effect on the public finances. Likewise, the Local Government Act 1894, which as noted in Section 4.2.1 affected the suffrage, did not directly change the fiscal powers and responsibilities of the Municipal Boroughs but renamed the urban sanitary districts as urban districts.

repair or create sewers, to control water supply, to regulate cellars and lodging-houses and to establish bye-laws for controlling new streets and buildings.

The second major intervention was the Local Government Act 1888, which came into force on April 1889. The act established county boroughs in 59 of the largest Municipal Boroughs (with a population of more than 50.000 in 1881). This gave these boroughs new tax and borrowing powers, new responsibilities for county assets, repair of county roads and bridges, and the responsibility for the establishment and maintenance of reformatory and industrial schools (Davis 2000, p. 269-71).

The third intervention was the Education (Balfour) Act 1902 (see Eaglesham 1967). Prior to the Balfour Act, primary education was outside the remit of the boroughs. The act gave the county boroughs significant new fiscal responsibilities and powers in relation to elementary and secondary schools in their jurisdiction.

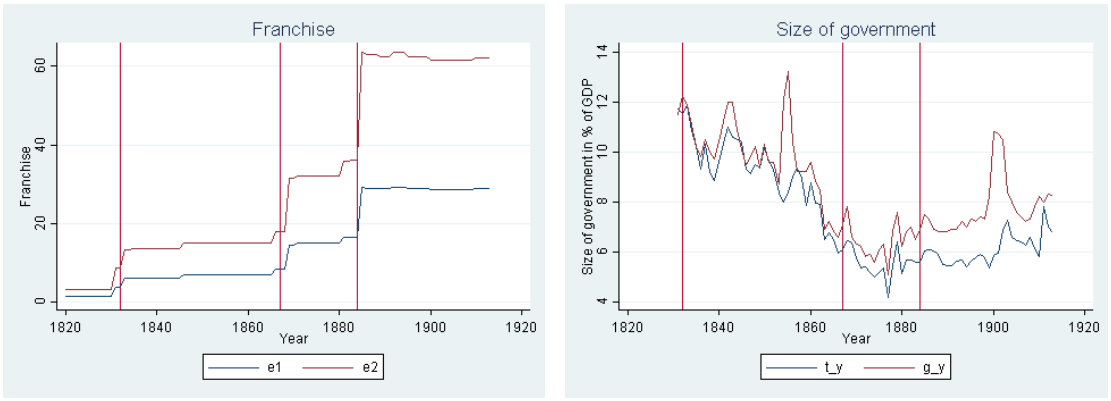
5 The data

We have collected historical data quantifying the gradual extension of the suffrage (e) and the evolution of the fiscal system (f). The central government data cover the long 19th century (1820 to 1913), while the local government data cover the period from 1867 to 1912.

We measure the extension of the franchise that governed elections to the House of Commons in two ways (Flora et al. 1983; Caramani 2000). The first measure is the proportion of the adult population with the right to vote in parliamentary elections (e_1); the second measure is the proportion of the adult male population with the right to vote in parliamentary elections (e_2) and thus excludes women from the denominator. Figure 1(a) plots the two suffrage series. The vertical red lines are the suffrage reforms of 1832, 1867 and 1884, respectively. We notice the step-pattern of these series which captures the process of including men with progressively lower income in the electorate and the the corresponding fall in the income of the median voter. We measure the extension of the franchise for elections to the Municipal Boroughs in England and Wales as the number of voters registered to vote for municipal elections relative to the size of the municipal population (e_1^L).[15] While elections took place yearly, only eight cross sections of suffrage data are preserved in the British Parliamentary Papers and Vine (1878). Based on these cross sections, we construct the time series shown in Figure 2(a).[16] The suffrage granted in 1835 gave the right to vote to about 6 percent of the population

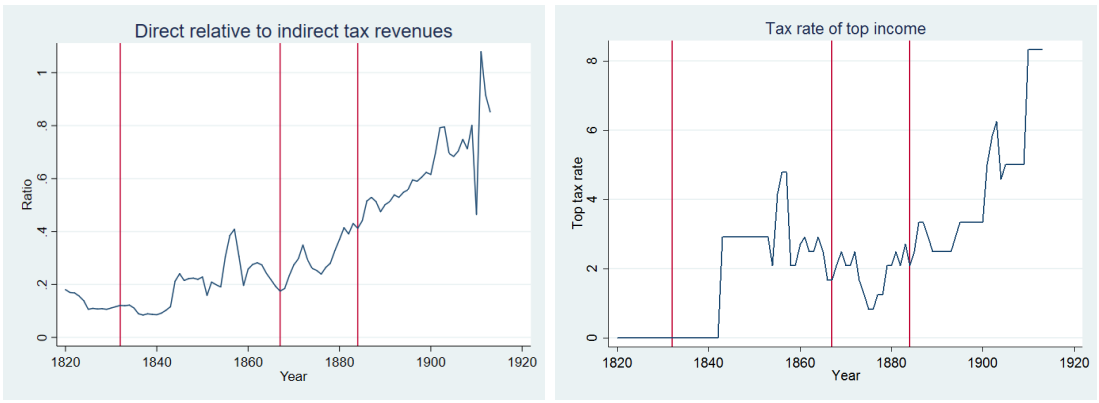
[15]It is not possible to obtain age- and gender-specific demographic data at the correct spatial units to calculate municipal voters per adult male and we also observe that a significant number of property-owning women had the right to vote in local elections.

[16]See Appendix A2 for details on data construction and primary sources.



(a) Franchise (e_1 and e_2)

(b) Government Size ($\frac{G}{Y}$ and $\frac{T}{Y}$)



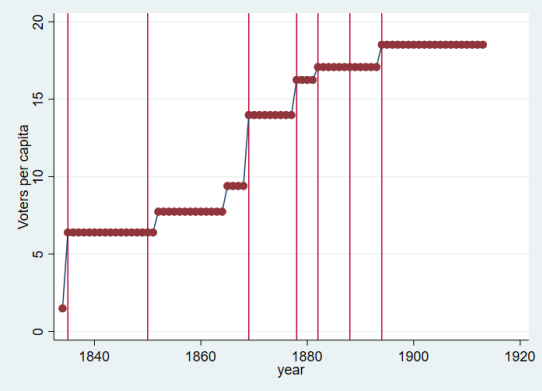
(c) Direct/indirect tax ($\frac{T^d}{T^{in}}$)

(d) Top Income Tax Rate (Top)

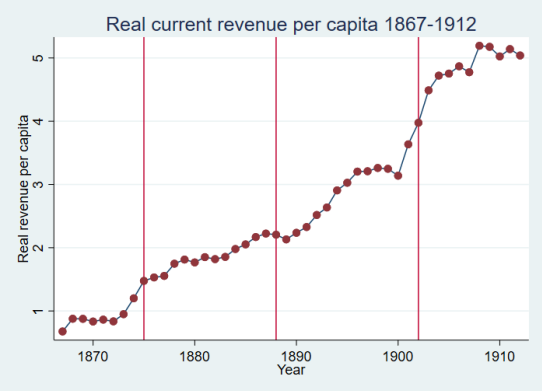
Notes: Panel (a) shows the two measures of the franchise for elections to the House of Commons. e_1 is the proportion of the adult population with the right to vote in parliamentary elections. e_2 is the proportion of the adult male population with the right to vote in parliamentary elections. Panel (b) shows the two measures of the size of government. $\frac{T}{Y}$ is total tax revenue out of GDP and $\frac{G}{Y}$ is total spending out of GDP. Panel (c) shows the ratio of direct to indirect tax revenue $\frac{T^d}{T^{in}}$. Panel (d) shows the top marginal income tax rate (Top). The red vertical reference lines indicate the years with suffrage reforms (1832, 1867 and 1884).

Figure 1: The evolution of the franchise and the fiscal variables for central government in the United Kingdom, 1820 to 1913

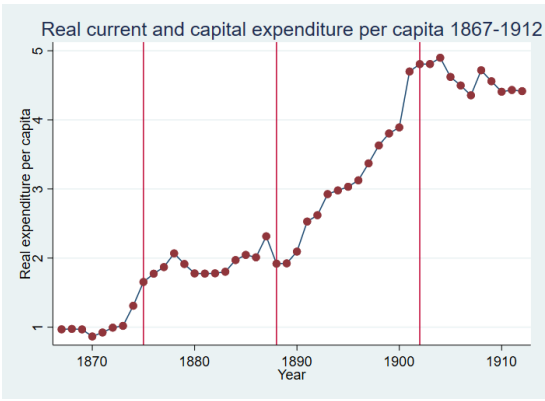
in the Municipal Boroughs. The Municipal Franchise Act of 1869, which lowered the residency requirement, had a big effect on the electorate: it increased from 9 to about 14 percent of the population. The Parliamentary and Municipal Registration Act of 1878 and the Municipal Corporation Act of 1882 increased the share from about 14 percent to 17 percent. The reforms in 1888 and 1894 had only a minor effect on the municipal electorate.



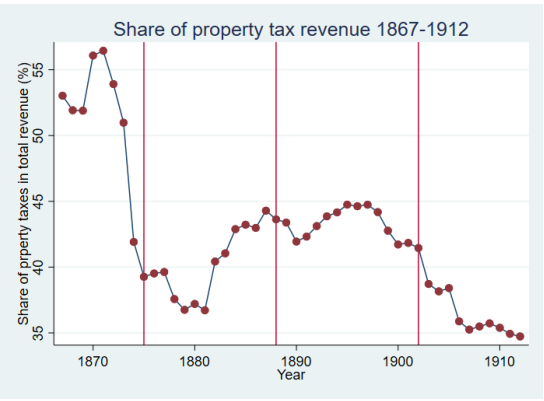
(a) Franchise e_1^L



(b) Local government size $\frac{T}{N}$



(c) Local government size $\frac{G}{N}$



(d) Tax composition $\frac{TAX}{T}$

Notes: Panel (a) shows the evolution of the number of voters per capita registered to vote in local elections to the Municipal Borough councils in England and Wales (e_1^L). Panel (b) and (c) show the evolution of total revenue including income from loans ($\frac{T}{N}$) and total expenditure including capital spending ($\frac{G}{N}$), respectively, expressed in real 1870 prices and per capita. Panel (d) shows the share of income from the property tax in percentage of total revenue ($\frac{\text{TAX}}{T}$). The data refer to the Municipal Boroughs in England and Wales which acted as Urban Sanitary Authorities. The vertical lines indicate years with a potential intervention in the suffrage or the fiscal system as identified in Table 4.2.1.

Figure 2: The evolution of the franchise and the fiscal variables in the Municipal Boroughs, 1867 to 1912

We measure three aspects of the fiscal system: scale, revenue composition, and tax rates. We quantify the scale of the central government's fiscal activities by total central government tax revenue relative to GDP ($\frac{T}{Y}$) and total central government spending relative to GDP ($\frac{G}{Y}$).[17] We measure the composition of tax revenues by the ratio of direct to indirect tax revenues ($\frac{T^d}{T^{in}}$). Direct taxes include property taxes and taxes on land and various assessed taxes, and from 1842 income taxes and later on corporation taxes. Indirect taxes include commodity taxes (excise duties) and taxes on international

[17]The series are constructed from data reported by Flora et al. (1983) and Maddison (2003).

trade. Direct taxes are, typically, better geared towards redistribution than indirect taxes and thus we associate an increase in $\frac{T^d}{T^{in}}$ with more redistribution.[18] We measure the rate structure by the top marginal income tax rates.[19] These capture the degree of progression in the income tax system.[20]

Figure 1(b) to (d) graph the central government's fiscal series. The vertical red lines indicate the years of the three suffrage reforms (1832, 1867 and 1884, respectively). Panel (b) shows the scale, measured either by total tax revenue out of GDP ($\frac{T}{Y}$) or total spending out of GDP ($\frac{G}{Y}$). We observe a long period of retrenchment between 1820 and 1875 where government spending and taxation fails to keep up with the expansion of the economy, and taxation and spending fall from about 10 percent of GDP to about 6 percent. The downwards trend breaks in the mid-1870s and by 1913, the expenditure and taxes stand at about 8 percent of GDP. The Crimean War (1853-56) and Boer War (1899-1902) are clearly visible spikes in spending (and in the income tax rate in Panel (d)). Panel (c) shows the evaluation of the composition of tax revenues $\frac{T^d}{T^{in}}$. At the beginning of the period, the indirect taxes contribute between 8 times as much as the direct taxes. The 1842 income tax reduces this to a factor of about 4 and the importance of indirect taxes continues to fall after that. By 1913, the ratio is close to one. Panel (d) shows the evolution of the top marginal income tax rate. The top income tax rate varies between a little under 1 percent to a maximum of 9 percent in 1909. We observe a gradual fall in the top rate between 1842 and 1879, after which it slowly increases by 8 percentage points. As discussed above, graduation in the income tax was not introduced till 1909, so for most of the period the marginal tax rate is the standard tax rate.

