The Taxman Tools Up: An Event History Study of the Introduction of the Personal Income Tax in Western Europe, 1815-1941.*

Toke S. Aidt† and Peter S. Jensen‡

December 3, 2007

Abstract

The introduction of income taxation was a landmark in the development of the fiscal state in Western Europe and elsewhere. This paper presents an event history study of the adoption of the income tax in 11 Western European countries between 1815 and 1941. We find evidence that social learning, reductions in tax collection costs and to a lesser extend spending pressures played a significant role for the adoption decision. Surprisingly, we also find evidence that the extension of the franchise *reduced* the likelihood of adoption of the income tax.

Key words: Voting franchise, social learning, tax collection technology, public finance, income taxation.

JEL classification: D7, H1.

1 Introduction

In 2006 more than 30 percent of total tax revenue derived from the personal income tax in the United Kingdom; in the years after it was first introduced in 1842, the yield was less than half of that. During the intervening 164 years, the income tax, not only in

*We would like to thank Stanley Winer, Hamish Low, Martin Paldam, Francisco Gonzalez, Chris Meissner, Wantje Fritschy, participants in the microeconometrics workshop at Cambridge and participants at the conference "Institutions, Public Policy and Economic Outcomes" at Jesus College, Cambridge for helpful discussions and suggestions. We thank Axel Dreher, Christianne Smit, Jukka Pirttila, Carmine Guerriero and Lone Aidt for assistance with dating the income tax in various countries. We would also like to thank the U.K. Economic and Social Research Council for financial support (grant no. RES 000-22-1743) and Ivy Ko and Chris Botham for competent research assistance.

†Corresponding Author: Faculty of Economics, University of Cambridge, Cambridge CB3 9DD, United Kingdom. E-mail: toke.aidt@econ.cam.ac.uk. Phone: +44 1223 335231. Fax: +44 1223 335475.

‡Department of Border Region Studies, University of Southern Denmark, Alsion 2, DK-6400 Sonderborg, Denmark. E-mail: psj@sam.sdu.dk.

the United Kingdom, but in all advanced democratic societies moved from the margins to the center of the fiscal state. The decision, by Western European states, to add the income tax to the arsenal of revenue raising instruments spanned almost a century. The first country to introduce a national-wide personal income tax was the United Kingdom in 1842; the last country to introduce a nation-wide (federal) income tax was Switzerland in 1941. In between these extremes, other European countries gradually introduced taxation of personal incomes.

One of the classical questions in public finance concerns the sources of growth in government? Research on this question has been motivated by the observation that the size of government however measured has increased dramatically in proportion to the economy in virtually all countries which are now considered part of the developed world (see, e.g., Tanzi and Schuknecht, 2000). The stylized pattern followed by most of these countries suggests that growth in government can be divided into four phases: i) moderate growth before World War I, ii) level shift in the interwar period; iii) rapid growth starting in the 1960s with iv) a plateau being reached in the late 1980s. A variety of economic and political explanations has been advanced to explain this pattern, but without reaching firm conclusions (see, e.g., Lindert, 2004a,b). One important idea, pushed for example by Becker and Mulligan (2003), is that efficient taxes are behind big government. If so, this begs a deeper question, namely what causes a society to innovate and adopt efficient taxes? The income tax is arguably one of the most important tax innovations of the past 150 years, only rivalled by value added tax. Gaining a better understanding of how and why it came into being would therefore provide another important stepping stone towards a better understanding of the sources of growth in government.

The purpose of this paper is to study the factors behind the differences in the timing of the income tax across Western European states. We are interested in four potential causes or reasons for delaying or adopting the income tax. The first set of reasons are related to political factors such as the extension of the franchise, the degree of political competition, and the importance of left-wing parties. The second set of reasons relates to the possibility of social learning or spatial diffusion, i.e., the idea that late adopters

observe the results achieved by early adopters and base their adoption decision on this. The third set of reason relates to tax collection technology, while the fourth set relates to wars and other pressures on the public purse.

There exists plenty of circumstantial evidence that these factors mattered for the decision to adopt the income tax. Peters (1991, p. 231), for example, argues that "with the extension of the franchise and the birth of labor movements and socialist parties came demands for greater redistribution through the tax system, and particularly for greater use of income and profit taxation to raise the needed revenues for increasingly active governments." At the same time, there is also much anecdotal evidence that politicians aiming at introducing an income tax at home took notice of what happened abroad. For example, the Austrian income tax introduced in 1849 was explicitly modelled on the English income tax (Sieghart, 1898). In France the repeated, but unsuccessful, attempts to introduce an income tax throughout the later part of the nineteenth century took much of its inspiration from Prussia and England.[1] Likewise, when Britain introduced the income tax in 1842, many Dutch politicians were inspired and a first proposal was made in parliament at that time. The minister responsible was, however, forced to resign and it was not until 1893 that the income tax was finally adopted in the Netherlands (Smit, 2002). Tax collection costs are also believed to matter by many authors. These include Riezman and Slemrod (1987) who view falling tax collection costs as one of the root causes of tax innovations. Moreover, economic history is littered with examples of early attempts to introduce income taxation that subsequently failed because of the lack of fiscal capacity to extract sufficient yields.[2] Finally, the fact that European countries ever since the Napoleonic Wars had resorted to temporary taxation of incomes in times of fiscal need strongly suggests that fiscal pressures might have played a key role in spreading income taxation across Europe during the nineteenth century.

[1] In 1893, for example, three income tax proposals were discussed. The first was simply an extension of the personal property tax (contribution personnelle et mobilière). The second was one based on the Prussian model of a lump-sum income tax (impôt global or impôt sur le revenu). The third was the stoppage-at-source (PAYE) income tax based on the English model, known as the impôt cédulaire or the impôt sur les revenue (Seligman, 1911, Chapter 2; Willis, 1895).

[2] One example is the income tax introduced in 1809-10 in Sweden which was given up in 1812 because of low yields (see Philip, 1965, chapter 9, which contains additional examples).

To quantify the importance of these factors, we offer an event history study of the timing of the adoption of lasting income taxation by the national governments of 11 Western European countries from 1815 to 1941. The idea is to model the (conditional) probability that a country which has not yet adopted the income tax adopts it in a given year as a function of the four factors outlined above. To this end, we estimate a duration model that allows for a time varying hazard rate and unobserved heterogeneity. We find evidence that social learning, reductions in tax collection costs and to a lesser extent spending pressures played a significant role for the adoption decision. Surprisingly, we also find evidence that the extension of the franchise *reduced* the likelihood of adoption of the income tax.

The paper is related to a growing literature on the historical roots of the fiscal state. Boix (2003), Lindert (2004a,b), Aidt et al. (2006), Aidt and Jensen (2007) and others report quantitative evidence on the factors that drove tax and spending levels as well as the composition of the public budget during the nineteenth and early twentieth century. Lindert (1994) studies the factors that triggered the introduction of social spending programs before 1930.[3] Our study differs from these in that it, to the best of our knowledge, is the first to apply event history techniques to study the historical origins of the income tax in a non-US context. Event study techniques are frequently used to study the factors that lead governments to adopt new policies in contemporaneous societies[4] as well as historically[5] and our paper builds on a well-established methodology (e.g., Walker, 1969; Berry and Berry, 1992). It is directly related to a small literature on tax innovations. Building on Hansen (1983) and Berry (1988), Berry and Berry (1992) study the timing of new taxes, related to income and gasoline, among US states during various sub-periods of the twentieth century. They report evidence that states were less likely to adopt an income tax in election years and more likely to adopt it if the state legislature was con-

[3]There is also a related literature on the development of the fiscal state after the Second World War. This includes among others Peltzman (1980), Reizman and Slemrod (1987), Kau and Rubin (1987), Husted and Kenny (1997), Rodrik (1998), Boix (2001), Becker and Mulligan (2003), Persson and Tabellini (2003), Kenny and Winer (2006) and Dusek (2006).

[4]Some examples are Berry and Berry (1990), who study adoption of state lotteries in the US, and Canon and Baum (1981), who study adoption of tort law innovations among US states.

[5]Meissner (2005), for example, studies the diffusion of the gold standard between 1870-1913 in a sample of 24 countries.

trolled by a liberal party.[6] Evidence from outside US states is more scarce, largely because tax innovations are relatively rare events. One exception is Ashworth et al. (2006) who study the adoption of environmental taxes among Flemish municipalities in the 1990s. They also find evidence that adoptions are less likely in election years and more likely when the municipality is controlled by a left-wing party. In addition, they find evidence of social learning in that the likelihood of adoption in a particular municipality increases if its neighbors (both geographically and ideologically) have adopted in the past. We shall comment in more detail on these findings when we discuss our results in Section 5. A virtue of our setting relative to the previous studies is that it spans a period of time during which major political reforms took place. This allows us to investigate, for the first time, the role of institutional reform, such as the extension of the franchise, on tax innovations. One problem, however, is that tax adoption decisions are not made against the backdrop of a common institutional framework, as in the case of states or municipalities; another problem is that we observe relatively few innovations (11) spread over a long period of time (99 years). To deal with the former problem, we pay special attention to unobserved heterogeneity and report results that allow for random effects as well as some instrumental variable estimates. To deal with the latter problem, we report results that correct for the bias that arises when events are rare.

The rest of the paper is organized as follows. In section 2, we provide a brief history of income taxation in Western Europe. In section 3, we introduce and formulate the hypotheses to be tested and discuss how to measure the four sets of potential determinants of the decision to introduce the income tax. In section 4, we set out the empirical strategy. In section 5, we report the main results and consider various strategies to deal with unobserved heterogeneity and rare events. In section 6, we conclude. The data appendix at the end discusses the construction of the data set and its sources and provides a very brief history of income taxation in the 11 countries in our sample.

[6] Democrats are assumed "liberal" except in the Southern states where they are assumed "conservative".

2 Income Taxation in Western Europe

At the end of the Napoleonic Wars in 1815, the fiscal states of Western Europe still relied heavily on indirect taxes, such as customs and excise duties, for revenues, supplemented with land and inheritance taxes. But things were starting to change with significant ramifications for the development of the fiscal state. Forced by the need to raise revenues to finance the war against the French, an income tax was introduced in Britain in 1799.[7] It had most of the attributes of the modern income tax: citizens had to file a yearly tax return stating gross income from all sources which were then taxed at 10 per cent with an allowance for deduction of certain expenses, children and payments to life insurance. The tax was, however, repealed in 1802, but reintroduced in the following year when war broke out again. Other countries, including the Nordic countries, France, and some of the German states, also experimented with income taxation during the Napoleonic Wars, but only temporarily. The first country in Europe to make the income tax, understood as a tax on earned income,[8] a lasting feature of the tax system was Britain in 1842.[9] It was soon after followed by the Austria Empire in 1849[10] and eventually by 1941, when Switzerland introduced a federal income tax, all the countries in our sample (listed in Table 1) had a permanent income tax levied by the central or federal government. The precise years of adoption are reported in Table 1 along with information about temporary income taxes and income taxes levied by local government or by states within a federation. We also list the (approximate) year in which the revenue yield of the income tax reached five percent of total revenue. We interpret this as a measure of how important a contribution the income tax made to the fiscal capacity of the state. For most countries, the five percent threshold was reached shortly after the adoption, but there are significant exceptions. In Austria, where an income tax modelled in part on the English income tax (with different schedules)

[7] See Daunton (2001, chapter 1 and 4) for a detailed analysis of the income tax in the United Kingdom.

