Universität St.Gallen

Does Female Suffrage Increase Public Support for Government Spending? Evidence from Swiss Ballots

Katharina E. Hofer

January 2015 Discussion Paper no. 2015-02

School of Economics and Political Science,
Department of Economics

University of St.Gallen

Editor: Martina Flockerzi
University of St.Gallen
School of Economics and Political Science
Department of Economics
Bodanstrasse 8
CH-9000 St. Gallen
Phone +41 71 224 23 25
Fax +41 71 224 31 35
Email seps@unisg.ch

Publisher: School of Economics and Political Science
Department of Economics
University of St.Gallen
Bodanstrasse 8
CH-9000 St. Gallen
Phone +41 71 224 23 25

Electronic Publication: Fax +41 71 224 31 35
http://www.seps.unisg.ch

Does Female Suffrage Increase Public Support for Government Spending? Evidence from Swiss Ballots[1]

Katharina E. Hofer

Author's address:

Katharina E. Hofer-Jaronicki, Ph.D. des.
SEW-HSG
Varnbüelstrasse 14
CH-9000 St.Gallen
Phone +41 71 224 2312
Fax +41 71 224 2302
Email katharina.hofer@unisg.ch
Website https://sites.google.com/site/katharinaehofer/

[1] For valuable comments, I thank Monika Bütler, Patricia Funk, Martin Huber, and Alois Stutzer, as well as participants at the Meeting of the European Public Choice Society (April 2013, Zurich, Switzerland), Spring Meeting of Young Economists (May/June 2013, Aarhus, Denmark), European Political Science Association Annual Meeting (June 2013, Barcelona, Spain), CESifo Venice Summer Institute (July 2013, Venice, Italy), European Economic Association Congress (August 2013, Gothenburg, Sweden), and Jahrestagung Verein für Socialpolitik (September 2013, Düsseldorf, Germany). I appreciate helpful input by the following discussants at conferences and workshops: Andreas Bernecker, Krisztina Kis-Katos, and Silke Übelmesser.

Abstract

In this paper, I challenge the notion that women prefer larger governments than men which is why extending the franchise to women has led to an increase in government spending in many industrialized countries. I estimate the average treatment effect of being female on support for government spending, by analyzing the voting outcomes of two similar Swiss referendum votes concerning the federal government's authorization to levy income, capital and turnover taxes. The first ballot took place shortly before the extension of suffrage to women in February 1971, and the other one directly thereafter. Based on municipal voting data, I relate the increase in the electorate to the difference in acceptance rates for the two propositions. Surprisingly, I find that approval for government spending is higher among the male population.

Further, I conduct a mediation analysis based on post-ballot surveys after comparable votes in 1981, 1991, and 1993. The intrinsic direct effect of being female proves to be the driving force behind the negative gender gap. In contrast, socioeconomic mediators like employment status or education turn out to play a weaker role.

Keywords

Female Suffrage; Gender Preference Gap; Voting; Direct Democracy; Mediation.

JEL Classification

J16, H10, D72.

1 Introduction

Several contributions find a positive relation between female enfranchisement and a subsequent rise in government spending as well as revenue pointing towards the existence of stronger female preferences for large governments: analyzing historical data from the U.S., Lott and Kenny (1999) find that the introduction of female suffrage raised government spending and revenue as gradually more women made use of their voting rights. Following Lott and Kenny (1999), Aidt and Dallal (2008) confirm their results for six western European countries for which the long-run effects are significantly larger than the short-run effects. Similarly, Abrams and Settle (1999) find that the introduction of female suffrage in Switzerland increased government spending on welfare issues by 28%, and also total government spending grew by about 12%. Aidt et al. (2006) confirm these findings for Europe. Bertocchi (2011) provides empirical support that allowing women to vote increases government spending, however, only in non-Catholic countries in which the cost of disenfranchisement is relatively high.[1] The literature attributes the increase in spending to women prefering larger governments than men.

Female risk aversion with the increased need for insurance from the state, or the breakdown of the families leading to higher divorce rates constitute some of the most commonly provided explanations for gender gaps in taxation preferences (Edlund & Pande, 2002). In marriage, husbands tend to earn more and transfer income to their wives who specialize in household production and care for the children (Becker, 1974). While income differences and specialization are internalized in marriage, the possibility of divorce, however, makes women more vulnerable economically since they might be rendered with a low income to care solely for the children.