For local government, the Local Taxation Returns (contained in the British Parliamentary Papers) record for each Municipal Borough and Urban Sanitary Authority detailed expenditure and revenue accounts. We have aggregated these and constructed time series representing the local government fiscal system.[21] We measure the scale of fiscal activity by real revenue per capita ($\frac{T}{N}$) and real expenditure per capita ($\frac{G}{N}$). We normalize with the total municipal population to account for population growth and also for the fact the number of Municipal Boroughs expanded over time. We measure the revenue composition as property taxes as a share of total revenue ($\frac{\text{TAX}}{T}$). The two main alternative sources of revenues to property taxes were estate income and income from municipal operated

[18] Flora et al. (1983) report information on revenue sources which we supplement with information from Mitchell and Deane (1962, Chapter XIV).

[19] The source of these data is Scheve and Stasavage (2016).

[20] We have experimented with many other fiscal variables, including total real spending and taxation, a finer decomposition of taxes by source, public debt and bond rates, and the composition of spending. The three aspects we focus on in the reported analysis, we believe, are a good representation of the evolution of the fiscal system's redistributive characteristics between 1820 and 1913 at the level of central government.

[21] It is not a trivial matter to construct these series. Appendix A2 explains in detail now we dealt with a range of data constituency and aggregation issues.

public services (e.g., water or gas works). We interpret a higher share of property taxes as a more progressive local tax system. It is not possible to obtain consistent data on the property tax rates.

Figure 2(b) and (c) graph total real revenue and total expenditure per capita for the Municipal Boroughs between 1867 and 1912 while panel (d) graphs the share of property tax revenue relative to total revenue. The vertical red lines indicate potential intervention years (see Table 4). We observe a gradual increase in both revenue and expenditure over time, with an acceleration in the 1890s and a flatting out before World War I. The share of property tax revenues fell dramatically in the first half of the 1870s reflecting that many boroughs got involved in the running of local utilities (gas and water) which were major revenue generators (Millward 2000).

6 The statistical analysis

The statistical analysis proceeds in two steps: first, pre-filtering, and then, structural breaks analysis.

6.1 Filtering the data for confounders

A prerequisite for applying the Hoover approach is that reasons for breaks in the marginal and conditional distributions of e and f other than those directly related to interventions in the two processes themselves have been removed prior to testing for structural breaks. Failure to do so appropriately casts doubt on whether the statistical pattern of structural breaks are informative of the true casual relationship between e and f, as those structural breaks could be caused by a common confounding intervention.[22] The two prime candidates for confounding interventions that can induce structural breaks in both e and f are wars and shocks to the economic structure. For local government, wars are not likely to be a confounding factor, but shocks to the economy structure are.

First, war may affect both e and f for reasons unrelated to the Redistribution Hypothesis. Janowitz (1976), Ticchi and Vindigni (2008), Dincecco (2011) and others argue that wars are associated with suffrage reform. One reason, explored by Ticchi and Vindigni (2008), is that conscripted citizens are only willing to put effort into fighting wars if they are promised some amount of income redistribution in return. To make such promises credible, it may be necessary to relinquish political power to citizen-soldiers through an extension of franchise. At the same time, wars are obviously associated with fiscal (war-related) expansion but, on top of that, Scheve and Stasavage (2016) argue

[22] Appendix A1 provides a deeper explanation of why we need to pre-filter the data.

that progressive income taxation that falls on the rich compensates the poor for making an unequal sacrifice in mass warfare. We have collected information about United Kingdom's involvement in wars (including across the British Empire) from Britannica (1911, 2009) and Singer and Small (1994) and code the dummy variable *war* as one if the United Kindom was at war in a given year and zero otherwise.[23]

Second, interventions in the economic structure constitute a potential source of structural breaks in both e and f. The feedback from income to democratization is, in general, contested[24] but Brückner and Ciccone (2011) show that economic shocks are important for democratization and they open up possibilities for constitutional bargains (Congleton 2011, Chapter 12). Interventions in the economic structure have a direct effect on the tax structure (see, e.g., Musgrave 1969; Hettich and Winer 1999) because they affect tax collection costs (Aidt and Jensen 2009a), because the income responsiveness of various tax bases are different, and because government tax policy generally adapts to the evolution of the relative size of tax bases and to economic factors influencing their structure (Kenny and Winer 2006). In addition, economic growth and structural change may alter the franchise under existing rules by its effect on the wealth and income of citizens (Congleton 2004). We capture interventions in the economic structure with movements in real GDP per capita using data from Maddison (2003).

We use these data to filter e and f for the effect of war and for interventions in the economy structure by regressing the six series related to central government on the dummy variable *war* and on real GDP per capita. The residuals from these regressions isolate the evolution of the suffrage and the fiscal system net of interventions or shocks caused by war and structural change. Any structural breaks that are left are then not caused by these two potential confounders. For the four series for local government, we filter for the effect of structural change by regressing the series on real GDP per capita. We conduct the tests for structure breaks in the marginal and conditional distributions on these residuals, rather on the raw series. However, to avoid cumbersome notation and terminology, we refer henceforth to these filtered data series (the residuals) by the notation and name of the underlying series.

6.1.1 Structural breaks tests

We begin the structural breaks analysis by establishing the dynamic properties of the (filtered) data series and check that none of them are integrated of a higher order than

[23]We have coded the three largest wars during the period which were the Napoleonic Wars (till 1815, KIA: 26,489), Crimea War (1853-56, KIA: 22,000), and Boer War (1899-1902, KIA: 21,942), where KIA means killed in action, and two smaller wars: The Anglo-Afghan war in 1878 and Anglo-Egyptian war in 1882.

[24]See, e.g., Acemoglu et al. (2008) and Gundlach and Paldam (2009).

one. Table A1 in Appendix A3 reports the details. We find that the series are either I(0) or I(1). With a mixture of stationary I(0) and non-stationary I(1) variables, the appropriate models for the marginal and conditional distributions are Autoregressive-Distributed Lag (ARDL) models with an error-correction mechanism.[25] The marginal distributions for the series are AR(p) models where each variable is regressed on a constant and on a number of own lags (p). The conditional distributions are specified as ARDL(p,q) models in which the dependent variable is explained by lagged values of itself and by the current value and successive lags of the conditioning variable. For example, the ARDL(p,q) model for e_1 conditional on the $\frac{T}{Y}$ incorporates p lags of e_1 and the contemporaneous value and q lags of $\frac{T}{Y}$.

The optimal lag orders in the AR(p) and ARDL (p,q) models are determined by minimising Akaike's information criterion (AIC). The models are estimated by Ordinary Least Squares (OLS). Table A2 and A3 in Appendix A3 show the detailed specifications of these models. To increase the power of the structural break tests, we select the most parsimonious ARDL (p,q) model for each series and only keep the terms (both lagged terms of the dependent variable and the contemporaneous and lagged terms of the independent variable) which are significant in the full ARDL(p,q) model. Those parsimonious representations of the marginal and conditional distributions are the starting point for the structural breaks tests.

We use the structural breaks tests developed by Bai and Perron (2003) rather than the Chow test used by Hoover and Sheffrin (1992) and Hoover and Siegler (2000).[26] There are five varieties of the Bai-Perron test and we report results for all of them.[27] Two of these uses local search algorithms. The "Sequential test, all subsets" method splits the data into subsets and tests for an additional break point in each of these subsets, for a given pre-specified number of break points. The "Sequential L+1 breaks vs. L" method chooses the single added break point that reduces the sum of squares of the test

[25]The error-correction model is suitable for a mixture of I(0) and I(1) data, i.e., when some series are I(1) while others are stationary (there is also the possibility of co-integration among some of the I(1) variables). As most of the series that we consider are I(1), the model should be written in first-differences of the dependent variable and explanatory variables. This is, however, a simple re-parametrization of the level equation once the lagged dependent variable and the contemporaneous value of the independent variable are included. As both yield the error-correction form and it is easier to implement the structural break tests in levels, we use the ARDL(p,q) models in levels to conduct the structural breaks test.

[26]The Bai-Perron tests (Bai and Perron 2003) are superior to sequential Chow tests for a number of reasons. First, the Bai-Perron tests allow for multiple unknown structural breaks. Second, the Chow test requires us to pre-specify a "tranquil period", i.e., a period where we are sure that no structural breaks took place. This requires a judgement and the test results are sensitive to this, and it is not obvious which periods we should treat as "tranquil" in our application. Third, the Chow test can only identify one break point at the time in a given (known) year, and estimating break points outside "tranquil periods" one by one without a global maximisation procedure that takes the joint pattern of multiple break points into account can bias the results.

[27]We carry out the structural breaks tests in EViews 9.

statistics the most. The remaining three tests ("Global vs. none", "Global L+1 vs. L" and "Global information criteria" methods) employ global optimization procedures. The "Global information criteria" method uses the Schwarz criterion or the LWZ criterion (which is a modification of the Schwarz criterion) to approximately determine structural breaks. There are four sub-methods associated with the "Global L breaks vs. none" method. We use the method "Selecting Highest Significant". It chooses the largest number of breaks from among the significant tests. We make the same choice for the "Global L+1 vs. L" method.[28]

7 The results

To implement test A and B (Table 2) and establish the causal relationship between e and f from the structural breaks tests, we follow two rules. Firstly, we allow for a window of two years around the intervention years identified by the historical narrative analysis (as indicated by the narrative intervals in column (4) of Table 3 and Table 4) and judge a statistical break to be consistent with the narrative if it falls within one of these intervals. The rationale is that our statistical specifications, as detailed above, include lags of one to four years. Secondly, we use five variants of the Bai-Perron test. We apply two alternative criteria to accept a statistical break in a given interval: the strict criterion requires that two of the tests find a break; the weak test requires that one does. We report results with both criteria below and indicate with brackets which are based on the weak criterion.

7.1 Central government

Table 5 records (at the top) the years in which the Bai-Perron tests find break points in the marginal distribution of the series representing the franchise for the House of Commons and the fiscal system of central government. Two tests find a break in the franchise series around the 3nd reform act (1885) and one finds a break around the 2nd reform act (1866) but only in the series e_1 that normalizes the number of voters with the adult population. The reason the Bai-Perron tests do not consistently find break points around the three reform bills is that two franchise series after filtering for the effect of war and GDP per capita do not retain the ladder structure from Figure 1. The two variables clearly absorb some of the structural breaks in the franchise series.

The structural breaks in the marginal distributions of the fiscal series coincide in many cases with the interventions highlighted by the narrative analysis of the fiscal system. The beginning (1853-57) and, in particular, the end (1894-1900) of the Gladstonian Fiscal

[28] Appendix A4 contains more information about how we implement these tests.

Table 5: Break Points in the Marginal Distributions of the Fiscal and Franchise Series: Filtered with War Dummies and GDP per Capita for the Central Government Data and for GDP per Capita for the Local Government Data

	Sequential L+1 vs. L	Sequential all	Global vs. none	Global L+1 vs. L	Global information criteria
Central government					
e_1	None	None	1866, 1885, 1889	1885, 1889	1885, 1889
e_2	None	None	None	None	None
G/Y	None	None	1875, 1879, 1884	1875, 1879, 1884	None
T/Y	1836, 1840, 1846, 1857	1836, 1840, 1846, 1857	1857, 1875, 1879, 1899, 1903	1857, 1875, 1879, 1899, 1903	None
T^d/T^{in}	1858, 1899, 1903, 1909	1858, 1879, 1899, 1903, 1909	1871, 1879, 1899, 1903, 1909	1871, 1879, 1899, 1903, 1909	1909
Top	1901	1901	1876, 1880, 1888, 1895, 1900	1876, 1880, 1888, 1895, 1900	None
Local government					
e_1^L	None	None	1869, 1872, 1909	1869, 1872, 1909	None
TAX/T	1877	1877	1882, 1892, 1901	1882, 1892, 1901	1871, 1873, 1875
G/N	1905	1905	1891, 1895, 1905	1891, 1895, 1905	None
T/N	1905	1905	1882, 1891, 1905	1905	1908, 1910

Note: The table reports the years with break points in the marginal distribution of the franchise and fiscal series. We perform five varieties of Bai and Perron tests in Eviews 9. Details are explained in the text and Appendix A4.