[8] Various "proxy" taxes, such as wealth taxes, window taxes and property taxes, partly aimed at taxes something correlated with income, were widely used. These taxes are not part of our definition of income taxation.

[9] Although the tax was initially introduced as a temporary measure to replace revenues lost by the repeal of the Corn Laws and had to be approved by annual votes in Parliament, it persisted all attempts to abolish it (Peters, 1991, chapter 7).

[10] See Sieghart (1898).

was introduced in 1849, the yields stayed below five percent until 1905.

<Table 1: The timing of the income tax in Western Europe.>

As is clear from the Table, most countries experimented with temporary income taxes during periods of war or financial crisis and some federal countries, notably Germany and Switzerland, had income taxation at the state level long before it was added to the toolkit of central government (Seligman, 1911). Likewise, some form of income taxation at the municipality level was common in the Nordic countries long before nation-wide income taxes were introduced (Philip, 1965).[11] It is therefore important to stress that we mainly focus on the timing of the adoption of a *permanent* income tax levied by *central* government, although we for completeness also report some results for specifications in which we model both permanent and temporary income tax adoptions.

3 The Main Hypotheses

It is clear from Table 1 that the income tax was gradually added to the fiscal toolbox in the years following the end of the Napoleonic Wars. We want to understand what explains the timing of the adoption of the income tax by studying its diffusion across Europe. To fix ideas, it is useful to develop a simple theoretical framework. Consider a risk neutral country, indexed by i, which has not yet at time t adopted the income tax. Let $B_{it}(.)$ be the expected present value of the benefit of adoption. The benefits include *diversification gains* that arise because the same revenues when raised with more tax instruments can be collected at lower economic and political cost and an *expansion gain* that arises because more revenues can be raised at the same cost (see, e.g., Ashworth et al., 2005). The present value of the cost of adoption as perceived at time t, denoted $C_{it}(.)$, includes the fixed cost of adoption as well as the variable costs of tax collection once the tax has been introduced. We can think of these costs and benefits as being partly

[11] This has caused a certain amount of confusion in the literature about the dating of the income tax as authors are not always explicit neither about whether the income tax is lasting or temporary nor about the level at which it is levied. We have checked the dates recorded in Table 1 against multiple sources including national encyclopaedia. The Data Appendix provides a concise history of income taxation in each country and lists the sources used to date the adoption.

political (e.g., lost or gained popularity) and partly economic (e.g., the deadweight and administrative costs or the benefits of being able to finance new spending programs). Under democratic rule, for example, political parties would factor the electoral consequences of new taxes into their adoption decision, while autocratic rulers might be more concerned with the consequences for regime stability.[12] As stressed by Winer and Hettich (1991), the economic costs and benefits of adopting new taxes depends on the *entire* fiscal system, including the availability of alternative tax bases, how hard these bases are pushed to raise revenues, what the revenue needs are and so on.

Each period there is a random shock to the cost of adoption ε_{it}. It is independently distributed across time and space with $E(\varepsilon_{it}) = 0$ and distribution function F. The shock is observed before the adoption decision is made. Given that, the country adopts at time t if

$$B_{it}(P_{it}, S_{it}, L_{it}) - C_{it}(S_{it}, T_{it}, L_{it}) \geq \varepsilon_{it} \tag{1}$$

and the probability of adoption at time t given that adoption has not yet happened is then given by

$$p_A(P_{it}, L_{it}, S_{it}, T_{it}) = F(\varepsilon_{it} \leq B_{it}(.) - C_{it}(.)). \tag{2}$$

The (conditional) probability of adoption, p_A, depends on four main factors.

Political institutions and politics [P]. The decision to adopt or not is filtered through the political institutions in place in the country at the time. The costs and benefits of adoption are, therefore, affected not only by the constituency of government and the party composition of parliament, but also by how easy it is to contest political power by different factions in society. Given that the income tax is a potentially powerful tool for redistribution and that a broader franchise lowers the income of the decisive voters, the benefits of adoption should be higher under universal than under restricted suffrage. Moreover, as the constituency of government expands (as a consequence of suffrage reforms), it becomes more attractive for governments to provide public goods. This creates a need for

[12] Hettich and Winer (1988, 1999) and Ashworth et al. (2006) develop a simple framework for analysing tax innovations within the context of a competitive democratic system which is useful for conceptualising some of these effects.

extra revenues. We, therefore, expect that the extension of the voting franchise (to poorer citizens) increases the probability that the income tax is adopted:[13]

Hypothesis 1 *The extension of the franchise makes adoption of the income tax more likely.*

Arguably, the newly enfranchised citizens need representation in parliament to make their demands heard. Left-wing parties, in particular, are likely to have played an important role in the propagation of income taxation in Europe as they have been found to do in other contexts (Berry and Berry, 1992; Ashworth, 2006).[14] For left-wing parties the benefit of new tax instruments is relatively high because they represent constituencies that favor more spending. The political cost of adoption might also be relatively low because many of their voters would not, at least initially, pay income tax. We therefore conjecture that:

Hypothesis 2 *The income tax is more likely to be adopted under left-wing government.*

In societies with regular democratic elections, there is ample evidence that politicians time fiscal decisions within the election cycle to maximize the chance of reelection.[15] Insofar as new taxes are unpopular with at least some segments of the voter population, politicians looking to minimize the political cost of a tax innovation are least likely to adopt it immediately before an election and most likely to do it immediately after an election. As pointed out by Berry and Berry (1992, p. 719) "such timing would give the public the maximum amount of time to forget the government's unpopular action before the next election". We, therefore, expect

[13] Theoretical work by Meltzer and Richard (1981) suggests a positive link between franchise reform and redistribution. The survey by Tridimas and Winer (2005) discusses of franchise reform within the framework of the probabilistic voting model and reach a similar conclusion. The CGE simulations of Winer and Rutherford (1993), however, cast some doubt on how big a fiscal expansion franchise extension causes in practice.

[14] There is also substantial evidence that left-wing parties played an important role in building welfare states in Western democracies after the Second World War (see, e.g., Hicks and Swank, 1992; or Blais et al., 1993).

[15] See, for example, the survey by Paldam (1997).

Hypothesis 3 *The income tax is less likely to be adopted in election years and more likely to be adopted the longer there is to the next election.*

Another important political factor that might have influenced the decision to adopt the income tax is the degree of political competition. By political competition, we mean the extent to which political power can be contested by organized factions and the executive can be held accountable by parliament and/or voters. Enhanced political competition can affect the adoption probability through a number of competing channels. Firstly, it may increase the probability of adoption because it forces political parties to innovate and to seek new and more effective ways of raising revenues.[16] Secondly, by allowing power to be more freely contested, enhanced political competition may lead to government fragmentation. This can lead to grid-lock and be an obstacle for tax innovations. Thirdly, compared to an autocracy or a restricted democracy, a fully competitive political system may increase political uncertainty (and government turnover). This may shorten the time horizon of politicians. Insofar as a large fraction of the cost of adoption is paid up front and the benefits arrive later, this could cause a myopic bias against adopting new taxes. Overall, a priori political competition has an ambiguous effect.

Social learning and geographical diffusion [L] A country with no experience with the income tax might look to neighboring countries, which has already adopted the tax, to learn about its effectiveness in terms of raising revenues and about the administrative procedures and costs needed to implement it (Walker, 1969; Berry and Berry, 1992). Looking at successful adoptions in neighboring countries can reduce some of the uncertainties surrounding the costs and the benefits of adoption. Insofar as successful adoption in neighboring countries demonstrates the revenue potential of the income tax, the benefit of adoption should increase in country i. Likewise, if the adoption elsewhere demonstrates cost effectiveness or introduces a particular innovation in tax collection technology country i should revise its cost estimate downwards. Taken together, these information externalities suggest that

[16] Fundamentally, this hypothesis derives from the belief that political competition is efficiency enhancing. For theoretical work supporting this idea, see, e.g., Wittman (1989).

Hypothesis 4 *The income tax is more likely to be adopted in a particular country if other countries have already introduced the tax.*

It seems clear that a successful (permanent) adoption of the income tax in one country may send a strong (positive) signal to neighboring countries about the costs and benefits of income taxation. However, countries may also learn from temporary, and therefore in some sense unsuccessful, adoptions in other countries. The information content in observing a temporary adoption is, however, likely to be very different. One may, in fact, conjecture that temporary adoptions may have a negative effect on the likelihood that neighboring countries adopt: the fact that the tax was given up again suggests that the benefit did not exceed the cost after all.

Tax collection costs [T]. Compared to trade taxes and taxes on property, the income tax is complex to administer and difficult to collect. While, for example, a tariff can be collected at ports by a small number of educated civil servants, income taxes require the participation of a large number of educated people, who can fill in complex tax returns, and require a large bureaucracy to administer (Riezman and Slemrod, 1987; Kenny and Winer, 2006). The cost of taxing incomes as they are earned is reduced when taxpayers acquire the skills needed to handle tax returns and the tax administration develops the capacity to audit and collect complex, broad-based taxes. The development of formal markets and adoption of modern accounting systems are also important factors. We expect that:

Hypothesis 5 *Improvements in the tax collection technology increase the probability that the income tax is adopted.*

Spending pressures [S]. The doctrine of a balanced budget was rigorously applied throughout the nineteenth century and was effectively *the* framework for public finances in Western European societies till after the Second World War when Keynesian ideas of deliberately unbalanced budgets became generally accepted (Webber and Wildavsky, 1986, chapters 6 and 8). In fact, with the possible exception of Italy and Germany, the norm of balance was so widely accepted that when governments began to spend proportionately

larger amounts on social programs and arms, they had to find new taxes to finance these expenses to maintain balance. Mounting spending pressures, e.g., triggered by wars or other types of fiscal crises decrease the risk of adopting new taxes. We can summarize what Berry and Berry (1992) call the fiscal health hypothesis as follows:

Hypothesis 6 *The income tax is more likely to be adopted in times of fiscal crisis.*

Economic development and the associated changes in demographics and modernization of society more broadly (Wagner, 1883) also put pressures on the public purse while at the same time enlarging the potential tax base from which income taxes can be collected. This makes tax innovations more acceptable to public decision makers and we expect that

Hypothesis 7 *The income tax is more likely to be adopted in more developed societies.*

To estimate p_A, we obviously need suitable variables to represent P, L, T and S. Table 2 lists the variables that we have chosen for the purpose. The choice is somewhat constrained by the fact that our analysis starts in 1815. This makes data availability a problematic issue.