However, the literature does not convincingly show where gender gaps originate from: are they due to "being female" or can they partly be explained by observable differences between men and women? For example, Meltzer and Richard (1981) famously hypothesized that enfranchising new constituents such that the median voter is poorer than before, increases demand for redistribution. Husted and Kenny (1997) exploit the repeal of literacy tests and poll taxes in the U.S. which hitherto prevented poor and foreign men from voting. They find a 15% increase in welfare spending from enfranchising poorer men. Consequently, one would expect that extending suffrage to women, who on average have lower incomes than men, should have a positive effect on redistributive spending. Not controlling for socioeconomic gender gaps, however, might falsely attribute some of the effects to the fact that women can vote, instead of having enfranchised a new group which is on average poorer than the former electorate.[2] The literature mostly relies on analyses of representative democracies.

[1] In contrast, Stutzer and Kienast (2005) who use the variation in the timing of female suffrage in Swiss cantons, the 26 Swiss states, surprisingly find that government expenditures at cantonal level decreased after enfranchising women. They conclude that the negative effect might stem from the existence of direct democracy instruments in Swiss cantons for which previous research shows that they are likely to lead to smaller government size (Feld and Matsusaka (2003) provide some evidence).

[2] In Switzerland, women earned 51% of the male hourly wage in 1930, 66% in 1971 and 67% in 1995 which shows how big the gender wage gap was despite its tendency to decrease over time (Swiss Economic and Social History Online

A potential mechanism through which female voting could increase government spending is through politicians' behavior. They either change their policies, or women elect new, more spending-friendly politicians. To understand whether the relationship between higher expenditures and female voting is causal or influenced by a third force, like more liberal thinking, the analysis of elections and politicians' behavior would be required. Yet, the literature is relatively silent on the mechanism which would lead to higher expenditures.[3]

Literature confirms that women vote more often in favor of higher expenditure, but the outcomes depend on the issue the money is spent for. For example, Aidt et al. (2006)) find that female voting rights increased spending on health, welfare and education. Miller (2008) documents rising levels of public health expenditure to enhance child welfare that can be attributed to the enfranchisement of women. While there exist clear-cut predictions and empirical evidence for gender preference gaps on several spending categories, explaining why women would prefer larger governments per se turns out to be more difficult.

This paper provides a direct way of analyzing gender preference gaps for government spending. Instead of deriving preference gaps indirectly via the development of government expenditures around the time of female enfranchisement, this paper relies on results from referendum votes constituting real voter choices. I analyze the voting outcomes of referendums laying down the constitutional basis for the Swiss government to levy income, capital and goods turnover taxes. They are a measure of preferences for the federal government's spending: without popular approval at the ballot, the Swiss government would be deprived of its authorization to levy federal taxes which are crucial for financing government expenditures. While taxation of income and consumption is commonly found all over the world, it is a Swiss particularity that voters even nowadays need to accept it's legislative basis in a referendum every few years. Therefore, over time a large number of comparable votes on the federal financial system exists.

The goal of the analysis is twofold: first, I estimate the size of the gender preference gap for government spending. Second, I decompose it into a direct gender effect and an indirect effect due to socioeconomic mediators like employment. For the total gender effect, I analyze the voting outcomes of two very similar referendum ballots in Switzerland concerning federal taxation of which one took place shortly before the extension of suffrage to women in 1971, and the other directly thereafter. The first proposition in November 1970 with a men-only suffrage was rejected at ballot. But the second proposition, which took place 7 months later with universal suffrage, was accepted. Since the two ballots took place under two distinct suffrage regimes, changes in voting outcomes can be directly attributed to changes in the electorate after accounting for the differences between the two propositions. The analysis is based on data from 2,143 Swiss municipalities. To isolate female approval for government spending, a similar reasoning to Lott and Kenny (1999) is employed. How much of the increase in voter approval for government spending can be explained by female voting

Database).