Table 6: Break Points in the Conditional Distributions of Fiscal Variables and Franchise Extension: Filtered with War Dummies and GDP per capita for the central government data and with GDP per capita for the local government data

	Sequential L+1 vs. L	Sequential all	Global vs. none	Global L+1 vs. L	Global information criteria
Panel A: The Conditional Distributions of the Fiscal series conditional on the Franchise series, $D(f\,\vert\, e)$					
Central government					
$D(G/Y\,\vert\, e_1)$	None	None	1853, 1857, 1878, 1899, 1903	1853, 1857, 1878, 1899, 1903	1853, 1857, 1878, 1899, 1903
$D(T/Y\,\vert\, e_1)$	1828, 1834, 1839, 1853, 1857	1828, 1834, 1839, 1853, 1857	1857, 1870, 1879, 1886, 1903	1857, 1870, 1879, 1886, 1903	1878, 1884
$D(T^d/T^{in}\big\vert\, e_1)$	1871, 1903, 1908	1871, 1879, 1884, 1903, 1908	1871, 1880, 1884, 1903, 1909	1871, 1880, 1884, 1903, 1909	1908
$D(Top\vert\, e_1)$	1869, 1873, 1878, 1883, 1887	1878, 1883, 1887, 1896, 1900	1853, 1862, 1876, 1886, 1900	1853, 1862, 1876, 1886, 1900	1873, 1886
$D(G/Y\,\vert\, e_2)$	None	None	1849, 1858, 1876, 1885, 1899	1849, 1858, 1876, 1885, 1899	1849, 1858, 1876, 1885, 1899
$D(T/Y\,\vert\, e_2)$	1834, 1846, 1857, 1885, 1903	1834, 1846, 1857, 1885, 1903	1857, 1870, 1885, 1896, 1905	1857, 1870, 1885, 1896, 1905	None
$D(T^d/T^{in}\big\vert\, e_2)$	None	None	1853, 1862, 1871, 1885, 1903	1853, 1862, 1871, 1885, 1903	1871, 1885, 1903
$D(Top\vert\, e_2)$	1888, 1899	1834, 1869, 1878, 1888, 1899	1834, 1858, 1871, 1885, 1899	1834, 1858, 1873, 1886	1870, 1886
Local government					
$D(\mathrm{TAX}/T\vert\, e_1^L)$	1876, 1882, 1891	1876, 1882, 1891	1874, 1889, 1883	1876, 1882, 1891	1874, 1879, 1883
$D(\mathrm{G}/N\vert\, e_1^L)$	1874, 1888, 1901	1874, 1888, 1901	1872, 1888, 1901	1874, 1888, 1901	1872, 1888, 1901
$D(\mathrm{T}/N\vert\, e_1^L)$	1874, 1884, 1901	1874, 1884, 1901	1874, 1884, 1901	1874, 1884, 1901	1874, 1884, 1901
Panel B: The Conditional Distributions of the Franchise Series conditional on the Fiscal Series, $D(e\vert\, f)$					
Central government					
$D(e_1\vert\, G/Y)$	None	None	1869, 1873, 1883, 1887, 1900	1869, 1873, 1883, 1887, 1900	1882, 1886
$D(e_1\vert\, T/Y)$	None	None	1831, 1844, 1869, 1885, 1894	1831, 1844, 1869, 1885, 1894	1869, 1885, 1894
$D(e_1\vert\, T^d/T^{in})$	None	None	1866, 1870, 1885, 1895, 1900	1866, 1870, 1885, 1895, 1900	1866, 1870, 1885, 1895
$D(e_1\vert\, Top)$	None	None	1834, 1869, 1873, 1885, 1889	1834, 1843, 1869, 1885, 1895	1869, 1885, 1889
$D(e_2\vert\, G/Y)$	1885, 1894, 1905	1885, 1894, 1905	1833, 1869, 1882,1894, 1905	1833, 1869, 1882,1894, 1905	1869, 1882, 1894
$D(e_2\vert\, T/Y)$	None	None	1857, 1869, 1884, 1893, 1903	1857, 1869, 1884, 1893, 1903	1869, 1884, 1893
$D(e_2\vert\, T^d/T^{in})$	None	None	1869, 1885, 1892, 1902, 1909	1869, 1885, 1892, 1902, 1909	1869, 1885, 1894
$D(e_2\vert\, Top)$	None	None	1835, 1858, 1869, 1884, 1893	1835, 1858, 1869, 1884, 1893	1835, 1869, 1884, 1893
Local government					
$D(e_1^L\big\vert\, \mathrm{TAX}/T)$	None	None	1870, 1878, 1882	1878, 1882, 1909	1870, 1878, 1882
$D(e_1^L\big\vert\, G/N)$	1872, 1879	1872, 1879	1872, 1876, 1881	1874, 1879	none
$D(e_1^L\big\vert\, T/N)$	1872, 1879	1872, 1879	1872, 1876, 1880	1872, 1879	1872

Note: The table reports the years in which there are (statistical) break points in conditional distribution of the fiscal series conditional on franchise series (panel A) and franchise series conditional on the fiscal series (panel B), respectively. We perform five varieties of Bai and Perron tests in Eviews 9. Details are explained in the text and Appendix A4.

Constitution, are associated with breaks in the scale and composition of taxation and in the top income tax rate. In contrast, only one of the structural breaks in the fiscal series overlap with the interventions in the franchise series identified by the narrative analysis, namely $\frac{G}{Y}$ which breaks around the 3rd reform act (in 1884). Finally, there are also break points that are inconsistent with the narratives and in some of the tests based on local search algorithms, no break points are identified at all.

Table 6 reports the break years in the distribution of fiscal series conditional on the franchise series (top of panel A) and the franchise series conditional on the fiscal series (top of panel B). In panel A, the 3nd reform act (1883-87) is identified as a break year in the series for the fiscal system conditional on the franchise in many of the tests. Occasionally, the 2nd reform act (1866-70) and the 1st reform act (1830-34) are also identified. In panel B, the 2nd and 3rd reform acts are almost universally identified as break points in the franchise series conditional on the fiscal series and there are also statistical breaks around the 1st reform act.

Based on these tests and the historical narrative, we apply the algorithm outlined in Section 3 to determine the nature of the causal relationship between the franchise for the House of Commons and the fiscal system of the central government (if there is one). This requires applying test A and B systematically to the pattern of breaks we find and check consistency with the historical narrative. Appendix A5 reports the details. Convincing evidence for a particular causal pattern must come from both tests: Test A tells us which causal relationship is consistent with historically verified interventions in the franchise series and test B tells us which causal relationship is consistent with the historically verified interventions in the fiscal system. Table 7 summarizes the results, with panel A showing the years when test A is conclusive and panel B showing the years when test B is conclusion.

We find one case in which test A, test B *and* the historical narrative are consistent with the Redistribution Hypothesis. First, the break pattern around the 3rd reform act in 1883-87 – statistical breaks in all series except $D(e_1|\frac{G}{Y})$ – implies that e_1 causes $\frac{G}{Y}$ (test A). Second, the break pattern around the beginning of the Gladstonian Fiscal Constitution – statistical breaks in all series related to $\frac{T}{Y}$ except $D(e_2)$ – implies that e_2 causes $\frac{T}{Y}$. However, the evidence should be judged against the fact that for the three other fiscal series than $\frac{G}{Y}$, test A suggests reverse causality from f to e around the 3rd reform act (1883-87) and the evidence from the 2nd reform act (1866-70) either points to independence or reverse causality from the fiscal system to the suffrage.[29]

[29]The indications of reverse causality from test A are not confirmed with any evidence from test B that fiscal innovation causes suffrage reform, and so is not compelling.

The fact that we find that the franchise causes government expenditure to grow, but not taxes raises the question as to what is in the evolution of $\frac{G}{Y}$ that is not in $\frac{T}{Y}$. The obvious answer is the deficit, which, in turn, suggests that shocks to the interest rate paid on debt may be confounding the relationship between the franchise and government size. This suspicion is bolstered by simple (autoregressive) regressions of $\frac{G}{Y}$ and of $\frac{T}{Y}$ on GDP per capita, the war dummy and the consol rate for government bonds, in which the consol rate is statistically significant in the former regression but not in the latter. If we add the consol rate to the set of variables used to filter the franchise and fiscal series, the result in Table 7 that e_1 causes $\frac{G}{Y}$ under test A disappears. The question, then, is whether it is appropriate to pre-filter with the consol rate. It is straightforward to argue that debt interest may cause breaks in government expenditure independently of the franchise, but less so to argue that the consol rate causes breaks in the franchise independently of its effect on the fiscal structure. In any event, however, the first of these channels of influence is enough to argue for pre-filtering (by OLS) both e and f with the consol rate, which is equivalent to running an OLS regression of (unfiltered) e on (unfiltered) f with the consol rate added as a control variable.[30]

In summary, the causal evidence that the Third Reform Act 1884 triggered an increase in government spending relative GDP has a question mark hanging over it because we do not find a similar effect for government revenue and because the pattern of statistical breaks required to pass test A is not robust to filtering with the consol rate. We must, therefore, conclude that we do not find clear evidence that e causes f for central government.

7.2 Local government

Table 5 reports (at the bottom) the years with structural breaks in the marginal distributions of the series that describe the municipal voting franchise and the evolution of the fiscal system of the Municipal Boroughs. We see that the two tests detect a break coinciding with the Municipal Franchise Act of 1869. The other potential interventions in the municipal franchise are not identified statistically as structural breaks. The marginal distributions of $\frac{T}{N}$ and $\frac{G}{N}$ exhibit a structural break coinciding with the Local Government Act of 1888. $\frac{\text{TAX}}{T}$ has a structural break (in 1871) consistent with the Municipal Franchise Act of 1869.

[30]The Frisch-Waugh-Lovell Theorem (see, e.g., Davidson and MacKinnon 1993, p. 19-23)) shows that instead of using a regression of, say, f on e along with control variable(s) (such as *war* and GDP per capita) added, the coefficient of interest (here on e) can be estimated by first regressing each of f and e on the control variables, and then regressing the residuals for f on the residuals for e, provided that all regressions are estimated with Ordinary Least Squares.

Table 7: Central government: Direction of Causality Based on the Patterns of Structural Break Points

		Years with			
Franchise	Fiscal	$e \perp f$	$e \leftrightarrow f$	$e \rightarrow f$	$f \rightarrow e$
Test A					
e_1	T^d/T^{in}	(1866-70)[a]			1883-87
e_2	T^d/T^{in}				
e_1	G/Y	(1866-70)[a]		1883-87[c]	
e_2	G/Y				
e_1	T/Y				1883-87, (1866-70)[a]
e_2	T/Y				
e_1	Top				1883-87, (1866-70)[a]
e_2	Top				
Test B					
e_1	T^d/T^{in}	1807-11			
e_2	T^d/T^{in}				
e_1	G/Y				
e_2	G/Y				
e_1	T/Y	1853-57			
e_2	T/Y	1844-48, 1994-00		1853-57	
e_1	Top	1894-00		(1894-00)[b]	
e_2	Top	1894-00			

Note: All data filtered with War Dummies and GDP per capita. Based on Test A and B from Table 2, the table reports the years in which it is possible to draw a causal conclusion based on the structural break test result reported in Table 5 and 6 and the historical narrative analysis summarized in Table 3. Test A requires that there is an intervention in the franchise series according to the historical narrative in the years with statistical breaks. Test B requires that there is an intervention in the fiscal series according to the historical narrative in the years with statistical breaks. $\perp$ means independent, $\leftrightarrow$ means causality in both directions, and $\rightarrow$ means direction of causality. For test A, the years 1883-87 are the 3rd reform act. For test B, the years 1844-48 are the repeal of the corn laws, 1853-57 are the beginning of the Gladstone's fiscal constitution, 1894-1900 is the end of the Gladstone's fiscal constitution, and 1904-08 is Asquith's budget. We allow for plus/minus two years around the event identified in the narrative analysis. a. One break only in $D(e_1)$ in this period. b. One break inly in $D(e_1|Top)$ in this period. c. The result is not robust to adding the consol rate on government bonds to the set of variables used to pre-filter the series.

Table 6 reports the break years in the distribution of the fiscal series conditional on the municipal franchise series (bottom of panel A) and the franchise series conditional on the fiscal series (bottom of panel B). In panel A, all the conditional distributions are stable in the interval around the Municipal Franchise Act of 1869, but occasionally exhibit breaks around the other potential interventions in the franchise series (e.g., in 1882 and 1884). Panel B records few instances in which the statistical breaks overlap

Table 8: Local Government: Direction of Causality Based on the Patterns of Structural Break Points

Franchise	Fiscal	Years with $e \perp f$	$e \leftrightarrow f$	$e \rightarrow f$	$f \rightarrow e$
Test A					
e_1^L	TAX/T	1868-72		(1868-72)[a]	
e_1^L	G/N				1868-72
e_1^L	T/N	1868-72			
Test B					
e_1^L	TAX/T	1873-77, 1900-04			
e_1^L	G/N	1887-91			
e_1^L	T/N	(1887-91)[b]			

Note: Based on Test A and B from Table 2, the table reports the years in which it is possible to draw a causal conclusion based on the structural break test result reported in Table 5 and 6 and the historical narrative analysis summarized in Table 3. Test A requires that there is an intervention in the franchise series according to the historical narrative in the years with statistical breaks. Test B requires that there is an intervention in the fiscal series according to the historical narrative in the years with statistical breaks. $\perp$ means independent, $\leftrightarrow$ means causality in both directions, and $\rightarrow$ means direction of causality. For test A, the years 1868-72 are the Municipal Franchise Act 1869. For test B, the years 1873-77 are the Public Health Act 1875, 1900-04 are the Education (Balfour) Act, and 1887-91 are the Local Government Act of 1888 We allow for plus/minus two years around the event identified in the narrative analysis. a. Only one test finds a break in $D(\frac{\text{TAX}}{T})$. If we use the stricter criterion that we need two tests to confirm a break, we conclude that $\frac{\text{TAX}}{T} \perp e_1^L$. b. Only one test finds a break in $D(\frac{T}{N})$ in this period.

with the interventions in the fiscal system identified by the narrative analysis, but some of the tests do suggest that $D(e_1^L \Big| \frac{G}{N})$ exhibits a break coinciding with the Public Health Act of 1875.