<Table 2: Overview of explanatory variables and mnemonics>

We use five variables to capture political factors. Our main political variable is *suffrage*. It is a direct measure of the extension of the franchise and records the fraction of the (male) population that could vote for elections to the lower chamber of parliament and is coded zero if no elections took place (Flora et al., 1983).[17] We also make use of four other political variables. The first of these is a variable called *left-wing parties*. It is an attempt to capture the importance of left-wing parties in the lower chamber of parliament. It is measured as the seat share won by left-wing parties (Flora et al., 1983; Caramani, 2000; Cook and Paxton, 1998). We expect that both of these variables have a positive effect on the adoption probability. To capture the effect of the election cycle, we use the dummy

[17] Women's suffrage was, except in Germany and Belgium, introduced after the adoption of the income tax (see Aidt and Dallal, 2007). Thus, we do not make any attempt to model this aspect of enfranchisement.

variable *election year*. It is equal to 1 in election years and 0 otherwise. We also use the variable *years to next election*. It counts the number of years to the next election. The last variable makes most sense if there are regular elections. We expect a negative effect of *election year* and a positive effect of *years to next election*. We use the variable *political competition* to capture institutional reforms that induced more political competition for control of government and/or imposed more checks on the executive. The variable is equal to the polity IV index constructed by Marshall and Jaggers (2000).[18] The index is measured on a scale from -10 (little political competition) to 10 (lot of political competition).[19] As discussed above, we do not have a clear prediction regarding this variable.

We use two alternative measures of "positive" social learning. The idea behind both is that a country that has not yet adopted the income tax is more likely to adopt it if countries "close" to it has adopted in the past. The two measures differ in their definition of "closeness". *Geographical closeness* uses the inverse of the distance between the capital in the country under consideration and that of the other countries which have adopted the income tax in the past to define closeness.[20] *Trade connection* uses the extent of bilateral trade relative to GDP to define closeness. The later may be theoretically more satisfactory insofar as information about social innovations is transmitted mostly through trade interactions and is only indirectly related to physical distance. However, *trade connection* cannot, due to data limitations, be traced further back in time than 1870. Thus, in the statistical analysis, we use *geographical closeness* as the principal measure of (positive) social learning. We expect a positive effect of both of these variables. To capture that temporary adoption may also induce social learning, but most likely by giving a negative signal to surrounding countries, we have, using physical distance between capitals, coded

[18] The polity IV index summarizes different indicators of political authority patterns to measure three key aspects of a country's political system. The three aspects are: i) competitiveness and openness in the process of executive recruitment; ii) constraints on the chief executive and iii) competitiveness and regulation of political participation. A weighted sum of these components is used to construct two summary variables, measuring democracy on a scale from 0 to 10 and autocracy from -10 to 0. The polity IV index is the sum of these two sub-indexes.

[19] See Aidt and Eterovic (2007) for a detailed discussion of the link between the polity IV index and political competition.

[20] Walker (1969), Berry and Berry (1992) and Ashworth et al. (2006) use neighboring states to define geographical closeness. The geography of Western Europe makes this an uncompelling choice for our purposes and we prefer to use the capital-to-capital distance instead.

the variable *learning from temporary adoptions*.

As argued by Riezman and Slemrod (1987), Kenny and Winer (2006), Aidt and Jensen (2007) and others, the relative administrative cost of levying income taxes falls as literacy becomes more widespread. Literacy and basic numerical skills are *necessary* for filing tax returns. The spread of these skills in the population, therefore, reduces the cost of collecting taxes that require individuals and businesses to file a return and to keep detailed records of their transactions.[21] In short, a minimum level of literacy and numeracy is necessary for successful income taxation. As Riezman and Slemrod (1987, p. 546), who first proposed the proxy, point out: it "seems reasonable [that the collection cost of tariffs relative to income and sale taxes depends positively on literacy levels] because operating a tariff system requires a small number of educated civil servants, while income or sales taxation, to function effectively, requires the participation of a large number of educated people".[22] In addition to this, it is necessary to keep track of where individuals live and work. Accordingly, as more accurate census information becomes available the administrative cost of income taxation should fall. Finally, experience with either local income taxation or with temporary nation-wide income taxes, e.g., at times of war helps build institutional capacities that lower the administrative cost of a permanent adoption. We have constructed a *cost index* as described in Table 2 to capture these forces and expect that an increase in the index (representing a fall in relative tax collection cost) to increase the probability of adoption. It can be argued that urbanization is another good proxy for administrative costs (see, e.g., Kenny and Winer, 2006). The idea is that it is cheaper to collect income taxes in an urban environment. We, therefore, expect that *urbanization* increases the probability of adoption. While the constituents of *cost index*

[21] These tax returns were complex documents and it was not a simple matter to fill them in. An example of a tax return from 1887-88 from the United Kingdom can be viewed on http://www.econ.cam.ac.uk/faculty/aidt/papers /papers2.htm.

[22] There are two other reasons why literacy is a good proxy for the relative administrative cost of income taxation. Firstly, education is correlated with earnings, and average education attainment provides a measure of the earnings potential of the population that is independent of the business cycle. Income taxation is ineffective and costly if the vast majority of the population live on subsistence wages. As education levels improve, more people earn incomes above this level. This, in turn, reduces the cost of collecting taxes levied on income. Secondly, it is easier (and cheaper) to recruit a well-qualified bureaucracy in societies with a larger pool of educated individuals. This, in turn, allows the government to build institutional capacities.

can be tracked back to 1815, data on the degree of urbanization is not available for the entire period and, for this reason, *urbanization* is not one of the constituents of the cost index and is a somewhat problematic measure of tax collection costs in our context. We expect a positive effect of both of these variables.

Spending pressures are likely to increase the probability of adoption of the income tax and can be measured in four alternative ways. Firstly, wars create an acute need for public finance and many countries introduced temporary income taxes during times of war. We have coded a dummy variable called *war* to capture this possibility. Second, economic development (measured by *GDP per capita*) and population pressures (measured by the size of the population) may increase the demand for public goods and other government services (Wagner, 1883). A larger population also means that there are more shoulders to bear any fixed costs of adoption. Third, ageing of the population (measured as the percentage of the population above 65 years of age) is another likely contributor to spending pressures as demands for social security and pensions increase (Lindert, 1994). Finally, we use a measure of the budget deficit, which within the context of a balanced budget regime becomes a direct measure of spending pressures. While the variables *war*, *GDP per capita* and *population* cover the entire period from 1815, the variables *age structure* and *deficit* cannot be traced further back than 1860 for most countries. We expect that all these variables have a positive effect on the probability of adoption.

4 Empirical Specification

In our baseline model, we code the dependent variable y_{it} as one if country i adopts the income tax in year t and zero in the years before that. A country drops out of the sample when it has adopted the income tax.[23] We deal with the problem of left censoring, i.e., the problem that we do not know precisely when the spell without income taxation starts for each country, in the following way. We have chosen the end of the Napoleonic Wars (1815) as the point in time when countries that were independent at the time became at "risk" of

[23]Since we focus on the arrival of lasting income taxation there is, by definition, no issue of "reentry" of countries that give up the income tax at some future point in time.

adopting the income tax permanently. The rationale for this is that many countries (e.g., the United Kingdom, the Nordic countries and France) partly financed the Napoleonic Wars by taxing incomes. This establishes beyond any doubt that the technology to do so existed at that point in time. For countries that were not independent in 1815, we assume that they become at "risk" of adopting the income tax as soon as they did become independent.[24] All in all, this gives us a data set of up to 642 country-year observations.

Our data are grouped duration data and it is natural to use a duration model to estimate the relationship between the potential explanatory factors introduced above and the time conditional probability of adopting the income tax (the hazard rate). Since it is unlikely that the hazard rate is independent of the tax history of the countries, we allow and test for duration dependence. We follow Beck et al. (1998) and estimate the following discrete logistic model:[25]

$$P\left(y_{it}=1|\, x_{it}, y_{it-1}=0\right)=\frac{1}{1+e^{-(x_{it}\beta+H(t-t_i))}}$$

where x_{it} is a vector of explanatory factors (chosen from among those listed in Table 2) and β is the vector of parameters of interest. $H(t-t_i)$ is a smooth function of the number of years a country has been at "risk" of adopting the income tax[26] and allow us to model duration dependence in a flexible way and to test the assumption of a constant baseline hazard rate. We estimate $H\left(.\right)$ using natural cubic splines and use the estimated spline coefficients along with the cumulation of years spend without income taxation to trace out the path of duration dependence.[27]

In addition to the baseline model, we estimate two alternative models. Firstly, to increase the number of events, we have recoded the dependent variable, y_{it}, to take into account temporary as well as permanent adoptions.[28] Secondly, it can be argued that the

[24] This means that Belgium and the Netherlands enter the data set in 1830; that Switzerland enters in 1848 (when a federal stucture was established); that Italy enters in 1861; and that Germany enters in 1871.

[25] Beck et al. (2004, Appendix) show that this is the discrete hazard model corresponding to the well-known continuous time proportional hazard model (Cox, 1975).

[26] t_i representes the year in which country i enters the risk set (i.e., either 1815 or the year of independent).

[27] We have determined the number of knots by a sequence of F-tests and have settled on a specification with two knots.

[28] For a country that had, say, one temporary adoption before the income tax was finally adopted for

income tax is not "really" introduced until it contributes with more than a token to total tax revenues. To allow for this possibility, we have recoded y_{it} such that the income tax is "adopted" in the year when the revenue yield exceeds five percent of total tax revenues. We have experimented with different cut-offs, but the results are fairly similar, at least within the range from five to ten percent of total revenues.

5 The Results

We have divided the discussion of the results into four sub-sections. Firstly, we present the baseline results. Secondly, we present results that take into account unobserved heterogeneity. Thirdly, we present results from specifications that model both temporary and permanent income tax adoptions and results estimated with a statistical technique that takes into account that tax adoptions are rare events. Fourthly, we report some results that use the yield-based definition to date the introduction of the income tax.