[3] Lott and Kenny (1999) also look at the politicians' voting behavior in the U.S. senate and find that after the introduction of female suffrage politicians voted more liberally. However, they do not show that women were more likely to vote for liberal politicians and did so because they desired higher government spending.

depends crucially on the intensity with which women made use of their voting rights. I therefore estimate the impact of the change in the number of voters on the change in the number of yes votes. I refine the estimates taking into account the share of women in municipal populations and canton fixed effects. Moreover, I take into account that the two ballot propositions, though very similar, are not identical but differ in so far that the second proposition included a time limit. This means that even in case of acceptance the second proposition would have required a new ballot after 10 years while the first one did not. Traditionally, permanent federal financial orders have been rejected in Switzerland, suggesting that the inclusion of a time limit is an important factor influencing voter decisions. I utilize voting results from a similar ballot in 1963 under the males-only suffrage to proxy for the difference in the content of the two ballots which might have led to some men changing their voting behavior between the two ballots. I also provide extensive evidence for the validation of this approximation.

The results show that men were significantly more likely to favor taxation and thus government spending than women. These results contradict the notion that women are per se more likely to support large governments.

In the second part, I disentangle the total average treatment effect of gender on support behavior into its direct and indirect component. As argued above, being female has an indirect effect on support for government spending which is mediated by socioeconomic variables like income and employment that are generally known to differ across men and women. Moreover, being female has a direct channel influencing support behavior which is intrinsically female. To decompose both potential channels, I use a mediation framework to non-parametrically estimate the (average) direct and indirect effects of being female. Estimates are based on inverse propensity score weighting, which allows to control for potentially confounding factors affecting both mediators and the outcome variable (Huber, 2013, 2014; Imai, Keele & Yamamoto, 2010).

I conduct the mediation analysis based on individual-level data from randomized post-ballot surveys after three comparable referendum votes about the federal fiscal order in 1981, 1991 and 1993. The results confirm that women are on average less likely to support the taxation propositions. The mediation analysis reveals a strong negative direct channel of being female on support for government spending. Though negative, there is little consistent evidence for the mediated effect. Consequently, mediating socioeconomic variables with negative gender gaps like full-time employment or having a high school degree play a smaller role for female preferences for government spending than direct effects.

This paper adds to the existing literature on the effects of franchise extension on government spending and revenue as well as gender preference gaps. The main innovation in my approach is to directly analyze the outcomes of ballots instead of relating suffrage to government spending. With the exception of Funk and Gathmann (2012) who explore gender preference gaps for different spending categories by utilizing post-ballot polls in Switzerland, literature has so far only analyzed the effect of female suffrage on the size of state expenditure. This approach, however, is imperfect since voters only elect politicians who finally decide about policies. By analyzing outcomes of referendum ballots, I provide evidence for how voters directly decide on taxation and consequently

government spending. I complement literature which emphasizes the importance of distinguishing between spending items when it comes to analyzing gender preference gaps. While women might be more likely to care for redistributive spending as can be inferred from Meltzer and Richard (1981), they might be indifferent or even opposed to other spending categories.

The remainder of the paper is organized as follows. Section 2 provides information on the institutional setting. Section 3 contains the empirical framework, data and results for the estimation of the total gender effect. In a similar manner, Section 4 deals with estimating the direct and indirect gender effect. The concluding remarks are in Section 5.

2 Institutional Setup

Beginning with the foundation of the Swiss state in 1848, duties were the main revenue source at federal level.[4] It took until the First World War, collapsing international trade and growing state expenditure before an income tax was introduced. But income was only taxed in times of need such as during the war, or when budgetary problems got out of hand in the 1930ies. In 1941 the Wehrsteuer (defense tax, an income and capital tax; referred to as direct federal tax in what follows) was introduced to finance growing military expenditure. After the Second World War, the direct federal tax remained in place to finance other state expenditure like the social security system but also new spending fields like education and culture. In addition, a goods turnover tax (Warenumsatzsteuer) on goods but not on services, resembling a value-added tax, was introduced also in 1941 (Stockar, 2007). However, both taxes lacked a constitutional basis, and were a product of an increased need of state revenue during war and emergency times.