To determine the causal relationship between e and f for the Municipal Boroughs from these tests and the historical narrative, we, again, apply the Hoover algorithm. Table 8 reports the results. Based on test A, we find one indication that the Municipal Franchise Act in 1869 may have caused the subsequent drop in $\frac{\text{TAX}}{T}$. The required break in the marginal distribution for $\frac{\text{TAX}}{T}$ is, however, only detected by one of the Bai-Perron tests and if we use the stricter criterion that a break must be detected by two of the five tests, test A implies that the two series are independent. Figure A2 in Appendix A3, which plots the residuals or filtered series for $\frac{\text{TAX}}{T}$ and e_1^L over time, the break in 1868-72 appears genuine but hard to detect with the tests because we only have few years of fiscal data prior to the break. Test B, however, does not detect any instances consistent with $e \rightarrow f$

and the suffrage and fiscal structure appear to be independent. This casts doubt on the reliability of the finding from test A.

In conclusion, we do not find clear evidence that e causes f at the local level. If anything, the patterns of structural breaks suggest that $e \rightarrow f$ in the years around the Municipal Franchise Act 1869 but at a time where property taxes as a share of total revenue fell dramatically, possibly reflecting the fact that many of the newly enfranchised taxpayers wanted fiscal economizing and retrenchment, not fiscal expansion and redistribution (Szreter 1988; Daunton 2001; Aidt et al. 2010).

7.3 The choice of pre-filtering variables

An important consideration in the implementation of the Hoover approach is the way the data is filtered prior to doing the tests for structural breaks. We argue it is important to filter for the effect of war and shocks to the economic structure and possibly also in some cases for the effect of the consol rate. This may, however, lead to "over-smoothing" of the series, i.e., the possibility that genuine breaks in a series are masked by over-fitting it. Tables A5 to A7 in Appendix A6 reports the results from an analysis of central government data that are only filtered for war. With this more parsimonious pre-filtering, we do not find that e_1 causes $\frac{G}{Y}$ in 1883-87 because the marginal distribution of $\frac{G}{Y}$ conditional on e_1 exhibits a break (which is not present when we also pre-filter with GDP per capita and thus may be related to economic shocks), while the marginal distribution for $\frac{G}{Y}$ is stable. The fact that we can find a break in $\frac{G}{Y}$ when we filter with both *war* and GDP per capita but not when we only filter for *war* is obviously not a result of over-fitting, but may be due to the magnitude of the "break" in 1884 being small relative to the one in 1853-57. This makes it hard for the global Bai-Perron tests to detect any break in government size in the mid 1880s without first filtering with GDP per capita, which mitigates the magnitude of the breaks in 1853-57 and makes 1883-87 stand out. The evidence that the pattern of structural breaks associated with the beginning and end of Gladstone's fiscal rules are consistent with the Redistribution Hypothesis, however, remains irrespectively of how we filter the data. Overall, these robustness checks do not suggest that over-fitting is a serious concern, but they clearly show that the conclusion drawn in section 7.1 from Test A, which aims at finding evidence that franchise extension causes fiscal innovation, is sensitive to how the series are filtered.

8 Conclusion

We study the effect of franchise extension on fiscal structure in the United Kingdom between 1820 and 1913 in order to revisit the Redistribution Hypothesis - that extension of suffrage to lower income voters leads to more redistribution.

The Redistribution Hypothesis is one of the fundamental propositions in political economics. Empirically, it has proved difficult to establish a causal (positive) relationship between voting rights granted to low-income men and redistributive activity. To reconsider the hypothesis and to provide new causal evidence, we adopt the novel method proposed by Hoover (2001). The method is based on a combination of detailed historical narrative analysis and structural breaks tests in the conditional and marginal distributions of the franchise series and the fiscal series. We apply the method to the United Kingdom between 1820 and 1913. The United Kingdom is a promising case to study because of the way the franchise was gradually extended to lower income men in a sequence of reforms, but the approach we take can be adopted to other countries and time periods, and it would be valuable to do so.

Our analysis, both at the central and at the local government level, does not find clear and compelling evidence in support of the Redistribution Hypothesis, and, in this sense, mirrors the inconclusive findings from the many cross national studies. For central government, we find one instance, around the Third Reform Act 1884, where the historical narrative and the pattern of structural breaks are consistent with a causal effect from the suffrage to government spending relative to GDP. However, we cannot find a similar effect on the revenue side and the break pattern needed to support a causal interpretation disappears if we filter the series for the consol rate or if we do not filter for GDP per capita. This casts doubt on this finding. We also find instances in which interventions to the fiscal system - in particular the fiscal rules laid down by William Gladstone in the 1850s and dismantled in the 1890s - and the pattern of structural breaks are consistent with the Redistribution Hypothesis, in the sense that these structural shifts in fiscal policy did not affect the franchise process, which was evidently independent of them. For local government, we also find one instance - the Municipal Franchise Act 1869 - where the historical narrative of the suffrage and the statistical break patterns are consistent with the Redistribution Hypothesis, but this cannot be backed up with similar evidence from the narrative of the evolution of the fiscal system, which makes that evidence uncompelling.

To conclude, we speculate on the reasons why we are not, despite its logical appeal, able to clearly confirm the Redistribution Hypothesis. There are at least three possibilities. First, there is the problem of measuring redistribution. We do not have direct measures

of redistributive intent by the governments of the day. This would be the best outcome indicator since redistributive actions may not be effective. We do not even have rough measures such as the difference between the market or before fisc GINI coefficient and the after fisc (after tax and transfer) GINI. Second, there is the problem of complexity. There are many economic, social and political changes occurring at the same time in the history we (and others) explore that make the task of controlling for unrelated events difficult. The finding that the Gladstonian revolution in fiscal budgeting provides evidence in favor of the Redistribution process may be an example of this problem. Third and finally, the Hypothesis may be 'incomplete'. For example, it may be that the context in which franchise extension occurs is almost always one in which suffrage and redistribution are determined more or less simultaneously as part of broader social developments. In other words, the franchise, fiscal structure and redistribution may be joint indicators of broader or deeper social developments, the exogenous shocks to, or drivers of which are somewhere outside of what we think of when we formulate one or another version of the Redistribution Hypothesis. It is not obvious which, or which combination of these possibilities is the right one. Further work dealing with all of these problems would be highly valuable.

References

Acemoglu, D. (2006). A simple model of inefficient institutions. *Scandinavian Journal of Economics 108*(4), 515–546.

Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2008). Income and democracy. *The American Economic Review 98*(3), 808–842.

Acemoglu, D., S. Naidu, P. Restrepo, and J. Robinson (2015). Democracy, redistribution, and inequality. *Handbook of Income Distribution 2*, 1885–1966.

Acemoglu, D. and J. Robinson (2000). Why did the west extend the franchise? Democracy, inequality, and growth in historical perspective. *Quarterly Journal of Economics 115*(4), 1167–1199.

Aidt, T. S. and B. Dallal (2008). Female voting power: The contribution of women's suffrage to the growth of social spending in western europe (1869-1960). *Public Choice 134*(3-4), 391–417.

Aidt, T. S., M. Daunton, and J. Dutta (2010). The retrenchment hypothesis and the extension of the franchise in England and Wales. *Economic Journal 120*(547), 990–1020.

Aidt, T. S., J. Dutta, and E. Loukoianova (2006). Democracy comes to europe: Franchise extension and fiscal outcomes, 1830-1938. *European Economic Review 50*(2), 249–283.

Aidt, T. S. and D. Eterovic (2011). Political competition, electoral participation and public finance in 20th century Latin America. *European Journal of Political Economy 27*(1), 181–200.

Aidt, T. S. and P. S. Jensen (2009a). Tax structure, size of government, and the extension of the voting franchise in Western Europe, 1860-1938. *International Tax and Public Finance 16*(3), 362–394.

Aidt, T. S. and P. S. Jensen (2009b). The taxman tools up: An event history study of the introduction of the personal income tax. *Journal of Public Economics 93*(1-2), 160–175.

Aidt, T. S. and P. S. Jensen (2013). Democratization and the size of government: evidence from the long 19th century. *Public Choice 157*, 511–542.

Aidt, T. S. and P. S. Jensen (2014). Workers of the world, unite! Franchise extensions and the threat of revolution in Europe, 1820-1938. *European Economic Review 157*, 52–75.

Aidt, T. S. and P. S. Jensen (2017). From open to secret ballot: Vote buying and modernization. *Comparative Political Studies 50*(5), 555–593.

Aydelotte, W. O. (1967). The country gentlemen and the repeal of the corn laws. *The English Historical Review 82*(322), 47–60.

Babbage, C. (1852). *Thoughts on the principles of taxation with reference to a property tax and its exceptions* (3rd edition ed.). London: John Murray, Albemarle Street.

Bai, J. and P. Perron (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics 18*(1), 1–22.

Baysinger, B. and R. Tollison (1980). Chaining Leviathan: the case of Gladstonian finance. *History of Political Economy 12*, 206–213.

Bertocchi, G. (2011). The enfranchisement of women and the welfare state. *European Economic Review 55*(4), 535–553.

Blewett, N. (1965). The franchise in the United Kingdom 1885-1918. *Past and Present 32*, 27–56.

Boix, C. (2003). *Democracy and Redistribution*. Cambridge University Press.

Brennan, G. and J. M. Buchanan (1980). *The Power to Tax: Analytic Foundations of a Fiscal Constitution*. Cambridge, UK: Cambridge University Press.

Britannica, E. (1911). *Encyclopaedia Britannica* (11th edition ed.). Cambridge, UK: Cambridge University Press.

Britannica, E. (2009). *Encyclopaedia Britannica* (15th edition ed.). Cambridge, UK: Cambridge University Press.

Brock, M. G. (1973). *The Great Reform Act*. London: Hutchinson University Library.

Brückner, M. and A. Ciccone (2011). Rain and the democratic window of opportunity. *Econometrica 79*(3), 923–947.

Cannon, J. A. (1973). *Parliamentary reform, 1640-1832*. Cambridge, UK: Cambridge University Press.

Caramani, D. (2000). *Elections in Western Europe since 1815. Electoral results by constituencies*. Basingstoke, Hampshire, UK: Palgrave Macmillan Publishers Ltd.

Chapman, J. (2018). Democratic reform and opposition to government expenditure: Evidence from nineteenth-century britain. *Quarterly Journal of Political Science 13*(4), 363–404.

Congleton, R. (2011). *Perfecting parliament*. Cambridge, UK: Cambridge University Press.

Congleton, R. D. (2004). Economic development and democracy, does industrialization lead to universal suffrage? *Homo Economicus 21*, 283–311.

Corneo, G. and H. Gruner (2000). Social limits to redistribution. *American Economic Review 90*(5), 1491–1507.

Daunton, M. (2001). *Trusting Leviathan: The Politics of Taxation in Britain, 1799-1914*. Cambridge, UK: Cambridge University Press.

Davidson, R. and J. G. MacKinnon (1993). *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Davis, J. (2000). Central government and the towns. In M. Daunton (Ed.), *The Cambridge Urban History of Britain 1840-1950*, Volume III of *The Cambridge Urban History of Britain*, Chapter 9, pp. 261–286. Cambridge, UK: Cambridge University Press.

Dincecco, M. (2011). *Political transformations and public finances: Europe, 1650-1913*. Cambridge, UK: Cambridge University Press.

Doyle, B. M. (2000). The changing functions of urban government: Councillors, officials and pressure groups. In M. Daunton (Ed.), *The Cambridge Urban History of Britain 1840-1950*, Volume III of *The Cambridge Urban History of Britain*, Chapter 10, pp. 287–314. Cambridge, UK: Cambridge University Press.

Eaglesham, E. J. R. (1967). *The Foundations of Twentieth-Century Education in England*. Abingdon-on-Thames, UK: Routledge and Kegan Paul.

Evans, E. J. (2000). *Parliamentary reform, 1770-1918*. London: Pearson.

Fitzgerald, G. A. R. (1876). *The Ballot Act, 1872*. London: Stevens and sons.

Flora, P., J. Alber, R. Eichenberg, J. Kohl, F. Kraus, W. Pfenning, and S. Kurt (1983). *State, economy and society in Western Europe 1815-1975, vol. I and II*. Frankfurt, Germany: Campus Verlag.

Glen, W. C. (1885). *The Representation of the People Act, 1884 by Great Britain*. London: Shaw and Sons.

Gundlach, E. and M. Paldam (2009). A farewell to critical junctures: Sorting out long-run causality of income and democracy. *European Journal of Political Economy 25*(3), 340–354.

Hansard (1848). *Commons Sitting of March 3*, Volume 97 of *3rd series*, pp. Cols 192. London: House of Commons.

Harms, P. and S. Zink (2003). Limits to redistribution in a democracy: A survey. *European Journal of Political Economy 19*(4), 651–668.

Hettich, W. and S. L. Winer (1999). *Democratic choice and taxation. A A theoretical and empirican analysis*. Cambridge, UK: Cambridge Press.

Hicks, D. (2013). War and the political zeitgeist: Evidence from the history of female suffrage. *European Journal of Political Economy 31*, 60–81.

Hoover, K. (1991). The causal direction between money and prices. An alternative approach. *Journal of Monetary Economics 27*(3), 381–423.

Hoover, K. and S. Sheffrin (1992). Causation, spending, and taxes: Sand in the sandbox or tax collector for the welfare state? *American Economic Review 82*(1), 225–248.

Hoover, K. and M. Siegler (2000). Taxing and spending in the long view: The causal structure of us fiscal policy, 1791-1913. *Oxford Economic Papers 52*(4), 745–773.

Hoover, K. D. (2001). *Causality in Macroeconomics*. Cambridge UK: Cambridge University Press.