5.1 The Baseline Results

The main results are reported in Tables 3 and 4. The baseline model (1) is relatively parsimonious and includes only the key variables from each of the four groups of variables (*suffrage*, *cost index*, *geographical closeness*, *war*, *GDP per capita*, and *population*). The other models are permutations of the baseline where we have added (or deleted) variables. The models shown in Table 3 add variables for which we have data for the entire sample period, while the models shown in Table 4 add variables that reduce the sample size considerably and it is, in these cases, not possible to tell if the baseline results change because the sample is shortened or because of additional covariate(s). A comparison between model (1) and model (2) indicates strong duration dependence in the hazard rate. The Likelihood ratio test reported at the bottom of Table 3 rejects the hypothesis of a constant hazard rate at the one percent level.

Our measure of the extension of the franchise, *suffrage,* is significant but with a negative

good, the dependent variable is coded zero up to the point of the temporary adoption and then one in the year of adoption. The country drops out of the data set during the period of the temporary tax and reenters when it is abolished.

coefficient in the baseline model. The result is very robust across the different specifications shown in the two tables and suggests that the extension of the franchise *reduced* the likelihood of adopting the income tax. Based on the point estimate from the baseline model, the odds of adopting the income tax is about 100 percent higher in a society without elections than in one with universal (male) suffrage.[29] This result is surprising and a clear rejection of Hypothesis 1. One possible explanation is that the estimate in the baseline model picks up the effect of elections rather than the effect of the franchise as such. This would be consistent with the idea that tax innovations are less likely to be adopted in election years (Hypothesis 3). However, model (3), which adds *years to next election* and *election year* to the baseline specification, allows us to rule this out: neither of these additional variables are significant and the coefficient on *suffrage* continues to be negative and highly significant. Not only does this fail to provide an explanation for the rejection of Hypothesis 1, it also provides evidence that the election cycle did not exercise a major influence on the adoption of the income tax in Western Europe. This stands in sharp contrast to the evidence reported by Berry and Berry (1992) for U.S. states and by Ashworth et al. (2006) for Flemish municipalities. Both of these studies found strong evidence that tax innovations were not adopted in election years. This may indicate that the election cycle played a less important role in the early days of democracy in Europe than it does today. In addition and also in contrast to Berry and Berry (1992) and Ashworth et al. (2006), as model (4) in Table 3 shows, we find virtually no evidence that left-wing parties were leading the drive towards income taxation (Hypothesis 2). Not only is the effect statistically insignificant, the point estimate is *negative* rather than positive. Finally, model (4) also shows that the degree of political competition does not seem to matter much for the likelihood of adopting the income tax.

We find some evidence that high tax collection costs reduce the probability of adoption (Hypothesis 5). In the baseline model, the *cost index* has the predicted positive sign and is significant at the ten percent level. Based on this estimate, a one point reduction

[29] The odds ratio is calculated as $exp(-0.054 * 100) = 0.044$ where the "100" represents the move from no elections to universal male suffrage. The percentage increase in the odds that the income tax will be adopted following a full extension of the franchise is $(exp(-0.054 * 100) - 1) * 100 = -99.6$.

in the *cost index* increases the odds of adoption by more than 300 percent. In models (5) to (7) in Table 3, we investigate the contribution of the sub-components of the cost index individually.[30] The different sub-components have conflicting effects. *Temporary income tax*, which captures mainly wartime income taxes, seems to reduce the probability of a permanent adoption, while *local income tax* has a significant positive and significant impact. Education attainment is not close to being significant. This suggests that experimentation with local income taxes, either at the state or at the municipal level, were the main driver of the technological innovations that eventually reduced the cost of rolling out income taxation at a national level. The specification with the alternative measure of tax collection cost, *urbanization*, brings further support to Hypothesis 5 as the coefficient on *urbanization* is positive and highly significant in Table 4, model (9).

We also find evidence of social learning (Hypothesis 4). In the baseline model, *geographical closeness* has the predicted positive sign and is significant at the five percent level. The effect is robust across specifications using the full data set, but it becomes insignificant in the specifications with a reduced sample reported in Table 4. The variable *geographical closeness* only takes into account the learning effects that might arise from permanent adoptions in other countries. It is, however, possible that neighboring countries would also learn something from observing temporary adoptions. To allow for this possibility, we have in model (8) included *learning from temporary adoptions*. We note that this variable is insignificant but with a negative sign. If anything this suggests that the learning experience from permanent and temporary adoptions was very different. A specification with the alternative measure of social learning, *trade connection*, is reported in Table 4, model (10). We notice that despite the fact that the sample is reduced to 263 observations, the estimate is significant at the ten percent level. All in all, we interpret the evidence as being mildly supportive of Hypothesis 4.

The evidence on the role of spending pressures is more mixed. While the dummy variable *war* is not statistically significant in any of the specifications reported in Tables

[30]It is not possible to enter the forth component, *census*, in the estimations on its own when using logit. We have estimated the relation using Maximum Penalized likelihood (see note to Table 4) and find a positive, but insignificant impact of *census*.

3 and 4, it does have the expected positive sign. However, the direct measure of spending pressures, *deficit*, brings some support to Hypothesis 6: the estimate from model (11) in Table 4 suggests that a one percentage point increase in the budget deficit increases the odds of adopting the income tax by about 7 percent. Moreover, *population* has a positive and significant impact on the adoption probability in some specifications (e.g., model (1)).

Finally, modernization theory (Hypothesis 7) receives no support. *GDP per capita* seems to have a negative effect on the adoption probability, albeit an effect that is mostly insignificant. *Age structure* does have a significant impact (model (12) in Table 4), but one that is negative. In addition to rejecting modernization theory, this result also suggests that the "grey power" effect that aging populations were one of the key drivers behind the rise of social spending in the West (Lindert, 1994) was not making it more likely that income taxation was adopted to finance this expansion.

<Table 3: Logit estimates of the probability of adopting the income tax permanently, 1815-1941>

<Table 4: Logit (or MPL) estimates of the probability of adopting the income tax permanently, additional control variables>

5.2 Unobserved Heterogeneity

One concern about the (surprising) finding that the extension of the franchise made the introduction of income taxation less likely is that it may be driven by omitted variables bias. However, if fiscal and political reforms were driven by common unobserved preference shifts, we would expect that the logit estimate is upwards, not downwards, biased. Nevertheless, we make an attempt to deal with this concern by taking an instrumental variables approach. We use two instrumental variables for *suffrage*. The first instrument is a dummy variable coded one if a revolution or major revolt took place in one of the 11 countries in the sample or elsewhere in Europe in a particular year (*revolution*). This variable is motivated by the theoretical model of Acemoglu and Robinson (2000) and is a

crude attempt to quantify the (perceived) threat of revolution.[31] Second, the extension of the franchise depends on its own past value. This captures the strong path dependency in the evolution of franchise institutions in Europe. While the first instrument is clearly rooted in theory, one might question whether the lagged value of the franchise is uncorrelated with any unobserved factor that affects fiscal reform.[32] To implement this strategy, we estimate a linear probability model with 2SLS.[33] The results are reported in Table 5.

<Table 5: 2SLS Linear Probability model and Random Effects Logit Model.>

We might begin by noting that the J-test does not reject the validity of the instruments. Thus, if we believe on a priori grounds that *revolution* is a valid instrument, then we can interpret this test result as a validation of the lagged value of *suffrage* as an instrument. However, although the two instruments are highly significant in the first stage regression reported in model (13), this is largely driven by the lagged value of *suffrage*, and not by *revolution*. More importantly, we see that the impact of suffrage reforms on the probability of adoption of the income tax continues to be negative and highly significant (model (14)). Accordingly, we conclude that the rejection of Hypothesis 1 is unlikely to be a by-product of omitted variables.

Another concern that one may have about the results reported in Tables 3 and 4 is that the baseline hazard rate is not allowed to have a country-specific component. To accommodate this concern, we report in model (15) of Table 5, the results of estimating a random effects logit model. This specification allows the baseline hazard rate to be affected by idiosyncratic country-specific shocks. A comparison between the estimates from this model and those from the baseline model reveals no substantial differences and thus we believe that the conclusions drawn from the baseline model are robust to allowing for a heterogenous baseline hazard rate. Further, tests of the null hypothesis of homogeneity do

[31] For alternative theories of franchise extension see Congleton (2004) or Justman and Gradstein (1999).

[32] We use lags of three years. Under the assumption that the lagged value of *suffrage* is uncorrelated with the unobserved component of the income tax equation, this procedure produces consistent estimates in panels with a long time dimension. This, however, requires that the relevant error terms are uncorrelated over time.

[33] Angrist and Krueger (2001) and Wooldridge (2002) recommend the use of 2SLS even in cases when the dependent variable is binary.

not reject the null.

5.3 Temporary Income taxation and Rare Events Logit

Tax innovations are rare events and the innovations that led to the adoption of the income tax in Europe were spread out over almost 100 years. This is what makes it challenging to estimate the causes of these adoptions. To rise to this challenge, we have explored two alternative routes. Firstly, we have increased the number of events by including temporary as well as permanent adoptions in the analysis. While the advantage of this is more events, the downside is that it is unclear if the decisions that led to a temporary adoption were of the same nature as those that led to a permanent adoption. For example, as is clear from Table 1, many temporary adoptions had to do with involvement in wars. On the other hand, when the income tax was introduced in United Kingdom in 1842, it was intended as a temporary measure and was subject to an annual vote, yet it lingered on and effectively became a permanent feature of the tax system. The results of the estimations with all income tax adoptions, temporary or permanent, are reported in Table 6 (which has a layout similar to that of Table 3). The results are very similar to those reported in Table 3 with two exceptions. First, there is less evidence of duration dependence and, in fact, we cannot in some of the specifications reject that the hazard rate is constant. Secondly, spending pressures as measured by the variable *war* are now highly significant. This is not surprising as many of the temporary income tax adoption in the sample can be attributed directly to war.

Secondly, it is well-known that the maximum likelihood estimate of the parameters of a logit model are biased in small samples (less than 200 observations). While our sample has more than 600 observations, we are concerned that whatever bias there is, it is amplified by the fact that tax adoptions are rare events.[34] We therefore need to take the possibility of a systematic downwards bias seriously. King and Zeng (2001) has developed an estimator that corrects for this bias. It also improves the efficiency of the estimates. We have re-run our estimations using this estimator and report a subset of the results in

[34] Tax adoptions constitute less than two percent of our cases.

Table 7. We show four specifications, two with permanent adoptions only and two with temporary or permanent adoptions. We continue to find that suffrage reforms reduced and tax innovations (measured either by *cost index* or by *local income taxes*) increased the adoption probability. However, in these estimations the evidence of social learning is weaker than before. The coefficient on *geographical closeness* is mostly insignificant albeit positive. In short, although the estimated coefficient on *suffrage* is numerically smaller that in the baseline model, the results reported in Table 7 allow us to rule out that the puzzling effect of suffrage reforms is caused by "rare events" bias.