Besides revenues from duties, the goods turnover tax and the direct federal tax were the most important revenue sources for the Swiss government. In the 1960ies, roughly 10 to 15 percent of revenues came from the direct federal tax, and around 25 percent from the goods turnover tax. Revenues from duties then dropped by 10 percentage points. (Swiss Statistical Office, 1973) The main reason for the decline was the increasing international integration and the general trend to reducing duties in connection with the World Trade Organization's rounds (Federal Announcement 1969 II, p.754). The lack of a permanent constitutional basis for levying federal taxes left some uncertainty about how to finance growing government expenditure. The main items of expenditure at federal level were defense and the social security system which together accounted for nearly 50 percent of total expenses. Other growing and new expenditure categories were infrastructure and energy, as well as culture and sports. Agricultural expenditure remained relatively stable at around 10 percent of total expenditure (Swiss Statistical Office, 1974).

A proposition to allow the state collecting a direct federal tax as well as the goods turnover tax on a constitutional basis without time limitations was issued in 1953. Because it involved amending a constitutional article, the issue was subject to a mandatory referendum[5], of which the outcome

[4] Information about the history of the Swiss Federal Tax are from Grütter (1968). Oechslin (1967) gives an overview of the overall development of the Swiss tax system.

[5] In Switzerland, all changes to the constitution have to be approved by the voters. When the change is proposed

Table 1: Chronology of Ballots Concerning the Swiss Federal Tax System

Ballot date	Time limit	Decision	% yes votes	Accepting cantons
06.12.1953	unlimited	rejected	42.0	3
24.10.1954	1955 - 1958	approved	70.0	21
11.05.1958	1959 - 1964	approved	54.6	17 1/2
08.12.1963	1964 - 1974	approved	77.6	22
15.11.1970	**unlimited**	**rejected**	**55.4**	**10**
06.06.1971	**1972-1982**	**approved**	**72.7**	**22**
12.06.1977	unlimited	rejected	40.5	1
20.05.1979	unlimited	rejected	34.6	0
29.11.1981	**1982-1994**	**approved**	**69.0**	**23**
02.06.1991	**unlimited**	**rejected**	**45.6**	**2 1/2**
28.11.1993	**1994-2006**	**approved**	**66.7**	**22**
28.11.2004	2006-2020	approved	73.8	22

Note For approval, the referendum needs more than half of total votes and at least 13 accepting cantons. In 1971, 19 cantons are "full" cantons while 6 cantons count only as "half" cantons. Data about acceptance are available on the homepage of the Swiss Federal Chancellery, http://www.bk.admin.ch. The time limits are from Federal Announcements published by the Swiss Federal Archive. See appendix for information on how to access the Federal Announcements. Votes in bold are used in empirical part.

is binding in Switzerland (Linder, 2007). Since public finances are a core element of a state, a wealth of similar ballots concerning the federal government's admission to file taxes exists. Table 1 gives an overview of all relevant ballots between 1953 and 2004. Even nowadays, it remains a Swiss particularity that citizens have to approve the federal financial order. Without acceptance, the federal government would not have the right to levy federal taxes. For a referendum to be successful, the majority of voters and a majority of cantons is required. In 1971 there were 19 cantons and 6 half cantons who's votes counted as a 1/2 vote.

The 1953 proposition was rejected. Only one year later, a similar proposition to include the federal right of levying income, capital and goods turnover taxes in the constitution but with a time limit of four years was put to the vote, and eventually approved by the male voting population. It was followed by another temporary financial order from 1959 to 1964.[6] The time limit forced the government to prepare a new legislation regarding the financial order in 1962. Essentially, it was an extension of the old provision for another 10 years with some minor changes (Federal Announcement 1962 I, p.997)[7]. Again, the proposition was accepted at ballot.

by the parliament, this requires a referendum vote. In contrast, changes proposed by (groups of) citizens are called initiatives.

[6] The comparably low acceptance rate of this financial order with time limit is most likely due to a heated debate of the large interest groups (Bolliger, 2010).

[7] All federal announcements (Bundesblatt) are collected by the Swiss Federal Archive (Schweizerisches Bundesarchiv) and published by the Federal Chancellery (Bundeskanzlei). A detailed list and possibility of online access is described in the online references section in the appendix.