Husted, T. A. (1989). Demand for income redistribution benefits: The case of AFDC. *Southern Economic Journal 55*(3), 710–727.

Husted, T. A. and L. W. Kenny (1997). The effect of the expansion of the voting franchise on the size of government. *Journal of Political Economy 105*(1), 54–82.

Janowitz, M. (1976). Military institutions and citizenship in western societies. *Armed Forces & Society 2*(2), 185–204.

Keir, D. L. (1953). *The constitutional history of modern Britain, 1485-1951* (5th edition ed.). London: A. and C. Black.

Keith-Lucas, B. (1952). *The English Local Government Franchise*. Oxford: Basil Blackwell.

Keith-Lucas, B. (1977). *English Local Government in the Nineteeth and Tweentieth Centuries*. General Series 90. London: The Historical Association.

Kenny, L. and S. Winer (2006). Tax systems in the world: An empirical investigation into the importance of tax bases, administration costs, scale and political regime. *International Tax and Public Finance 13*(2-3), 181–215.

Kim, W. (2007). Social insurance expansion and political regime dynamics in Europe, 1880-1945. *Social Science Quarterly 88*(2), 494–514.

Leathers, C. G. (1986). Gladstonian finance and the Virginia school of public finance: Comment. *History of Political Economy 18*, 515–521.

Lindert, P. (1994). The rise of social spending, 1880-1930. *Explorations in Economic History 31*(1), 1–37.

Lindert, P. H. (2004a). *Growing public. Social spending and economic growth since the eighteenth century. Vol I: The story.* Cambridge UK: Cambridge University Press.

Lindert, P. H. (2004b). *Growing public. Social spending and economic growth since the eighteenth century. Vol II: Further evidence.* Cambridge UK: Cambridge University PRess.

Macnicol, J. (1998). *The Politics of Retirement in Britain 1878-1948.* Cambridge, UK: Cambridge University Press.

Maddison, A. (2003). *The world economy: Historical statistics.* Paris: OECD.

Mares, I. and D. Queralt (2015). The non-democratic origins of income taxation. *Comparative Political Studies 48*(14), 1974–2009.

Meltzer, A. and S. Richard (1981). A rational theory of the size of government. *Journal of Political Economy 89*(5), 914–927.

Meltzer, A. and S. Richard (1983). Tests of a rational theory of the size of government. *Public Choice 41*(3), 403–418.

Millward, R. (2000). The political economy of urban utilities. In M. Daunton (Ed.), *The Cambridge Urban History of Britain 1840-1950*, Volume III of *The Cambridge Urban History of Britain*, Chapter 11, pp. 315–350. Cambridge, UK: Cambridge University Press.

Mitchell, B. and P. Deane (1962). *Abstract of British historical statistics.* Cambridge, UK: Cambridge University Press.

Mitchell, B. R. (1988). *British historical statistics.* Cambridge, UK: Cambridge University Press.

Mood, A. and F. Graybill (1963). *Introduction to the Theory of Statistics* (Second Edition ed.). McGraw-Hill Series in Probability and Statistics. New York: McGraw-Hill Book Company, Inc.

Mueller, D. C. and T. Stratmann (2003). The economic effects of democratic participation. *Journal of Public Economics 87*(9), 2129 – 2155.

Musgrave, R. (1969). *Fiscal Systems.* New Haven: Yale University Press.

Pearl, J., M. Glymour, and N. P. Jewell (2016). *Casual Inference in Statistics: A Primer.* Wiley.

Peters, G. (1991). *The politics of taxation. A comparative perspective*. Oxford, UK: Basil Blackwell.

Piketty, T. (1995). Social mobility and redistributive politics. *Quarterly Journal of Economics 110*(3), 551–584.

Plumper, T. and C. W. Martin (2003). Democracy, government spending, and economic growth: A political-economic explanation of the Barro effect. *Public Choice 117*(1-2), 27–50.

Profeta, P., R. Puglisi, and S. Scabrosetti (2013). Does democracy affect taxation and government spending? Evidence from developing countries. *Journal of Comparative Economics 41*(3), 684 – 718.

Redlich, J. and F. W. Hirst (1958). *The History of Local Government in England*. London: MacMillan and co.

Sabine, B. E. F. (1966). *A history of income tax*. London: George Allen & Unwin.

Scheve, K. and D. Stasavage (2016). *Taxing the Rich: A History of Fiscal Fairness in the United States and Europe*. Princeton, US: Princeton University Press.

Schonhardt-Bailey, C. (2006). *From the Corn Laws to Free Trade: Interests, Ideas, and Institutions in Historical Perspective*. Cambridge, MA: MIT Press.

Seghezza, E. and P. Morelli (2019). Suffrage extension, social identity, and redistribution: the case of the Second Reform Act. *European Journal of Economic History*. In Press.

Seligman, E. R. A. (1911). *The income tax*. London: Macmillan.

Singer, J. D. and M. Small (1994). *Correlates of war project: International and civil war data, 1816-1992*. Ann Arbor, Michigan: Inter-University Consortium for Political and Social Research.

Smellie, K. B. (1946). *A History of Local Government*. London: London: Allen and Unwin.

Smith, F. B. (1966). *The making of the Second Reform Bill*. Cambridge, UK: Cambridge University Press.

Szreter, S. (1988). The importance of social intervention in Britain's mortality decline 1850-1914: A reinterpretation of the role of public health. *Social History of Medicine 1*, 1–37.

Ticchi, D. and A. Vindigni (2008). War and endogenous democracy. IZA Discussion Paper 3397, Institute for the Study of Labor, Bonn, Germany.

Tridimas, G. and S. Winer (2005). The political economy of government size. *European Journal of Political Economy 21*(3), 643–666.

Vine, J. S. (1878). *The Municipal Corporations*. London: Waterlow and Sons Limited.

Vine, J. S. (1882). *The English Municipal Code*. London: Waterlow Bros and Layton.

Waller, P. J. (1983). *Town, City and Nation, England 1850-1914*. Oxford, UK: Oxford University Press.

Winer, S. L. (2019). The political economy of taxation: Power, structure, redistribution. In B. G. R. Congleton and S. Voigt (Eds.), *The Oxford Handbook of Public Choice*, Volume 2, Chapter 24. Oxford, UK: Oxford University Press.

Supplementary Appendix.

A1 An Alternative Explanation of Why Filtering for War and GDP Is Important

A further explanation on why a pre-filtering for war and GDP per capita is important for determining causal relationships based on Hoover's method is provided by considering the following directed acyclic graph (DAG). (See for example, Pearl et al. (2016). Here E refers to the franchise. V is voting or electoral competition, and F represents fiscal structure. S refers to the social and economic conditions that might affect, E, F and V. The relations shown need not be linear. Other DAG's can be drawn, for example by including additional and/or unobserved variables.

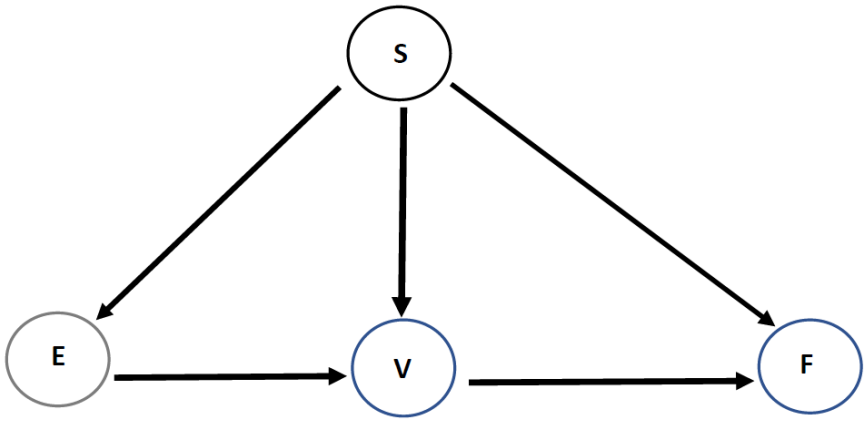


Figure A1: A directed acyclic graph for the Redistribution Hypothesis

We are interested in the causal effect of E on F. Here that effect is comprised of the causal effect of E on V, compounded with the causal effect of V on F. If we are dealing with stochastic relations, this compound effect will be the product of two conditional probabilities.

To estimate the causal effect of E on V, we need to block any back door path (giving rise to spurious correlation) between E and V. There are two such back-door paths:

$$E \leftarrow S \rightarrow V \text{ and } E \leftarrow S \rightarrow F \leftarrow V.$$

Here S is a confounder, and F is a collider. The collider at F blocks the second back-door path. It is sufficient to condition on S to block the first back door path.

To estimate the causal effect of V on F, we must block the back door paths:

$$V \leftarrow E \leftarrow S \rightarrow F \text{ and } V \leftarrow S \rightarrow F.$$

Conditioning on S is sufficient to block both paths. Now we are in a position to compare our application of Hoover's method to the situation illustrated in the figure. This comparison exposes two key assumptions that we make. First, there is the issue of representation of S. We pre-filter time series on E and on F using a war dummy and real GDP per capita, which are the only elements of S explicitly in our model of it. We are assuming that pre-filtering in this way is sufficient to block all back door paths from E to F. Second, we combine the effects of E on V and of V on F, which may in general be nonlinear, using the conditional distribution of F on E. If the relationships involved are both linear, then we can solve for this conditional distribution using the two linear casual relationships. So linearity is a second key assumption.

A2 Construction of the local government dataset

The data on local government have been constructed from primary sources. The British Parliamentary Papers contain from 1867 to 1913 the Local Taxation Returns revenue and expenditure accounts information for a variety of difference local government units in England and Wales. The governance structures for these different units differed. The data recorded are for the Municipal Corporations and for the Local Boards, Improvement Boards, Urban Sanitary Districts and Urban District Boards where the Corporation acted as these authorities. The names (and authority) of these boards changed over the period; for simplicity we will refer to them as Municipal Boroughs as a collective and it is understood that this include only the Urban Districts where the Municipal Corporation acted. Other local government bodies (education boards, Urban District not controlled by the Corporations, county councils, poor law unions etc.) are not included because they were not governed by the suffrage rules laid down by the House of Commons for the Corporations. Urban districts that were not Municipal Borough Corporations are not included. The rationale is that these (688) districts were not governed by the same suffrage as the boroughs till 1894 and we do not have any data on the voters in this districts which operated under a system of plural voting with richer taxpayers having more votes prior to that.

The British Parliamentary Papers and the book by Vine (1878) report information on the number of electors enfranchised for elections to the Municipal Boroughs for selected years.

Information about the population and the number of inhabited houses are obtained from the decennial Population Census of 1861, 1871, 1881, 1891, 1901 and 1911 and interpolated geometrically between the census years. The data is recorded for each Municipal Boroughs and for the area covered by the Urban District which till around 1901 differed.

Below we detail how the aggregate time series were constructed from the individual accounts for the Boroughs and from the information on the number of voters.

Definitions

- N: The size of the population to which the fiscal variable refers.
- $\frac{G}{N}$: Real (1871 prices) current and capital spending net of expenditure to repay debt (repayment of principal, interests and payment to sinking funds) and expenditures related to Burial Boards, Harbour and Port Authorities per capita.
- $\frac{T}{N}$: Real (1871 prices) current revenues net of income from Burial Boards, Harbour and Port Authorities per capita.
- $\frac{\mathbf{TAX}}{N}$: Real (1871 prices) property tax revenue (for the Corporations borough rates, general district rates, highway rates and other rates; for the Urban Districts borough rates, highway rates, general district rates and other public rates) per capita.
- $\frac{\mathbf{TAX}}{T}$: The share of property tax revenue as a percent of total current revenue.
- e_1^L: The share of the total population that registered to vote in municipal elections.

List of primary sources

- House of Commons, British Parliamentary Paper (1867). 'Return of the Number of Registered Voters on the List of All Municipal Boroughs in England and Wales and of the Numbers who Voted at the Municipal Elections in Those Boroughs for the Year 1852 and all Subsequent Years', vol. LVI, pp. 355-433.
- House of Commons, British Parliamentary Paper (1872). 'Return Showing With Respect to Each Municipal City and Borough in England and Wales, the Total Number of Municipal Electors on the Register now in Force', vol. XLVII.
- House of Commons, British Parliamentary Paper (1884-85). 'Return of Municipal Boroughs in England and Wales, Showing the Population, Number of Inhabited Houses, Number of Persons on the Burgess Rolls, Distinguishing Men From Women', vol. LXVII, pp. 23-31.
- House of Commons, British Parliamentary Paper (1897). 'Return Showing the Population and Number of Inhabited Houses (according to the Census of 1891), and the Number of Persons now on the Local Government Register of Electors each Administrative County (excluding the Municipal Boroughs), and for the County Borough and Municipal Borough in England and Wales'.
- House of Commons, British Parliamentary Paper (1867-1913), Local Taxation Returns.
- Population Census of Great Britain (1861, 1871, 1881, 1891, 1901 and 1911).

Construction of the fiscal series

The starting point is a digitalized version of the complete accounts for Municipal Corporations and the Local Boards, Improvement Commissions, Urban Sanitary Authorities and Urban District Boards. These exist starting with fiscal year 1867-68 and ending with the accounts of 1912-13. The accounting year ran from June to June and use the convention to index each set of accounts with the year in which the accounts starts (i.e., 1867-68 is indexed by 1867).