<Table 6: Logit estimates of the probability of adopting the income tax temporary or permanently, 1815-1941>

<Table 7: Rare events logit estimates of the probability of adopting the income tax, 1815-1941>

5.4 A Five Percent Threshold

Above we have addressed the question of what determines the *introduction* of the income, irrespective of whether the tax became a major revenue source or not. As indicated by Table 1, the income tax, in fact, did reach reasonably high yields soon after its introduction in most countries, but with Austria being a notable exception. Motivated by this, we want to ask a slightly different question in this section: what determined whether or not a country reached a given yield threshold, conditional on having adopted the income tax? We have experimented with different thresholds. We report the results from estimations using a five percent threshold in Table 8.[35] We note that suffrage reforms not only reduced the probability that the income tax was introduced, but also reduced the probability that once it was there, the yields hit the five percent threshold. Tax collection technology and social learning, on the other hand, did not play an major role. The latter result is perhaps not all that surprising as one would expect that the signal emitted by a country that

[35] The five percent threshold assumes that a country has not adopted the income until it yields at least five percent of total tax revenues (see Table 1). The dependent variable is coded accordingly.

adopts a new tax is much stronger than the signal emitted when a certain yield threshold is reached.

<Table 8: Logit estimates of the probability of reaching a yield of at least five percent, 1815-1942>

6 Discussion and Conclusion

We have estimated the conditional probability that a country adopts the income tax for a sample of 11 Western European countries. We find evidence that social learning, reductions in tax collection costs and to a lesser extent spending pressures played a significant role for the adoption decision. Surprisingly, we also find evidence that the extension of the franchise *reduced* the likelihood of adoption of the income tax. Our analysis suggests that this puzzling effect is unlikely to be caused by confusion between election year effects and the effect of franchise reforms or be due to omitted variables or to rare events bias.

We can offer one tentative explanation that might help resolve the puzzle. The elites that extended the franchise in Western Europe accepted that the poor could vote, but at the same time they introduced new or enhanced old checks and balances that prevented or delayed large scale Robin Hood type redistribution.[36] For example, it was common practice to have two legislative chambers and legislation either needed approval of both or could be blocked by the upper chamber. By maintaining control over the upper chamber, while widening the franchise for the lower, the elites effectively put constitutional constraints on the extent of redistribution. This may have delayed the adoption of the income tax for two reasons. Firstly, it was less attractive as a tool of redistribution. Secondly, its value as a revenue raiser was reduced by the fact that large scale expansion of spending programs favored by the newly enfranchised voters were blocked. This interpretation is consistent with other available evidence. Aidt et al. (2006) report that the franchise extension in Western Europe did not lead to a large and immediate expansion of total government spending, but shifted the portfolio of spending. This observation supports

[36] See Breyer and Ursprung (1998) for a theoretical model along these lines.

the view that income taxation as a revenue raiser might have been limited at the time of franchise extension.

7 Data Appendix

Construction of the event history data sets We use two data sets in the analysis. The *complete* data set contains information on *suffrage*, *left-wing parties*, *political competition*, *election year*, *years to next election*, *cost index* (and its constituents), *geographical closeness*, *GDP per capita*, *population*, *learning from temporary adoptions* and *war* for all 11 countries from 1815 (or the time of independence) to the time when the income tax was adopted. For *suffrage* and *left-wing parties* data are only recorded in election years. In off-election years, we have no data. For *suffrage,* we have made the assumption that the variable stayed constant between elections, something which is obviously true for left-wing parties' share of seats. Both variables are coded zero for periods without democracy and elections of any sort. The *incomplete* data set (or sets) contains four additional variables (*urbanization*, *trade connection*, *deficit* and *age structure*) which cannot be tracked back to 1815 (or time of independence). In addition to this, for some of the variables in the incomplete data set, the original data contains gaps for the time span they do cover. In these cases, we have interpolated the missing observations linearly. Linear interpolation seems a reasonable solution, although more sophisticated methods are available. Little (1992) suggests, for example, to use predicted values from least squares regressions in which the other explanatory variables available in the data set are used as regressors to fill in the gaps. For most variables, we, however, have too few observations to make this a sensible procedure and we did not pursue this alternative.

Data sources The data on the extension of the franchise are from Flora et al. (1983) and Cook and Paxton (1998). The data on the share of seats held by left-wing parties is constructed from Caramani (2000) and Flora et al. (1983). The source of the polity IV index is Marshall and Jaggers (2000). The data on urbanization, educational attainment, population size and age structure come from Flora et al. (1983), Mitchell (1998), Vanhanen (2003) and Maddison (1991, 2003). Tax and public spending data are from Flora et al. (1983) and Mitchell (1998). The data on real GDP at international prices are from Maddison (2003). Trade and distance data are from Lopez-Cordova and Meissner (2005). Data on the existence of a population census come from Flora et al. (1983). The sources used to date the adoption of the income tax are listed below for each country.

1. Denmark (Sample period: 1815-1903). A tax on earned income was introduced permanently in 1903 after the so-called "system change" in 1901 where the peasants' party finally took hold of government. This date is confirmed by Webber and Wildavsky (1986), Peters (1991, p. 230), Encyclopaedia Britannica (1911, Vol. XIV p.357), Philip (1965, chapter 9) and the National Danish Encyclopaedia (1998, Vol. 9). The first temporary income tax was levied in 1789. Other temporary incomes

taxes were levied in 1809, 1848-50, 1864, and in 1867-70. In 1803, a sort of income tax at the municipal level (on wealth) was introduced. However, the first proper tax on earned income at the municipal level was the income tax levied in Copenhagen from 1861 (Philip, 1965, chapter 9).

2. Sweden (Sample period: 1815-1902). We date the introduction of the income tax to 1902 when a progressive income tax with self-declaration was adopted. This dating is supported by Philip (1965, chapter 9), Nationalencyclopedin (1995), Messere (1998, p. 328). Webber and Wildavsky (1986, p. 344), however, put the date as 1897 and this is repeated by Peters (1991, p. 320), but this seems to be a mistake that can be traced back to Seligman (1911). During the Napoleonic Wars an income tax was introduced in 1809-10 but given up in 1812 because of low yields. Although local taxes on property were widely used from 1843 onwards, it was not until 1920 that the municipalities levied a proper income tax (Svensk Uppslagsbok, 1953, Vol. 14).

3. Norway (Sample period: 1815-1892). At the end of the Napoleonic Wars, the Dano-Norwegian Oldenburg king was forced to cede Norway to the King of Sweden. Norway took this opportunity to declare independence, adopted a constitution based on American and French models but was soon after forced into a personal union with Sweden. Under this arrangement, Norway kept its liberal constitution and independent institutions, except for the foreign service. Philip (1965, p. 123) dates the introduction of the income tax in Norway to 1892; a date which is confirmed by Store Norske Leksikon (1999) and Encyclopaedia Britannica (1911, Vol. XIV, p. 358). Webber and Wildavsky (1986) and following them Peters (1991, Table 7.1) report that the income tax was introduced in 1905, but this date corresponds to the introduction of progressive income taxation. From 1882 an income tax was levied by the local authorities (kommune) and the modern income tax can be traced back to the Napoleonic Wars when a temporary tax on income was imposed (Store Norske Leksikon, 1999).

4. Italy (Sample period: 1861-1864). The income tax (imposte e reddito) in Italy was introduced three years after unification, i.e., in 1864, and followed the British model of schedules (Seligman, 1911, p. 340; Webber and Wildavsky 1986; Peters, 1991, Table 7.1; Enciclopedia Italiana, 1937, Vol. XVIII). The tax code was amended in 1866, 1867, 1870, 1873, 1874, 1877, 1894 and 1907. The income tax, introduced in 1864, unified all taxes on mobile wealth from the old states, where mobile wealth meant that "the tax was imposed upon all incomes except those subject to the real estate tax and excepting also the income from government securities" (Enciclopedia Italiana, 1937).

5. Germany (German Empire (1871-1918) and Weimar Republic (1919–33)) (Sample period: 1871-1920). The first federal income tax was introduced in 1920. Before then only the states could levy income taxes. During the Napoleonic Wars some of the states including Prussia in 1808 introduced a temporary income tax. State income

taxes were introduced on a permanent basis in Prussia in 1851, in Hessen in 1869, in Sachsen in 1874, and in Baden in 1884 (Philip, 1965, p. 122; Seligman, 1911, p. 233ff; Webber and Wildavsky, 1986; Encyclopaedia Britannica, 1911).

6. Austrian Empire (Sample period 1815-1849). The Austrian Empire (German: Kaisertum Österreich) officially lasted from 1804 to 1867, followed by the Austro-Hungarian Empire. It was a monarchy with no elections. The personal income tax was introduced permanently in 1849 (originally envisaged as a temporary measure) after the revolution in 1848. Groschen (1872, p. 45) reports that the yields from the (imperial) income tax was an inappreciable item in the years following its adoption. Temporary attempts at an income tax were noted in 1743 and again in the war taxes of 1778, 1789 and 1790 (Seligman, 1932, vol. VII; Webber and Wildavsky, 1986, p. 344; Peters, 1991, Table 7.1; Encyclopaedia Britannica, 1911, Vol. XIV, p. 357). It was not possible to establish for sure if there were any local income taxation in Austria before 1849, but the detailed discussion in Sieghart (1898) suggests that this was not the case. We code the local income tax dummy as zero from 1815 to 1849.

7. Netherlands (Sample period 1830-1993). At the congress of Vienna, Holland and Austrian Belgium were united to the Kingdom of Holland. However, Belgium declared its independence in 1830, and we treat the two as independent states from 1830 onwards. The income tax became a permanent feature of the tax system in 1893 when a tax on income from business and professions (belasting op bedrijf en beroep) was introduced (Seligman, 1932, Vol. VII). The year before a tax on wealth had been introduced and the income tax is sometimes thus dated (e.g., Peters, 1991, chapter 7). The origins of the income tax can be traced back to 1797 when something like an income tax was temporarily introduced (Webber and Wildavsky, 1986, p. 337). Stuart (1898) indicates that "several years before [the national income tax was levied] a great many communes levied already income-taxes" (p. 325). However, no precise date is given and it appears that the tax in question was a kind of poll tax in which the tax per head depended on wealth rather than on income tax as such (Fritschy, 1997). We have coded the local income tax dummy zero for the years before 1893.