The first of the two ballots at the core of this paper's analysis took place on 15 November 1970. The second referendum took place with a new electorate on 6 June 1971: Switzerland was the last European country to grant women voting rights at federal level on 7 February 1971. It came into force on 16 March 1971. Swiss women were demanding suffrage more intensively in the aftermath of both world wars when democratization was spreading all across Europe. They also received support from male politicians who recognized that the women's position in society had changed to a more active role in public live and private employment (Ruckstuhl, 1986). However, female suffrage in Switzerland could only be brought about by a constitutional amendment, which required the male population to hold a vote on extending the franchise. While at a first ballot in 1959 female suffrage was rejected with 66.9% of the male votes,[8] a second run in 1971 saw the majority of voters and majority of cantons accepting the constitutional amendment. The next paragraphs describe the propositions on government spending and voting results in more detail.

Ballot proposition 1: 15 November 1970

Facing a big budget deficit and the urgent need to ensure government revenue for the next years, the government and parliament proposed to discard the time limit and the maximum taxes from the constitution in the "federal enactment about the amendment of the federation's financial order"[9] (Federal Announcement 1969 II, p.749).

The new ballot proposition had to take into account that regardless of the good economy federal expenditures were expected to exceed revenues by a large amount. Therefore, income, capital and goods turnover taxes had to be increased and old rebates reduced. In more detail, the tax burden would be shifted from the direct income tax to the indirect goods turnover tax such that revenue from the goods turnover tax would increase considerably and revenue from income taxes would stay roughly constant. The proposition wanted to increase the goods turnover tax for retailers from 3.6 to 4 percent, and for wholesalers from 5.4 to 6 percent. The income tax set in progressively at an annual income of 8,500 Swiss Francs after deductions (7,700 Swiss Francs before). It allowed for deductions for married couples (2,500 Swiss Francs), children under 18 years and dependents (1,200 Swiss Francs) (Federal Announcement 1970 II, p.3). Regarding the income tax, high income households would be worse off with the new regulation than low income households because of a more progressive system. Married couples or families with many children would be better off than with the old regulation.

The government argued that an increase in goods turnover taxes to generate state revenue was the preferable revenue source: it was not a typical consumption tax because of various exemptions for goods of daily use like food. It mainly taxed investment goods purchased by firms and the government, in addition to goods like alcohol, tobacco, and clothing purchased by households (Federal Announcement 1969 II, p.778). However, there seemed to be a general uncertainty about who would carry the burden of the higher goods turnover tax. Presumably the biggest load would

[8] Only three francophone cantons, Geneva (60.0%), Neuenburg (52.2%), and Waadt (51.3%) had a majority favoring universal suffrage. They were also the first three cantons to introduce universal cantonal suffrage.

[9] Te original German title is "Bundesbeschluss über die Änderung der Finanzordnung des Bundes".

be paid by enterprises.

Critics of the proposition mostly pointed to an unsatisfactory regulation concerning the Swiss cantons (Année Politique Suisse, 2012). In particular it lacked a clear division of revenue and expenditures between the federal government and the cantons because direct income taxes were an important revenue source for cantons and municipalities (Federal Announcement 1969 II, p.773).

All major parties, associations and unions recommended their voters to accept the proposition. Exceptions were the small Liberal Party of Switzerland (LPS), and the Labor Party (PdA) who opposed the proposition for not being progressive enough (Année Politique Suisse, 2012). These almost unanimously positive voting recommendations indicate the importance of the issue at stake. On 15 November 1970 the Swiss voters - which was the male eligible population at that point - rejected the proposition in a mandatory referendum. Though 55.4% of the voters were in favor the proposition, it failed to accomplish a majority of cantons: in 13 of 22 cantons the approval rate was below 50 percent. The rejecting cantons were mainly concentrated in rural, German-speaking areas.

Ballot proposition 2: 6 June 1971

The Swiss government immediately prepared a new proposition[10] because it urgently needed more revenue sources to finance growing state expenditure (Federal Announcement 1970 II, p.1581). In the major parts, the new proposition was identical to the old proposition, but it had the following changes. The biggest change included a time limit of 10 years (Federal Announcement 1971 I, p.487). This meant that in case of approval at the polls, the federal financial order had to be voted upon again in 1980 at the latest. As a further change, income tax ceilings of 9.5 percent for natural persons and 8 percent for legal persons were included. The income tax schedule became slightly more progressive and started to tax individuals at incomes after deductions of 9,000 Swiss Francs. These measures were taken to account for price inflation. Importantly, the only essential change between the first and the second ballot proposition was the inclusion of the time limit. Comparing the precise wording of both legislative texts shows that they are almost identical in all paragraphs.[11] Consequently, if a man changes voting behavior between the ballots, content-wise the only obvious reason can be the time limitation of the second proposition.