Units The number of Municipal Borough Corporations expends over time as more and more towns get incorporated and more and more corporations decide act as sanitary authorities. We only track the sanitary authorities where the Corporation acted, i.e., control the fiscal decisions. The Public Health and Local Government Acts adopted over period by the British Parliament changed the nature of the sanitary authorities. Before 1872, the Local Board and Improvement Commissions mostly established under the 1848 Public Health Act oversaw spending on sanitation (water, sewers, gas, street draining). In 1872, the Urban Sanitary Authorities were established and many Corporation adopted to act. In 1894, the Urban Sanitary Authorities were replaced by the Urban District Boards. To simplify, we refer to these bodies as the "Urban Districts". In 1888, the biggest of the Municipal Borough Corporations became County Boroughs.

The fact that the number of Corporations expanded over time and more and more Corporation decided to act as sanitary authorities implies that we need to normalize the fiscal variables with the size of the relevant population for which the fiscal item refers to make them comparable over time. Since we have information on the population in each Municipal Corporation and in each Urban District which until around 1897 covered different geographical areas, we construct for each year the population covered for Municipal Corporations and for the Urban Boards and uses these to normative the total fiscal sums.

Consistent time series It is not straight forward to construct time series of total revenue and expenditure over time because the format of the accounts change significantly in 1882 and 1902. The calculation of property tax revenue (TAX) is not affected by this, but there are minor variations in the titles of the headings represented the rates levied by the Corporation and the Urban Districts over time which might have created minor inconsistencies. However, as the time series is "smooth" without major year on year fluctuations, we do not believe this is a significant issue.

Until 1881, current and capital spending and revenue are combined; from 1882 onwards they are recorded separately. From 1902 the revenues and expenditures related to Burial Boards, Harbour and Port Authorities are added to the accounts for the Urban Boards while they were previously not.

The change in the reporting format in 1881-82 causes two problems for the time series total revenue (T) and total expenditure G. For T the problem is that prior to 1881 income and expenditures from loans included loans used to repay old debt while after loans income is net of this. For this reason, we need to net loans income out prior to 1881 to get a consistent record of the evolution of current revenues. It is not possible to construct a series for loans income. For G, the problem is that the total before 1881

include spending out of loans and we cannot separate it out, so we need to construct an expenditure series that include spending out of loans but adding the total from the loans accounts after 1882 to the current spending. However, this requires netting out "Loans repaid with interest during the year" before and after 1881 because till 1881 this item included loans repaid with new loans while it does not after. Since "loans repaid with interest during the year" are sometime combined with payments to sinking funds, this must be netted out as well for all years. So, in short, for revenue we can construct a consistent current revenue series while for spending we can construct a consistent current plus capital spending series (net of loan repayments).

From 1902, we need to net out the revenues and expenditures related to Burial Boards, Harbour and Port Authorities.

The nominal revenue and expenditures series are converted to real quantities with the Sauerbeck-Statisk price index from Mitchell (1988) with base year 1871.

Construction of the suffrage series

The Municipal Corporation Act of 1835 specified that every man who was an inhabitant householder within the borough or within seven miles thereof and who had occupied any house, warehouse, counting house[31] or shop within the borough for the previous two and a half years, provided he had been rated for the whole of that period and had paid the rates, and was duly enrolled on the Burgess Roll. Those, who, within the previous twelve months, had received parochial relief or other alms (under the Poor Laws) were excluded, and provision was made to enable a tenant, in cases where the landlord was liable for the rate, to pay them himself and so qualify for the franchise (see Vine 1878, p 44).

While elections were held each year (with one-third on the Municipal Borough council up for election), a complete record of the numbers of electors is not preserved. The British Parliamentary Papers, however, contain four returns detailing the number of voters in each borough in 1865, 1871, 1884 and 1897 and Vine (1878) contains an addition cross section of data for 1878. For each borough, we calculate the number of voters per capita and then average across the boroughs in existence at each date. To create the time series, we assume that the number of voters increases in the years with a national reform 1869, 1878, 1882 and 1894 and that this is reflected in the cross section available at the time of the reform or immediately after. We assume that the number of voters per capita is constant in between the years with reform. Since the suffrage was linked to property tax payments and subject to the residential requirement, there will have been movements in some localities year-on-year but we cannot track this and assume that these movements are neutralized on average by movements in total population.

[31] A counting house, or computing house is the building, room, office or suite in which a business firm carries on operations, particularly accounting.

A3 Tables

Table A1: Test of Stationarity of Franchise and Fiscal Series

	Optimal lag order	ADF
Panel A: Central government, filter war and GDP		
e_1	1	I(0)
e_2	1	I(0)
G/Y	1	I(0)
T/Y	2	I(1)
T^d/T^{in}	4	I(1)
Top	2	I(1)
Panel B: Local government, filter GDP		
e_1^L	1	I(0)
G/N	2	I(0)
T/N	2	I(1)
TAX/T	2	I(1)

Note: The table shows the ADF test results for stationarity of franchise and fiscal series. We preform the tests on series filtered with the War Dummy and GDP per capita for the central government sample and GDP per capita only for the local government sample.

Table A2: Regression Model Specifications for the Central Government Sample: Filtered with War Dummies and GDP

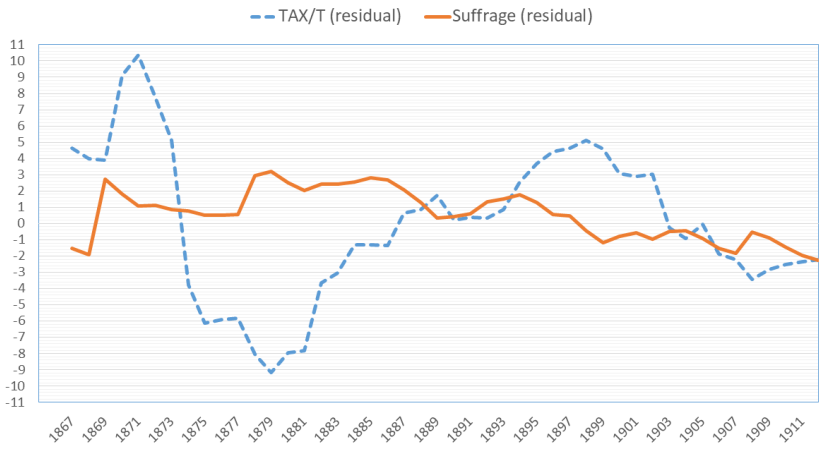
	ARDL (p,q)	Regression
Marginal distribution		
e_1	AR(1)	X = L.X
e_2	AR(1)	X = L.X
G/Y	AR(2)	Y = L.Y + L2.Y
T/Y	AR(3)	Y = L.Y + L3.Y + c
T^d/T^{in}	AR(2)	Y = L.Y + L2.Y + c
Top	AR(1)	Y = L.Y + c
Conditional Y on X		
G/Y on e_1	ARDL(1,0)	Y = L.Y + X
G/Y on e_2	ARDL(3,1)	Y = L.Y + X + L.X + c
T/Y on e_1	ARDL(3,1)	Y = L.Y + X + L.X
T/Y on e_2	ARDL(3,0)	Y = L.Y + L3.Y + X + c
T^d/T^{in} on e_1	ARDL(2,0)	Y = L.Y + L2.Y + X
T^d/T^{in} on e_2	ARDL(2,2)	Y = L.Y + L2.Y + X + c
Top on e_1	ARDL(1,1)	Y = L.Y + X + L.X
Top on e_2	ARDL(1,1)	Y = L.Y + X + L.X + c
Conditional X on Y		
e_1 on G/Y	ARDL(1,4)	X = L.X + L3.Y + L4.Y + c
e_2 on G/Y	ARDL(2,4)	X = L.X + Y + L.Y + L3.Y + L4.Y
e_1 on T/Y	ARDL(1,1)	X = L.X + Y + L.Y + c
e_2 on T/Y	ARDL(1,4)	X = L.X + L.Y + L2.Y + L3.Y + L4.Y
e_1 on T^d/T^{in}	ARDL(1,0)	X = L.X + Y
e_2 on T^d/T^{in}	ARDL(1,4)	X = L.X + Y + L3.Y
e_1 on Top	ARDL(1,1)	X = L.X + Y + L.Y
e_2 on Top	ARDL(1,3)	X = L.X + Y + L.Y + L2.Y + L3.Y

Note: The table shows the full ARDL(p,q) model as well as the parsimonious form of the model specification for further structural break tests on the data filtered with war dummies as well as GDP. Y and X refer to Y_t and X_t. $Lk.Y$ and $Lk.X$ refers to Y_{t-k} and X_{t-k}. c is the constant. AIC is used to select the lag structure. We also perform some post-estimation tests to validate the models. Firstly, we calculate Durbin-Watson (DW) statistics to detect the presence of autocorrelation in the residuals (prediction errors) from the regression analysis. We find that the errors of our models are serially independent. Secondly, we conduct Augmented Dickey Fuller (ADF) test on the residuals of the models to show that they are dynamically stable. These results are available upon request.

Table A3: Regression Model Specifications for the Local Government Sample: Filtered with GDP

	ARDL (p,q)	Regression
Marginal distribution		
e_1^L	AR(1)	X = L.X
G/N	AR(2)	Y = L.Y
T/N	AR(2)	Y = L.Y + L2.Y
TAX/T	AR(2)	Y = L.Y + L2.Y
Conditional Y on X		
G/N on e_1^L	ARDL(1,2)	Y = L.Y + X + L2.X + c
T/N on e_1^L	ARDL(1,2)	Y = L.Y + X + L.X + L2.X + c
TAX/T on e_1^L	ARDL(2,1)	Y = L.Y + L2.Y + X + L.X
Conditional X on Y		
e_1^L on G/N	ARDL(1,1)	X = L.X + Y + L.Y
e_1^L on T/N	ARDL(1,1)	X = L.X + Y + L.Y
e_1^L on TAX/T	ARDL(1,0)	X = L.X + Y

Note: The table shows the full ARDL(p,q) model as well as the parsimonious form of the model specification for further structural break tests on the data filtered with GDP. Y and X refer to Y_t and X_t. $Lk.Y$ and $Lk.X$ refers to Y_{t-k} and X_{t-k}. c is the constant. AIC is used to select the lag structure.



Notes: The figure plots the residuals from the series of $\frac{\text{TAX}}{T}$ and e_1^L filtered with GDP per capita.

Figure A2: The residuals from the series of $\frac{\text{TAX}}{T}$ and e_1^L filtered with GDP per capita.

A4 Details of the specification of the structural breaks

Stationarity tests

We need to establish the order of integration of all the variables for both central and local government data. We begin by finding the optimal lag order for the stationarity test of both franchise and fiscal variables. We first obtain lag-order selection statistics and find the optimal lag-order for stationarity test.[32] These are sequences of likelihood-ratio test statistics for all the full autoregression models of order less than or equal to the highest lag order. There are four selection criteria (the final prediction error, Akaike's information criterion, Schwarz's Bayesian information criterion and the Hannan and Quinn information criterion) and we select the lag order which is consistent with as many criteria as possible. The optimal lag orders are reported in the Appendix table A1.

Next, we conduct stationarity tests for both franchise and fiscal variables to see if any variables are I(2). We first use standard Augmented Dickey-Fuller Test (ADF) to identify unit roots. The lags are specified based on the optimal lag order already identified and the unit roots are identified at the 95% critical value. Many variables in the Appendix table A1 are I(1) but none of the series is I(2). This indicates that Autoregressive Distributed Lag (ARDL) models with an error correction form are appropriate for the structural breaks analysis.

As the series might incorporate structural breaks, we do a robustness check with the Zivot-Andrews Unit Root test which allows for a single break in intercept. This confirms that none of the series is I(2).

Implementation of the Bai-Perron test

Here are some details on how we operate Bai-Perron test in EViews 9. Standard errors in the model specification are ordinary for regressions without autocorrelation and HAC for series with autocorrelation in the error term (We choose HAC Newey-west method as the HAC option as is recommended on the official EViews manual). Among the whitening options we choose the Quadratic-spectral for kernel selection also as recommended. In the central government data, as there are 3 obvious break points in the franchise series, to maximise the power of break tests, we choose 3 as the maximal number of break points in the franchise series. For the fiscal series and all conditional distributions, we start from 3 and sequentially increase the number, until the number of break points does not further increase with the maximal number of break points we set. For the local government data, we choose 5 as the maximal number of break points in the franchise series based on the narratives. For the fiscal series and all conditional distributions, we set the maximal number of break points to be 3. It is also required that data at the very left and right of the interval should be trimmed. We use 5% level. If it leads to singular matrix, we then use the trimmed level which is as small as possible (10% or 15%). We test the structural breaks on both the coefficients of the level (or their lagged terms) of the independent variables as well as the constant.

[32] We set maximum lag order to 4 which is the default and use the confidence level 95%.

A5 Application of the Hoover algorithm

In this appendix, we explain how we applied test A and B to generate the results in Table 7 and Table 8. For ease of reference, the algorithm behind the two tests are repeated here.

Test A: (i) check that $D(e|f)$ and $D(e)$ break in year t; (ii) use Table 1A to check if in year t, $D(f)$ and $D(f|e)$ are stable or not and draw the inference indicated.