8. Belgium (Sample period: 1830-1922). Belgium became independent in 1830 (see notes for the Netherlands). The income tax was first introduced in 1922 (Peters, 1991, Table 7.1). The discussion in Goschen (1872, p. 44) suggests that there was no local (province, canton or commune level) income taxes in 1870. There are no mentioning of any income taxes, national or local, in Encyclopaedia Britannica (1911) or other sources. We, therefore, assume that no such taxes existed before 1922. The origins of the income tax can be traced back to 1797 when something like an income tax was temporarily introduced (Webber and Wildavsky, 1986, p. 337).

9. Switzerland (Sample Period: 1848-1941). At the end of the Napoleonic Wars in 1815, the independence of the 22 Swiss Cantons was guaranteed. However, it was not

until 1848 that a federation/central government was established. Although income taxation at the Canton level has a long history in Switzerland, it was not until 1941 that a permanent federal income tax was introduced (Flora et al., 1983). Temporary income taxes (on war profits) were introduced in 1915-19 and in 1934 (The Schweizer Lexikon, 1946, Vol. II; Der Grosse Brockhaus, 1954). The first Canton to adopt income taxation was Baselstadt in 1840 (Seligman, 1932; Webber and Wildavsky, 1986, p. 344; Encyclopaedia Britannica, 1911, Vol. XIV, p. 358).

10. France (Sample period: 1815-1911). Throughout the nineteenth century repeated attempts were made to introduce the income tax. They all failed and it was not until 1911 that an income tax law was finally approved by the Senate (Seligman, 1911, chapter 2; Philip, 1965, chapter 7; Webber and Wildavsky, 1986, p. 344). Peters (1991, Table 7.1) dates the income tax to 1909 but this was the year in which the tax bill was approved by the lower house and referred to the Senate. It appears, however, that it was not until 1914 that the legislation came into force and some sources date the French income tax to 1914 (e.g., Flora et al., 1983). We date the adoption of the income tax to 1911 when the relevant legislation was approved by the Senate, rather than to 1914/16 where revenues started to come in. A general income tax (the so-called compulsory loan) was introduced temporarily in 1793 as a sporadic war measure. Although various proxy income taxes, such as doors and windows taxes, were widespread throughout the 19th century, there appears not to have been any local income taxes before 1911. For a comprehensive discussion of the early attempts of income taxation in France, see Willis (1895).

11. The United Kingdom (Sample period: 1815-1842). The income tax, based on the schedule system, became a permanent feature of the tax system in 1842 (e.g., Daunton, 2001, Sabine, 1966). Income taxes had previously been used temporary during the Napoleonic Wars (1798-1802; 1803-1816). Incomes were not taxed at the local level before 1842 (or after).

List of Encyclopaedia

1. Encyclopaedia Britannica, 1911, eleventh edition, Cambridge University Press, Cambridge, U.K.

2. Den Store Danske Encyklopaedi (Danish National Encyclopaedia), 1998. Gyldendal, Copenhagen, Denmark.

3. Enciclopedia Italiana, 1937. Istitutio dell'Enciclopedia Italiana, Rome (with appendices published as recently as 2000).

4. Store Norske Leksikon, 1999. Kunnskapforlaget, Oslo.

5. Schweizer Lexikon, 1946. Encyclios-Verlag AG., Zurich.

6. Svensk Uppslagsbok, 1953. Forlagshuset Norden AB, Malmo.

7. Nationalencyclopedin, 1995. Bukförlaget Bra Böcket, AB, Höganäs.

8. Der Grosse Brockhaus, 1954. F.A. Brockhaus Wiesbaden, Munchen, Germany.

References

[1] Acemoglu, D., Robinson J.A., 2000. Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. *Quarterly Journal of Economics* 115(4), 1167-1199.

[2] Aidt, T.S., Dutta, J., Loukoianova, E., 2006. Democracy Comes to Europe: Franchise Extension and Fiscal Outcomes 1830-1938. *European Economic Review* 50, 249-283.

[3] Aidt, T.S., Dallal, B., 2007. Female Voting Power: The Contribution of Women's Suffrage to the Growth of Social Spending in Western Europe (1869-1960). *Public Choice*, in press.

[4] Aidt, T.S., Jensen, P., 2007. Tax Structure, Size of Government and the Extension of the Voting Franchise in Western Europe, 1860-1938. Faculty of Economics, University of Cambridge, Cambridge Working Papers in Economics 0715.

[5] Angrist, J. D., Krueger, A.B. 2001. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic. Perspectives* 15, 69–85.

[6] Ashworth, J., Geys, B., Heyndels, B., 2005. Determinants of tax innovation: The case of environmental taxes in Flemish municipalities. *European Journal of Political Economy* 22, 223-247.

[7] Beck, N., Katz, J.N., Tucker, R., 1998. Taking time seriously: time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science* 42(4), 1260-1288.

[8] Becker, G.S., Mulligan, C.B. 2003. Deadweight Costs and the Size of Government. *Journal of Law and Economics* XLVI, 293-340.

[9] Berry, F.S. 1988. *Tax Policy Innovations in the American States*. Ph.D. dissertation, University of Michigan, Ann Arbor, Michigan.

[10] Berry, F.S., Berry, W.D., 1990. State Lottery Adoptions as Political Innovations: An Event History Analysis. *American Political Science Review* 84, 395-415.

[11] Berry, F.S., Berry, W.D., 1992. Tax Policy Innovations in the States: Capitalizing on Political Opportunity. *American Journal of Political Science* 36, 715–742.

[12] Blais, A. Blake, D., Dion, S., 1993. Do Parties Make a Difference? Parties and the Size of Government in Liberal Democracies. *American Journal of Political Science* 37, 40-62.

[13] Boix, C., 2001. Democracy, Development, and the Public Sector. *American Journal of Political Science* 45(1), 1-17.

[14] Boix, C., 2003. *Democracy and Redistribution.* Cambridge University Press, Cambridge, UK.

[15] Breyer, F., Ursprung, H., 1998. Are the Rich Too Rich to be Expropriated: Economic Power and The Feasibility of Constitutional Limits to Redistribution. *Public Choice* 94, 135-156.

[16] Caramani, D., 2000. *Elections in Western Europe Since 1815: Electoral Results by Constituencies.* Basingstoke, Macmillian Reference, UK.

[17] Congleton, R.D., 2004. Economic Development and Democracy, Does Industrialization Lead to Universal Suffrage? *Homo Economicus* 21, 283-311.

[18] Cox, D.R., 1975. Partial likelihood. *Biometrika* 62, 269-76.

[19] Canon, B.C., Baum, L., 1981. Patterns of adoption of torts law innovations. *American Political Science Review* 75, 975-87.

[20] Cook, C., Paxton J., 1998. *European Political facts.* MacMillan Press, London.

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

[22] Dusek, L., 2006. Are Efficient Taxes Responsible for Big Government? Evidence From Tax Withholding. Unpublished Working Paper.

[23] Flora, P., with Alber, J., Eichenberg, R., Krausm, J.K.F., Pfenning, W., Seebohm, K., 1983. *State, Economy and Society 1815-1975.* Campus Verlag, Frankfurt, Germany.

[24] Fritschy, W., 1997. A History of the Income Tax in the Netherlands. *Revue Belge de Philologie et d'histoire* 75, 1045-1061.

[25] Goschen, S., 1872. *Local Taxation.* MacMillan, UK.

[26] Hansen, S.B., 1983. *The Politics of Taxation.* Preager, Westport.

[27] Hettich, W., Winer, S., 1988. Economic and Political Foundations of Tax Structure. *American Economic Review* 78, 701-12.

[28] Hettich, W., Winer, S., 1999. *Democratic Choice and Taxation: A Theoretical and Empirical Analysis*. Cambridge University Press, Cambridge, UK.

[29] Husted, T.A., Kenny, L.W., 1997. The Effect of the Expansion of the Voting Franchise on the Size and Scope of Government. *Journal of Political Economy* 105, 54-82.

[30] Hicks, A.M., Swank, D.H., 1992. Politics, Institutions, and Welfare Spending in Industrialized Democracies, 1960-82. *The American Political Science Review* 86, 658-674.

[31] Justman, M., Gradstein, M., 1999. The Industrial Revolution, Political Transition, and the Subsequent Decline in Inequality in 19th-Century Britain. *Explorations in Economic History* 36, 109-127.

[32] Kau, J.B., Rubin, P.H., 1981. The Size of Government. *Public Choice* 37, 261-274.

[33] Kenny, L., Winer, S., 2006. Tax Systems in the World: An Empirical Investigation Into the Importance of Tax Bases, Administrative Costs, Scale and Political Regime. *International Tax and Public Finance* 13, 181-215.

[34] King, G. and Zeng, L., 2001. Logistic Regression in Rare Events Data. *Political Analysis* 9(2), 137-163.

[35] Lindert, P.H., 1994. The Rise in Social Spending 1880-1930. *Explorations in Economic History* 31, 1-37.

[36] Lindert, P.H., 2004a. *Growing Public: Social Spending and Economic Growth Since the Eighteenth Century. Vol 1.* Cambridge University Press, Cambridge, UK.

[37] Lindert, P.H., 2004b. *Growing Public: Social Spending and Economic Growth Since the Eighteenth Century. Vol 2.* Cambridge University Press, Cambridge, UK.

[38] Little, R.J.A., 1992. Regression with Missing X's: a Review. *The Journal of the American Statistical Association* 87(420), 1227-1237.

[39] Lopez-Cordova, J.E., Meissner, C.M., 2005. The Globalization of Trade and Democracy, 1870-2000. National Bureau of Economic Research Working Paper Series, Working Paper 11117. Cambridge, Mass.

[40] Maddison, A., 1991. *Dynamic Forces in Capitalist Development: A Long-run Comparative View*. Oxford University Press, Oxford, UK.

[41] Maddison, A., 2003. *The World Economy: Historical Statistics*. OECD, Paris.

[42] Meissner, C.M., 2005. A New World Order: Explaining the International Diffusion of the Gold Standard, 1870-1913. *Journal of International Economics* 66, 385-406.

[43] Messere, K. (ed.), 1998. The Tax System in Industrialized Countries, Oxford University Press, Oxford, UK.

[44] Marshall, M.G., Jaggers K., 2000. Polity IV Project. Data set Users Manual. Center for International Development and Conflict Management, University of Maryland [www.bsos.umd.edu/cidcm/inscr/polity].

[45] Meltzer, A.H., Richard S.F, 1981. A Rational Theory of the Size of Government. *Journal of Political Economy* 89(5), 914-27.

[46] Mitchell, B. R., 1998. *International Historical Statistics: Europe, 1750-1993*. 4th ed., Macmillan Basingstoke, UK.

[47] Paldam, M., 1997. Political Business Cycles. In: D. Mueller (ed.), *Perspectives on Public Choice. A Handbook*. Cambridge University Press, Cambridge, UK.

[48] Peters, G., 1991. *The Politics of Taxation. A Comparative Perspective*. Basil Blackwell, Oxford, UK.

[49] Persson, T., Tabellini, G., 2003. *The Economic Effects of Constitutions.* Cambridge, Mass., The MIT Press.

[50] Peltzman, S., 1980. The Growth of Government. *Journal of Law and Economics* 23(2), 209-287.

[51] Philip, K., 1965. *Skattepolitik*. Gyldendal, Copenhagen, Denmark.

[52] Riezman, R., Slemrod, J., 1987. Tariffs and collections costs. *Weltwirtschaftliches Archiv* 123, 545-49.

[53] Rodrik, D., 1998. Why do More Open Economies Have Bigger Government. *Journal of Political Economy* 106, 1997-1032.

[54] Sabine, B. E. F., 1966. *A History of Income Tax*. George Allen & Unwin, London, UK.

[55] Seligman, E. R. A. 1911. *The Income Tax*. The Macmillan Company, UK.

[56] Seligman, E. R. A. (ed.), 1932. *International Encyclopaedia of the Social Science*. Macmillan, UK.

[57] Sieghart, R., 1898. The Reform of Direct Taxation in Austria, *Economic Journal* 8, 173-182.

[58] Smit, C., 2002. Omwille der billijkheid. De strijd over de invoering van de inkomstenbelasting in Nederland (For the Sake of Fairness. The Struggle for the Implementation of a Dutch Income Tax), Wereldbibliotheek, Amsterdam.

[59] Stuart, A.J.C., 1898. Progressive Taxation in Holland. *Economic Journal* 8, 325-332.

[60] Tanzi, V., Schuknecht, L., 2000. *Public Spending in the 20th Century.* Cambridge University Press, Cambridge, UK.

[61] Tridimas, G., Winer, S.L., 2005. The Political Economy of Government Size. *European Journal of Political Economy* 21(3), 643-666,

[62] Vanhanen, T., 2003. *Democratization: A Comparative Analysis of 170 Countries.* Routledge.

[63] Wagner, A., 1883. *Grundlegung der Politischen Oekonomie.* 3rd ed.C.F. Winter, Leipzig.

[64] Walker, J.L. 1969. The Diffusion of Innocations Among American States. *American Political Science Review* 63, 880-899.

[65] Webber, C., Wildavsky, A., 1986. *A History of Taxation and Expenditure in the Western World.* Simon and Schuster, New York.

[66] Winer, S.L., Rutherford, T., 1993. Coercive Redistribution and the Franchise. A Preliminary Investigation Using Computable General Equilibrium Modelling. In: A. Breton, G. Galeotti, P. Salmon, R. Wintrobe (eds.), *Preferences and Democracy*, Dordreht, Kluwer.

[67] Winer, S. L. and Hettich, W., 1991. Debt and tariffs. An Empirical Investigation of the Evolution of Revenue Systems. *Journal of Public Economics* 45, 215-242.

[68] Willis, H. P. 1895. Income Taxation in France. *Journal of Political Economy* 4(1), 37-53.

[69] Wittman, D., 1989. Why Democracy Produce Efficient Results. *The Journal of Political Economy* 97(6), 1395-1424.

[70] Wooldridge, J. M., 2002. *Econometric Analysis of Cross Section and Panel Data.* The MIT Press, Cambridge Mass.

[71] Zorn, C., 2005. A Solution to Separation in Binary Response Models. *Political Analysis* 13, 157-170.

Table 1: The timing of the income tax in Western Europe.

Country	Permanent Income Tax from	Episodes of Temporary Income Taxes	Local/State Income Taxes from	Revenue Yield reaches 5%
United Kingdom	1842	1798-1802, 1803-16	None	1844
Austrian Empire[f]	1849	1778, 1789-90	None[d]	1905
Italy	1864	None	<1861[a]	1877
Norway	1892	1809	1882	1892
Netherlands	1893	1797	None	1899
Sweden	1902	1809-12	1920	1903
Denmark	1903	1789, 1809, 1848-50, 1864, 1867-70	1861	1917[e]
France	1911	1793	None[d]	1918
Germany[c]	1920	1808	1851[b]	1924
Belgium	1922	1797	None[d]	1922
Switzerland	1941	1915-19, 1934	1840	1942

Notes: a. Income taxes levied by some of the old states before unification in 1861. b. Prussia introduced the income tax in 1851, followed by Hessen in 1869. c. German Empire (1871-1918) and Weimar Republic (1919–33). d. We could find no indication in the literature that local and/or state income taxes were used before the introduction of the national income tax. e. the threshold for Denmark was most likely reached before 1917. f. Austro-Hungarian Empire after 1867.
Sources: See Data Appendix.

Table 2: Overview of explanatory variables and mnemonics.

Category	Variable
Political factors (P)	1) The electorate (for parliamentary elections) in percentage of the enfranchised age and sex group, before women's suffrage, male population only (*suffrage*) 2) The share of seats held by left-wing parties in the lower chamber of parliament (*left-wing parties*) 3) Dummy variable equal to 1 in election years (*election year*) 4) The number of years until the next election (*years to next election*). 5) The polity IV index (*political competition*)
Social Learning & geographical diffusion (L)	1) *Geographical closeness* is defined as $CL_{ij} = \sum_i \frac{1}{dist_{ij}} A_j(t)$ where $dist_{ij}$ is distance between the capitals of country i and j. $A_j(t) = 1$ if country j adopted the income tax in year $\tau \leq t$, and is 0 otherwise. 2) *Trade connection* is defined as $TC_{it} = \sum_j T_{ij}(t) A_j(t)$ where $T_{ij}(t)$ is the volume of trade between country i and j in year t as a percentage of GDP in country i. 3) L*earning from temporary adoptions* is defined as geographical closeness, except that $A_j(t)=1$ if country j has a temporary income tax in year t and zero otherwise.
Tax collection cost (T)	1) *Cost index* is the sum of • *census* (dummy variable coded 1 in year t if the country had a population census at some $\tau \leq t$ and 0 otherwise). • *temporary income tax* (dummy variable coded 1 if the country at some point in the past had introduced a temporary income tax and the tax was later abolished and 0 otherwise). • *local income tax* (dummy variable coded 1 for the years after which a country started to levy income tax at the local, state or regional level and 0 otherwise). • *education attainment* (dummy variable coded 1 for the years after which enrollment in primary education as a percentage of all 5 to 14 years olds reached 60 per cent and 0 otherwise). 2) *Urbanization* measured as the percentage of the population living in towns with more than 10,000 inhabitants.
Spending pressures (S)	1) A dummy variable equal to 1 if a country was at war in year *t* and equal to 0 otherwise (*war*) 2) GDP per capita at 1990 International Geary-Khamis dollars (*GDP per capita*). 3) The size of the total population in 1,000s (*population*). 4) The percentage of the population above 65 years of age (*age structure*). 5) The percentage difference between total (central) government spending and total (central) government tax revenues, lagged one year (*deficit*).

Notes: The sources are listed in the Data Appendix.

Table 3: Logit estimates of the probability of adopting the income tax permanently, 1815-1941

Model	1	2	3	4	5	6	7	8
Suffrage	-0.054**	-0.025	-0.056**	-0.057*	-0.061***	-0.041**	-0.053**	-0.054**
	[-2.44]	[-1.59]	[-2.43]	[-1.89]	[-2.66]	[-2.12]	[-2.28]	[-2.43]
Cost index	1.472*	0.728	1.467*	1.636				1.476*
	[1.69]	[1.04]	[1.65]	[1.28]				[1.69]
Geographical closeness	146.00**	88.97***	150.80**	147.20*	124.40	151.10**	136.60	145.60**
	[1.99]	[2.72]	[2.03]	[1.95]	[1.38]	[2.09]	[1.59]	[1.99]
Log(GDP per capita)	-1.78	0.60	-1.61	-2.17	-1.36	-0.86	0.50	-1.73
	[-0.94]	[0.50]	[-0.83]	[-0.78]	[-0.69]	[-0.51]	[0.27]	[-0.92]
Log(population)	0.737*	0.27	0.822	0.728	0.963**	0.557	0.878*	0.718
	[1.65]	[0.77]	[1.60]	[1.58]	[2.09]	[1.30]	[1.75]	[1.61]
War	1.214	0.953	1.232	1.262	1.767	0.789	1.31	1.233
	[0.93]	[0.82]	[0.93]	[0.94]	[1.26]	[0.63]	[0.98]	[0.94]
Election year			0.326					
			[0.41]					
Years to next election			0.025					
			[0.42]					
Left-wing parties				-0.003				
				[-0.072]				
Political competition				0.032				
				[0.20]				
Local income tax					2.619**			
					[2.36]			
Education attainment						0.782		
						[0.56]		
Temporary income tax							-2.502	
							[-1.57]	
Learning from temp. adoptions								-4.508
								[-0.19]
Years without income tax	0.246		0.241	0.251	0.327	0.228	0.341	0.243
	[1.32]		[1.31]	[1.30]	[1.47]	[1.25]	[1.45]	[1.31]
Spline(1)	0.00026		0.00025	0.00027	0.00028*	0.00022	0.00032*	0.00026
	[1.64]		[1.54]	[1.59]	[1.65]	[1.50]	[1.72]	[1.62]
Spline(2)	-0.00018*		-0.00017*	-0.00019*	-0.00020*	-0.00015*	-0.00021*	-0.00018*
	[-1.83]		[-1.70]	[-1.76]	[-1.83]	[-1.68]	[-1.90]	[-1.81]
Constant	-4.820	-12.920	-7.174	-2.131	-9.001	-7.026	-19.620	-4.940
	[-0.34]	[-1.23]	[-0.46]	[-0.11]	[-0.65]	[-0.53]	[-1.27]	[-0.35]
Observations	642	642	642	642	642	642	642	642
LR $\chi^2(3)$ test for constant hazard		16.05***						

Notes: z statistics in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Logit estimates of the probability of adopting the income tax permanently, additional control variables.