Test B: (i) check that $D(f|e)$ and $D(f)$ break in year t; (ii) use Table 1B to check if in year t, $D(e)$ and $D(e|f)$ are stable or not and draw the inference indicated.

Table A4: How to determine the causal order from patterns of structural breaks

Table A

	$D(f)$		
		Stable	Unstable
$D(f\|e)$	Stable	$f \perp e$	$e \rightarrow f$
	Unstable	$e \leftarrow f$	$f \leftrightarrow e$

Table B

	$D(e)$		
		Stable	Unstable
$D(e\|f)$	Stable	$f \perp e$	$e \leftarrow f$
	Unstable	$e \rightarrow f$	$f \leftrightarrow e$

Notes: Panel A or B are used to draw conclusions about the causal order conditional on test A or test B having been passed. The arrows indicate the direction of causality and $\perp$ means that the two are independent.

A prerequisite for drawing causal inferences from test A and test B is that the historical narrative clearly indicates that there was an intervention in the suffrage (test A) or in the fiscal system (test B) in year t. We allow for an interval around each intervention identified in the historical narrative analysis of plus/minus two years and associate a statistical break point within these intervals with the intervention from the narrative analysis. For the central government data, the intervals for the interventions in suffrage series are defined related to the year of the first election under the new rules; for the local government data, we note that one-third of the councils were elected each year and the narrative intervals are defined relative to year after the reform (so that two-third of the councilor are elected on the new rules). For the interventions in the fiscal series which are timed to a particular year, we define the intervals relative to the corresponding fiscal year. We call these intervals the *narrative intervals*.

Constructing Table 7: Central government data

The national level raw series have been filtered with the war dummies and with real GDP per capita and the structural breaks tests are preformed on the residuals from the associated regressions.

Test A

For test A, the relevant narrative intervals are (see Table 3, column four)

- **First Reform Act 1882, first election in 1832**: 1830-34.
- **Second Reform Act 1867, first election in 1868**: 1866-70.
- **Secret Ballot 1872, first election in 1874**: 1872-76
- **Corrupt Practices Act 1883; Third Reform Act 1884; Redistribution Act 1885; first election in 1885**: 1883-87.
- **Parliamentary Act 1911**: 1909-13.

There are, according to Table 5, no statistical breaks in $D(e_2)$. This means that test A is not informative about the relationship between e_2 and the fiscal variables. $D(e_1)$ breaks in 1885 according to three of the five tests, i.e., in the interval 1883-87 but there is only one test that indicates a break in 1866, so this does not qualify as a break point under the "two breaks rule". We record below what inference can be drawn if we accept that break in the 1866-70 interval for e_1 based on one test only.

Variable $\frac{T^d}{T^{in}}$. Check if $D(e_1|\frac{T^d}{T^{in}})$ breaks within the interval 1883-87 and in 1866-70. From Table 6, panel B, we conclude

- $D(e_1|\frac{T^d}{T^{in}})$ breaks in 1885, i.e., in the 1883-87 interval.
- $D(e_1|\frac{T^d}{T^{in}})$ breaks in 1866, i.e., in the 1866-70 interval.

Check if $D(\frac{T^d}{T^{in}})$ breaks or is stable in the 1883-87 and 1866-70 intervals. From Table 5, we conclude that

- $D(\frac{T^d}{T^{in}})$ is stable in the 1883-87 interval.
- $D(\frac{T^d}{T^{in}})$ is stable in the 1866-70 interval (it breaks in 1871).

Check if $D(\frac{T^d}{T^{in}}\Big|e_1)$ breaks or is stable in the 1883-87 and 1866-70 intervals. From Table 6, panel A, we conclude that $D(\frac{T^d}{T^{in}}\Big|e_1)$ breaks in the 1883-87 interval (in 1884) but not in the 1866-70 interval (it breaks in 1871). From Table A4A, we conclude that $\frac{T^d}{T^{in}} \rightarrow e_1$ in 1883-87 and that $\frac{T^d}{T^{in}} \perp e_1$ in 1866-70.

Variable $\frac{G}{Y}$. Check if $D(e_1|\frac{G}{Y})$ breaks within the intervals 1883-87 and 1866-70. From Table 6, panel B, we conclude that

- $D(e_1|\frac{G}{Y})$ breaks in 1886 and 1887, i.e., in the 1883-87 interval.
- $D(e_1|\frac{G}{Y})$ breaks in 1869, i.e., in the 1866-70 interval.

Check if $D(\frac{G}{Y})$ breaks or is stable in the 1883-87 and 1866-70 intervals. From Table 5, we conclude that

- $D(\frac{G}{Y})$ breaks in 1884, i.e., in the 1883-87 interval.
- $D(\frac{G}{Y})$ is stable in the 1866-70 interval.

Check if $D(\frac{G}{Y}\Big|\, e_1)$ breaks or is stable in the 1883-87 and 1866-70 intervals. From Table 6, panel A, we conclude that $D(\frac{G}{Y}\Big|\, e_1)$ is stable in both intervals. From Table A4A, we conclude that $e_1 \rightarrow \frac{G}{Y}$ in 1883-87 and that $e_1 \perp \frac{G}{Y}$

Variable $\frac{T}{Y}$. Check if $D(e_1|\, \frac{T}{Y})$ breaks within the 1883-87 and 1866-70 intervals. From Table 6, panel B, we conclude that

- $D(e_1|\, \frac{T}{Y})$ breaks in 1885, i.e., in the 1883-87 interval.
- $D(e_1|\, \frac{T}{Y})$ breaks in 1869, i.e., in the 1866-70 interval.

Check if $D(\frac{T}{Y})$ breaks or is stable in the 1883-87 and 1866-70 intervals. From Table 5, we conclude that

- $D(\frac{T}{Y})$ is stable in the 1883-87 interval.
- $D(\frac{T}{Y})$ is stable in the 1866-70 interval.

Check if $D(\frac{T}{Y}\Big|\, e_1)$ breaks or is stable in the 1883-87 and 1866-70 intervals. From Table 6, panel A, we conclude that $D(\frac{T}{Y}\Big|\, e_1)$ breaks in 1870, 1884 and 1886, i.e., in the intervals 1866-70 and 1883-87. From Table A4A, we conclude that $e_1 \leftarrow \frac{T}{Y}$ in 1866-70 and 1883-87.

Variable Top. Check if $D(e_1|\, Top)$ breaks within the intervals 1883-87 and 1866-70. From Table 6, panel B, we conclude that

- $D(e_1|\, Top)$ breaks in 1885, i.e., in the 1883-87 interval.
- $D(e_1|\, Top)$ breaks in 1869, i.e., in the 1866-70 interval.

Check if $D(Top)$ breaks or is stable in the 1883-87 and 1866-70 intervals. From Table 5, we conclude that

- $D(Top)$ is stable in the 1883-87 interval (there is a break in 1888).
- $D(Top)$ is stable in the 1866-70 interval.

Check if $D(Top|\, e_1)$ breaks or is stable in the interval 1883-87 and 1866-70. From Table 6, panel A, we conclude that $D(Top|\, e_1)$ breaks in both intervals (in 1869, 1883, 1886, and 1887). From Table A4A, we conclude that $e_1 \leftarrow Top$ in 1883-87 and in 1866-70.

Test B

For test B, the relevant narrative intervals are (see Table 3, column four)

- **Income tax, 1842**: 1840-44.
- **Corn Law, 1846**: 1844-48.
- **Gladstone's fiscal constitution established**: 1853-57.
- **Gladstone's fiscal constitution abolished**: 1894-1900.

- **Asquith's budget, 1906**: 1904-08.
- **Lloyd George's budget, 1909**: 1907-11.

Variable $\frac{T^d}{T^{in}}$. The following statistical breaks in $D(\frac{T^d}{T^{in}})$ recorded in Table 5 coincide with the narrative intervals for the fiscal system listed above:

- $D(\frac{T^d}{T^{in}})$ breaks in 1899 and 1909 according to at least two tests and we conclude that $D(\frac{T^d}{T^{in}})$ breaks statistically within the narrative intervals 1894-1900 and 1907-11. We observe two breaks just outside the narrative intervals 1904-08 (in 1903) and 1853-57 (in 1858), but we maintain that they are outside.

Check if $D(\frac{T^d}{T^{in}}\big|\, e_1)$ and $D(\frac{T^d}{T^{in}}\big|\, e_2)$ break within the intervals 1894-1900 and 1907-11. From Table 6, panel A, we conclude that

- $D(\frac{T^d}{T^{in}}\big|\, e_1)$ breaks in 1908 and 1909, i.e., in the 1907-11 interval but is stable in the interval 1894-00.
- $D(\frac{T^d}{T^{in}}\big|\, e_2)$ is stable in both intervals, so no conclusion can be drawn from test B about the causal relationship between e_2 and $\frac{T^d}{T^{in}}$.

Check if $D(e_1)$ and $D(e_1|\, \frac{T^d}{T^{in}})$ break or are stable in the 1907-11 interval from Table 5 and Table 6, panel B. We find

- $D(e_1)$ is stable in the 1907-11 interval.
- $D(e_1|\, \frac{T^d}{T^{in}})$ is stable in 1907-11 interval.

From Table A4B, we conclude that $e_1 \perp \frac{T^d}{T^{in}}$ in 1907-11.

Variable $\frac{G}{Y}$. There are no statistical breaks in $D(\frac{G}{Y})$ recorded in Table 5 that coincide with the narrative intervals for the fiscal system. Thus, test B is not informative about the causal relationship between e_1 and e_2 on the one hand and $\frac{G}{Y}$ on the other.

Variable $\frac{T}{Y}$. The following statistical breaks in $D(\frac{T}{Y})$ recorded in Table 5 coincide with the narrative intervals for the fiscal system listed above:

- $D(\frac{T}{Y})$ breaks in 1840, 1846, 1857, and 1899 according to at least two tests; we conclude that $D(\frac{T}{Y})$ breaks statistically within the narrative intervals 1840-44, 1844-48, 1853-57 and 1894-1900.

Check if $D(\frac{T}{Y}\big|\, e_1)$ and $D(\frac{T}{Y}\big|\, e_2)$ break within the intervals 1840-44, 1844-48, 1853-57 and 1894-1900. From Table 6, panel A, we conclude that

- $D(\frac{T}{Y}\big|\, e_1)$ breaks in 1853 and 1857, i.e., within the interval 1853-57 but not in the other three intervals.
- $D(\frac{T}{Y}\big|\, e_2)$ breaks in 1846, 1857 and 1896, i.e., within the intervals 1844-48, 1853-57 and 1894-00.

Check if $D(e_1)$ and $D(e_1|\,\frac{T}{Y})$ break or are stable in the 1853-57 interval from Table 5 and Table 6, panel B. We find

- $D(e_1)$ is stable in the 1853-57 interval.
- $D(e_1|\,\frac{T}{Y})$ is stable in 1853-57.

From Table A4B, we conclude that $e_1 \perp \frac{T}{Y}$ in 1853-57.
Check if $D(e_2)$ and $D(e_2|\,\frac{T}{Y})$ break or are stable in the intervals 1844-48, 1853-57, and 1894-00 from Table 5 and Table 6, panel B. We find

- $D(e_2)$ is stable in the intervals 1844-48, 1853-57 and 1894-00.
- $D(e_2|\,\frac{T}{Y})$ breaks in 1857, i.e., in the 1853-57 interval but is stable in the other two, although we observe that there is a break in 1893.

From Table A4B, we conclude that $e_2 \perp \frac{T}{Y}$ in 1844-48 and 1894-00 and that $e_2 \rightarrow \frac{T}{Y}$ in 1853-57.

Variable Top**.** The following statistical breaks in $D(Top)$ recorded in Table 5 coincide with the narrative intervals for the fiscal system listed above:

- $D(Top)$ breaks in 1895 and in 1900 according to at least two tests; we conclude that $D(Top)$ breaks statistically within the narrative interval 1894-00.

Check if $D(Top|\,e_1)$ and $D(Top|\,e_2)$ break within the interval 1894-1900. From Table 6, panel A, we conclude that

- $D(Top|\,e_1)$ breaks in 1896 and 1900, i.e., within the interval 1894-1900.
- $D(Top|\,e_2)$ breaks in 1899, i.e., within the interval 1894-1900.

Check if $D(e_1)$ and $D(e_1|\,Top)$ break or are stable in the 1894-00 interval from Table 5 and Table 6, panel B. We find

- $D(e_1)$ is stable in the 1894-1900 interval.
- $D(e_1|\,Top)$ breaks according to one test in 1895, but this is not sufficient to conclude that there is a break in the 1894-1900 interval and we conclude it is stable.

From Table A4B, we conclude that $e_1 \perp Top$ in 1894-1900.
Check if $D(e_2)$ and $D(e_2|\,Top)$ break or are stable in the interval 1894-1900 from Table 5 and Table 6, panel B. We find

- $D(e_2)$ is stable in the interval 1894-1900.
- $D(e_2|\,Top)$ is stable in the interval 1894-1900, although we observe that there is a break in 1893.

From Table A4B, we conclude that $e_2 \perp Top$ in 1894-1900.

Constructing Table 8: Local government data

The local government level raw series have been filtered with real GDP per capita and the structural breaks tests are preformed on the residuals from the associated regressions.