Model	9	10	11	12
Suffrage	-0.186***	-0.017	-0.065**	-0.084***
	[-2.94]	[-0.98]	[-2.40]	[-2.67]
Cost index	4.808***	0.689	1.000	2.193**
	[2.79]	[0.70]	[1.23]	[2.10]
Geographical closeness	-30.562	-0.901	74.210	52.901
	[-0.45]	[-0.59]	[0.95]	[1.00]
Log(GDP per capita)	-5.873**	0.401	-2.667	-0.543
	[-2.42]	[0.78]	[-1.11]	[-0.34]
Log(population)	0.1302	0.4041	0.561	1.5744**
	[0.25]	[0.78]	[1.07]	[2.54]
War	-2.099	3.237	0.663	-0.514
	[-1.17]	[1.62]	[0.40]	[-0.31]
Urbanization	0.534***			
	[2.95]			
Trade connection		9.384*		
		[1.74]		
Deficit			0.067**	
			[2.54]	
Age structure				-1.983**
				[-2.56]
Years without income tax	-0.354	0.09	0.201	0.130
	[-2.04]	[0.60]	[1.07]	[1.03]
Spline(1)	0	0.0001	0.000233	0.0002
	[-0.24]	[0.97]	[1.40]	[1.42]
Spline(2)	-0.00010	-0.00010	-0.00018*	-0.0002*
	[-0.76]	[-1.17]	[-1.73]	[-1.95]
Constant	26.201	-4.144	4.54	-10.126
	[1.53]	[-0.32]	[0.26]	[-0.81]
Observations	345	263	517	490

Notes: z statistics in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.
Models (9), (10) and (12) have been estimated using Maximum Penalized Likelihood (MPL). When using Logit estimation, Stata drops the variable *war*, because some outcomes are predicted perfectly by this variable. Zorn (2005) refers to such a situation as 'separation' and recommends the use of MPL as the best defence against bias. We note that it matters little for the results to use this alternative estimation method. Because of missing data, we lose Austria, the Netherlands and UK, when estimating Model (9). With model (10) we lose Austria, Italy and the UK. When estimating model (12), we lose Austria.

Table 5: 2SLS Linear Probability model and Random Effects Logit Model.

Model	13	14	15	16	17	18
Dependent variable	Suffrage	Permanent income tax		Suffrage	Permanent or temp. income tax	
Suffrage		-0.00069**	-0.054**		-0.00069**	-0.028*
		[-2.10]	[-2.44]		[-2.10]	[-1.67]
Revolution	1.842			1.921		
	[0.86]			[0.96]		
Suffrage lagged three years	0.924***			0.945***		
	[37.8]			[42.5]		
Cost index	0.091	0.010	1.472*	-0.462	0.016	1.159
	[0.11]	[0.98]	[1.69]	[-0.51]	[1.39]	[1.60]
Geographical closeness	254.400	6.998	146.000**	228.600	7.824	97.840**
	[1.57]	[1.29]	[1.99]	[1.52]	[1.52]	[2.17]
Log(GDP per capita)	-0.279	-0.077	-1.779	0.727	-0.056	0.143
	[-0.19]	[-1.30]	[-0.94]	[0.50]	[-0.94]	[0.095]
Log(population)	0.750	0.006	0.737*	1.030**	-0.005	-0.270
	[1.50]	[1.15]	[1.65]	[2.34]	[-0.66]	[-0.70]
War	7.857	0.019	1.214	1.336	0.122	2.283**
	[1.42]	[0.37]	[0.93]	[0.39]	[1.33]	[2.55]
Years without income tax	0.036	0.00078	0.246	-0.087	-0.00038	0.049
	[0.31]	[0.41]	[1.32]	[-1.29]	[-0.16]	[0.56]
Spline(1)	-0.000066	0.0000018	0.00026	-0.00016	0.0000015	0.00012
	[-0.52]	[0.70]	[1.64]	[-1.55]	[0.44]	[0.96]
Spline(2)	0.000066	-0.0000016	-0.00018*	0.000089	-0.0000013	-0.000078
	[0.76]	[-0.86]	[-1.83]	[1.63]	[-0.67]	[-1.17]
Constant	-3.689	0.495	-4.819	-10.94	0.442	-5.988
	[-0.34]	[1.16]	[-0.34]	[-0.99]	[1.05]	[-0.48]
Observations	632	632	642	618	618	628
F-stat for exclusion of instruments	728.06***			923.06***		
J-statistic (p-value)		0.231			0.222	
Estimation technique	2SLS: first stage	2SLS: second stage	RE logit	2SLS: first stage	2SLS: second stage	RE logit

Notes: z statistics in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. The 2SLS model is a linear probability model. The J-statistic reports the p-value of a Hansen test of validity of over-identifying restrictions.

Table 6: Logit estimates of the probability of adopting the income tax temporary or permanently, 1815-1941.

Model	19	20	21	22	23	24	25	26
Suffrage	-0.024*	-0.016	-0.026**	-0.031*	-0.051***	-0.0096	-0.013	-0.044**
	[-1.77]	[-1.36]	[-1.97]	[-1.66]	[-2.73]	[-0.79]	[-1.13]	[-2.25]
Cost index	1.246*	0.922	1.147*	1.472*				1.534*
	[1.92]	[1.54]	[1.82]	[1.89]				[1.94]
Geographical closeness	94.300**	101.400***	89.330**	91.110**	31.060	85.930**	70.290	116.800**
	[2.23]	[2.92]	[2.14]	[2.09]	[0.71]	[2.13]	[1.53]	[2.28]
Log(GDP per capita)	-0.324	-0.295	-0.291	-0.989	2.331*	0.64	1.069	0.569
	[-0.30]	[-0.30]	[-0.28]	[-0.68]	[1.90]	[0.64]	[0.90]	[0.38]
Log(population)	-0.347	-0.423	-0.344	-0.39	0.0722	-0.249	-0.164	0.321
	[-1.06]	[-1.33]	[-1.11]	[-1.12]	[0.23]	[-0.81]	[-0.49]	[0.77]
War	2.129**	2.289***	2.281***	2.163**	3.120***	1.957**	2.174**	0.217
	[2.54]	[2.89]	[2.68]	[2.54]	[3.33]	[2.29]	[2.39]	[0.14]
Election year			0.076					
			[0.11]					
Years to next election			-0.045					
			[-0.86]					
Left-wing parties				0.02				
				[0.50]				
Political competition				0.041				
				[0.47]				
Local income tax					3.638***			
					[3.26]			
Education attainment						-0.101		
						[-0.10]		
Temporary income tax							-0.645	
							[-0.67]	
Learning from temp. ad.								224.400
								[0.60]
Years without income tax	0.0489		0.0586	0.0489	0.145	0.0358	0.0408	0.142
	[0.59]		[0.70]	[0.57]	[1.38]	[0.42]	[0.47]	[1.05]
Spline(1)	0.00012		0.00015	0.00012	0.00030**	0.0001	0.00012	0.00024
	[1.04]		[1.22]	[0.94]	[1.98]	[0.86]	[1.02]	[1.37]
Spline(2)	-0.00003		-0.00009	-0.00007	-0.00017**	-0.00006	-0.000079	-0.00014
	[-1.22]		[-1.41]	[-1.09]	[-2.22]	[-1.04]	[-1.21]	[-1.54]
Constant	-1.851	-0.595	-1.506	3.221	-23.630**	-7.030	-10.250	-16.500
	[-0.21]	[-0.07]	[-0.18]	[0.28]	[-2.15]	[-0.83]	[-1.06]	[-1.23]
Observations	628	628	628	628	628	628	628	628
LR $\chi^2(3)$ test for constant hazard		4.07						

Notes: z statistics in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. The p-value for the LR test is 0.25.

Table 7: Rare events logit estimates of the probability of adopting the income tax, 1815-1941.

Model	27	28	29	30
Dependent variable	Permanent income tax adoptions		Permanent and temporary income tax adoptions	
Suffrage	-0.045**	-0.051**	-0.021*	-0.046*
	[-2.03]	[-2.31]	[-1.66]	[-1.88]
Cost index	1.291*		1.148**	
	[1.77]		[2.04]	
Geographical closeness	112.900	92.240	90.150*	27.190
	[1.61]	[0.88]	[1.77]	[0.60]
Log(GDP per capita)	-1.410	-1.006	-0.419	2.073
	[-0.64]	[-0.40]	[-0.34]	[1.44]
Log(population)	0.644	0.830**	-0.263	0.117
	[1.31]	[2.12]	[-0.59]	[0.34]
War	1.455	1.995	2.048*	2.991**
	[0.87]	[0.98]	[1.73]	[2.36]
Local income tax		2.274*		3.397***
		[1.67]		[2.69]
Years without income tax	0.162	0.231	0.0319	0.114
	[1.02]	[1.00]	[0.35]	[0.79]
Spline(1)	0.00019	0.00022	0.000099	0.00026
	[1.20]	[1.05]	[0.73]	[1.21]
Spline(2)	-0.00014	-0.00016	-0.000066	-0.00016
	[-1.30]	[-1.13]	[-0.87]	[-1.38]
Constant	-4.208	-7.934	-1.122	-21.390
	[-0.25]	[-0.49]	[-0.10]	[-1.59]
Observations	642	642	628	628

Notes: Robust z statistics in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Estimates are corrected for rare events bias as suggested by King and Zeng (2001).

Table 8: Logit estimates of the probability of reaching a yield of at least five percent, 1815-1942.

Model	31	32	33	34	35[a]
Suffrage	-0.048**	-0.021	-0.058*	-0.044*	-0.041*
	[-2.08]	[-1.31]	[-1.81]	[-1.95]	[-1.65]
Cost index	0.069	0.878	0.367	0.166	-0.018
	[0.10]	[1.29]	[0.40]	[0.22]	[-0.03]
Geographical closeness	6.314	58.620**	7.290	5.081	15.010
	[0.15]	[1.96]	[0.17]	[0.12]	[0.28]
Ln(GDP per capita)	-4.040*	1.205	-4.675**	-4.338**	-3.577
	[-1.92]	[0.92]	[-2.00]	[-1.97]	[-1.25]
Ln(population)	0.598	0.281	0.544	0.562	0.510
	[1.30]	[0.76]	[1.13]	[1.20]	[0.88]
War	-0.400	0.922	-0.376	-0.646	-0.024
	[-0.30]	[0.79]	[-0.27]	[-0.47]	[-0.03]
Election year				-1.551	
				[-1.36]	
Years to next election				-0.093	
				[-0.55]	
Left-wing parties			0.007		
			[0.12]		
Political competition			0.081		
			[0.56]		
Years without income tax	0.032		0.023	0.019	-0.092
	[0.17]		[0.12]	[0.10]	[-0.62]
Spline(1)	-0.000017		-0.000023	-0.000016	-0.00006
	[-0.17]		[-0.24]	[-0.16]	[-0.57]
Spline(2)	-0.0000018		0.0000027	-0.0000033	0.000021
	[-0.03]		[0.05]	[-0.06]	[0.32]
Constant	17.140	-18.390	22.310	20.570	18.900
	[1.00]	[-1.59]	[1.18]	[1.14]	[0.78]
Estimation method	ML	ML	ML	ML	Rare Events
Observations	746	746	746	746	746
LR $\chi^2(3)$ test for constant hazard		15.25***			

Notes: z statistics in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. a. robust z statistics (see note to Table 7).