Test A

For test A, the relevant narrative intervals are (see Table 4, column four)

- **Municipal Franchise Art 1869**: 1868-72.
- **The Parliamentary and Municipal Registration Act 1878**: 1877-81.
- **Municipal Corporation Act 1882**: 1881-85.
- **County Electors Act**: 1887-91.
- **Local Government Act**: 1893-97.

There are, according to Table 5, $D(e_1^L)$ breaks in 1869 and 1872 according to two of the five tests, i.e., in the interval 1868-72.

Variable $\frac{\mathbf{TAX}}{T}$. Check if $D(e_1^L \Big| \frac{\text{TAX}}{T})$ breaks within the interval 1868-72. From Table 6, we conclude

- $D(e_1^L \Big| \frac{\text{TAX}}{T})$ breaks in 1870 according to two tests, i.e., in the 1868-72 interval.

Check if $D(\frac{\text{TAX}}{T})$ breaks or is stable in the 1868-72 interval. From Table 5, we conclude that

- $D(\frac{\text{TAX}}{T})$ breaks according to one test in 1871.

Check if $D(\frac{\text{TAX}}{T} \Big| e_1^L)$ breaks or is stable in the 1868-72 interval. From Table 6, there are no breaks in this interval, so we conclude that $D(\frac{\text{TAX}}{T} \Big| e_1^L)$ is stable in this interval. From Table A4A, we conclude based on one break in $D(\frac{\text{TAX}}{T})$ that $e_1^L \to \frac{\text{TAX}}{T})$. If we require two tests to confirm the break in 1871, we $\frac{\text{TAX}}{T} \perp e_1^L$ in 1868-72.

Variable $\frac{G}{N}$. Check if $D(e_1^L \Big| \frac{G}{N})$ breaks within the interval 1868-72. From Table 6, we conclude that

- $D(e_1^L \Big| \frac{G}{N})$ breaks in 1872 according to three tests, i.e., in the 1868-72 interval.

Check if $D(\frac{G}{N})$ breaks or is stable in the 1868-72 interval. From Table 5, we conclude that

- $D(\frac{G}{N})$ is stable in the 1868-72 interval.

Check if $D(\frac{G}{N} \Big| e_1^L)$ breaks or is stable in the 1868-72 interval. From Table 6, we conclude that $D(\frac{G}{N} \Big| e_1^L)$ breaks in this interval in 1872 according to two tests. From Table A4A, we conclude that $\frac{G}{N} \to e_1^L$ in 1869-72.

Variable $\frac{T}{N}$. Check if $D(e_1^L \Big| \frac{T}{N})$ breaks within the 1868-72 interval. From Table 6, we conclude that

- $D(e_1^L \Big| \frac{T}{N})$ breaks in 1872 according to two tests, i.e., in the 1869-72 interval.

Check if $D(\frac{T}{N})$ breaks or is stable in the 1868-72 interval. From Table 5, we conclude that

- $D(\frac{T}{N})$ is stable in the 1868-72 interval.

Check if $D(\frac{T}{N} \Big| e_1^L)$ breaks or is stable in the 1868-72 interval. From Table 6, we conclude that $D(\frac{T}{N} \Big| e_1^L)$ is stable in this interval. From Table A4A, we conclude that $e_1^L \perp \frac{T}{N}$ in 1868-72.

Test B

For test B, the relevant narrative intervals are (see Table 4, column four)

- **Public Health Act, 1875**: 1873-77.
- **The Local Government Act, 1888, implemented 1889**: 1887-91.
- **The Education (Balfour) Act**, 1902: 1900-1904.

Variable $\frac{\mathbf{TAX}}{T}$. The following statistical breaks in $D(\frac{\text{TAX}}{T})$ recorded in Table 5 coincide with the narrative intervals for the local fiscal system listed above:

- $D(\frac{\text{TAX}}{T})$ breaks in 1873, 1875 and 1877 according to several tests and we conclude that $D(\frac{\text{TAX}}{T})$ breaks in the interval 1873-77.
- $D(\frac{\text{TAX}}{T})$ breaks in 1901 according to two tests and we conclude that $D(\frac{\text{TAX}}{T})$ breaks in the interval 1900-04.

Check if $D(\frac{\text{TAX}}{T} \Big| e_1^L)$ break within the intervals 1873-1977 and 1900-04. From Table 6, we conclude that

- $D(\frac{\text{TAX}}{T} \Big| e_1^L)$ breaks in 1874 and 1876 according to several tests i.e., in the 1873-77 interval but is stable in the interval 1900-04.

Check if $D(e_1^L)$ and $D(e_1^L \Big| \frac{\text{TAX}}{T})$ break or are stable in the 1873-77 or 1900-04 interval from Table 5 and Table 6. We find

- $D(e_1^L)$ is stable in the 1873-77 and 1900-04 intervals.
- $D(e_1^L \Big| \frac{\text{TAX}}{T})$ is stable in the 1873-77 and 1900-04 intervals.

From Table A4B, we conclude that $e_1^L \perp \frac{\text{TAX}}{T^{in}}$ in the 1873-77 and 1900-04 intervals.

Variable $\frac{G}{N}$. The following statistical breaks in $D(\frac{G}{N})$ recorded in Table 5 coincide with the narrative intervals for the fiscal system listed above:

- $D(\frac{G}{N})$ breaks in 1891 according to two tests; we conclude that $D(\frac{G}{N})$ breaks statistically within the narrative intervals 1887-91.

Check if $D(\frac{G}{N}\Big| e_1^L)$ breaks within the 1887-91 interval. From Table 6, we conclude that

- $D(\frac{G}{N}\Big| e_1^L)$ breaks in 1888 according to all five test, i.e., within the interval 1887-91.

Check if $D(e_1^L)$ and $D(e_1^L\Big| \frac{G}{N})$ break or are stable in the 1887-91 interval from Table 5 and Table 6. We find

- $D(e_1^L)$ is stable in the 1887-91 interval.
- $D(e_1^L\Big| \frac{G}{N})$ is stable in 1887-91.

From Table A4B, we conclude that $e_1 \perp \frac{T}{Y}$ in 1887-91.

Variable $\frac{T}{N}$. The following statistical breaks in $D(\frac{T}{N})$ recorded in Table 5 coincide with the narrative intervals for the fiscal system listed above:

- $D(\frac{T}{N})$ breaks in 1891 according one test only; so we conclude that there is not break by the two breaks criterion. If we adopt a one break criterion, there is a break in the narrative intervals 1887-91. With the two breaks criterion test B is not informative about the causal relationship between $\frac{T}{N}$ and e_1^L. We the on test criterion, it is and we get the following:

Check if $D(\frac{T}{N}\Big| e_1^L)$ breaks within the 1887-91 interval. From Table 6, we conclude that

- $D(\frac{T}{N}\Big| e_1^L)$ is stable within the interval 1887-91.

Check if $D(e_1^L)$ and $D(e_1^L\Big| \frac{T}{N})$ break or are stable in the 1887-91 interval from Table 5 and Table 6. We find

- $D(e_1^L)$ is stable in the 1887-91 interval.
- $D(e_1^L\Big| \frac{T}{N})$ is stable in 1887-91.

From Table A4B, we conclude that $e_1 \perp \frac{T}{Y}$ in 1887-91 based on the one test criterion.

A6 Results with alternative filtering of the central government data

Table A5: Break Points in the Marginal Distributions of the Franchise and Fiscal Series: Filtered with War Dummies

	Sequential L+1 vs. L	Sequential all	Global vs. none	Global L+1 vs. L	Global information criteria
$D(e_1)$	1815, 1869, 1882	1815, 1833, 1869, 1882	1815, 1869, 1885	1815, 1869, 1885	1815, 1869, 1885
$D(e_2)$	1833, 1869, 1882	1815, 1833, 1853, 1869, 1882	1833, 1869, 1885	1833, 1869, 1885	1833, 1869, 1885
$D(G/Y)$	1853, 1857	1840, 1844, 1853, 1857	1853, 1857	1853, 1857	1853, 1857
$D(T/Y)$	None	None	1834, 1857, 1878, 1899	1834, 1857, 1878, 1899	None
$D(T^d/T^{in})$	1853, 1858, 1865, 1871, 1879	1853, 1858, 1865, 1871, 1879	1865, 1871, 1879, 1899, 1904	1865, 1871, 1879, 1899, 1904	1879
$D(Top)$	None	None	1857, 1892	1857, 1892	None

Note: The table reports the years with break points in the marginal distribution of the franchise and fiscal series. We perform five different Bai and Perron tests in Eviews 9. Details are explained in the text and Appendix A4.

Table A6: Break Points in the Conditional Distributions for the Franchise and Fiscal Series: Filtered with War Dummies

	Sequential L+1 vs. L	Sequential all	Global vs. none	Global L+1 vs. L	Global information criteria
Panel A: Conditional Distributions of Fiscal Variables on Franchise Extension, $D(f\,|\,e)$					
$D(G/Y\,|\,e_1)$	1885, 1896	1885, 1896	1878, 1887	1878, 1887	1878, 1887
$D(T/Y\,|\,e_1)$	None	None	1857, 1870, 1885	1857, 1870, 1885	None
$D(T^d/T^{in}\,\big|\,e_1)$	None	None	1857, 1879, 1898	1857, 1879, 1898	None
$D(Top\,|\,e_1)$	1808	1808	None	None	None
$D(G/Y\,|\,e_2)$	1849, 1869, 1885, 1896	1849, 1869, 1885, 1899	1846, 1859, 1878, 1896	1846, 1859, 1878, 1896	1846, 1859, 1878, 1896
$D(T/Y\,|\,e_2)$	1863, 1896	1863, 1896	1871, 1885	1863, 1896	None
$D(T^d/T^{in}\,\big|\,e_2)$	1885, 1904	1858, 1871, 1885, 1904	1867, 1887	1867, 1887	None
$D(Top\,|\,e_2)$	None	None	None	None	None
Panel B: Conditional Distributions of Franchise Extension on Fiscal Variables, $D(e\,|\,f)$					
$D(e_1\,|\,G/Y)$	1885, 1901	1885, 1901	1883, 1895	1883, 1895	1883, 1895
$D(e_1\,|\,T/Y)$	1857, 1869, 1882, 1903	1869, 1882, 1903	1857, 1870, 1885, 1901	1869, 1885, 1901	1869, 1885
$D(e_1\,|\,T^d/T^{in})$	1884, 1895	1884, 1895	1877, 1893	1884	1884
$D(e_1\,|\,Top)$	1815, 1853, 1903	1853	1846, 1869, 1891	1853	1853
$D(e_2\,|\,G/Y)$	1885, 1901	1885, 1901	1883, 1899	1883, 1899	1883, 1899
$D(e_2\,|\,T/Y)$	1869, 1885, 1894, 1903	1869, 1885, 1894, 1903	1869, 1885, 1895, 1904	1869, 1885, 1895, 1904	1869, 1885, 1903
$D(e_2\,|\,T^d/T^{in})$	None	None	1869, 1885, 1901	1869, 1885, 1901	1869, 1885, 1901
$D(e_2\,|\,Top)$	1869, 1885, 1900	1831, 1869, 1885	1831, 1863, 1885	1831, 1863, 1885	1853, 1885

Note: The table reports the years in which there are (statistical) break points in conditional distribution of the fiscal series conditional on franchise series (panel A) and franchise series conditional on the fiscal series (panel B), respectively. We perform five varieties of Bai and Perron tests in Eviews 9. Details are explained in the text and Appendix A4.

Table A7: Direction of Causality Based on the Patterns of Structural Break Points: Filtered with War Dummies

Franchise	Fiscal	$e \perp f$	$e \leftrightarrow f$	$e \rightarrow f$	$f \rightarrow e$
Test A					
e_1	T^d/T^{in}	1883-87			
e_2	T^d/T^{in}				1866-70, 1883-87
e_1	G/Y				1883-87
e_2	G/Y				1883-87
e_1	T/Y				1866-70, 1883-87
e_2	T/Y	1883-87			
e_1	Top				
e_2	Top	1830-34, 1866-70			
		1883-87			
Test B					
e_1	T^d/T^{in}	1853-57		1894-00	
e_2	T^d/T^{in}	1904-08			
e_1	G/Y				
e_2	G/Y				
e_1	T/Y			1853-57	
e_2	T/Y			1894-00	
e_1	Top				
e_2	Top				

Note: Based on test A and B from Table 2, the Table reports the years in which it is possible to draw a causal conclusion based on the structural break test result reported in Table A5 and A6 and the historical narrative analysis summarized in Table 3. Test A requires that there is an intervention in the franchise series according to the historical narrative in the years with statistical breaks. Test B requires that there is an intervention in the fiscal series according to the historical narrative in the years with statistical breaks. $\perp$ means independent, $\leftrightarrow$ means causality in both directions, and $\rightarrow$ means direction of causality. For test A, the years 1830-34 are the 1st reform act, 1866-70 are the 2nd reform act, 1883-87 are the 3rd reform act. For test B, the years 1853-57 are the beginning of the Gladstone's fiscal constitution, 1894-1900 is the end of the Gladstone's fiscal constitution, and 1904-08 is Asquith's budget. We allow for plus/minus two years around the event identified in the narrative analysis.