As in the first proposition, the parties and associations almost unanimously asked the voters to accept the proposition in their voting recommendations. Only the Labor Party (PdA), the Swiss Evangelic Party (EVP), and the Alliance of Independents (LdU) were opposed to the proposition because it disregarded deductions for working wives and was not progressive enough (Année Politique Suisse, 2012). With universal suffrage the ballot proposition concerning the Federal Tax System was accepted by a majority of voters (72.7%) and all cantons. Figure 1 shows the approving (white) and rejecting (grey) cantons for both ballots. The maps are based on swissvotes.ch.[12]

[10] "Federal enactment about the continuation of the federation's financial order". Original title in German is "Bundesbeschluss über die Weiterführung der Finanzordnung des Bundes"

[11] The comparison is available from the author on request.

[12] A project of the Institute of Political Science at the University of Bern, Switzerland, and the Année Politique Suisse.

Ballot propositions in 1981, 1991, and 1993

The votes in 1981, 1991, and 1993 are used in the empirical part to decompose the total gender preference gap into its direct and indirect components. The fiscal order approved by voters in 1971 was about to phase out in 1982. Both referendums in 1977 and 1979 which tried to change the tax system from the goods turnover tax as explained above to a value added tax (VAT) similar to those in other European countries were rejected at ballot. The financial order voted in 1981 and limited to the years 1982 to 1994 was therefore essentially a continuation of the old financial order from 1971. Minor changes included reliefs in the direct income tax which had to be compensated by increases in the goods turnover tax (Federal Announcement 1981 I, p.20). In 1991 government and parliament tried again to switch from goods turnover taxes to the VAT. Again, the proposition was rejected at ballot. It was argued that the proposition might have been a too complex package, which led to the rejection. Two years later a new financial order in a less complex proposition finally brought about the change to the VAT system, and secured the fiscal fundament for the federal state until 2006 (Année Politique Suisse, 2012).

Figure 1: Cantonal Approval Rates for Ballots 1 (15 November 1970) and 2 (6 June 1971)

Note: Accepting (white) and rejecting (grey) cantons. Based on swissvotes.ch

3 Average Gender Preference Gap for Government Spending

3.1 Baseline Empirical Framework

I aim at estimating approval for government expenditures by gender to find out whether a preference gap for government spending exists as suggested by the literature. A finite population is divided into two groups G, with the realizations $g = 1$ (female) and $g = 0$ (male). For each individual, outcome Y is observed which takes the value 1 if an individual is a supporter of government spending, and 0 else.

The total effect τ of gender on the probability of supporting government spending is the average treatment effect (ATE). It reflects the total effect of being female on the outcome variable. When averaging over the entire population, it can be written in the following way using expected values:

$$\tau = E[Y(1)] - E[Y(0)] \tag{1}$$

The ATE is the difference in expected outcomes between women and men. In fact, the ATE is the average difference in support for government spending by gender, and coincides with the total gender preference gap. Positive values of (10) point to stronger female than male preferences for government spending as suggested by the literature cited above. For negative values the opposite would be true. If the effect was zero, no gender preference gap would exist.

When conducting the analysis based on referendum votes, the outcome Y takes the value 1 if a voter votes "yes", and zero else. Hence, the empirical counterpart of $E[Y(1)]$ is the female acceptance rate and $E[Y(0)]$ the male acceptance rate. Then, the ATE is simply the difference between gender acceptance rates. Since the observed total acceptance rate can be written as the gender-weighted sum of female and male acceptance rates, it is enough to estimate the acceptance rate of women. The calculation of the ATE follows immediately.

3.2 Estimation Strategy: Identifying the Total Gender Effect

In my empirical strategy to estimate the female acceptance rate, I follow the idea of Lott and Kenny (1999) who recognize that the effect of female suffrage on voting outcomes depends on how intensely women make use of their voting rights. The intuition is that changes in voting outcomes can be explained by changes in the electorate's composition. For identification, I use the voting results of the two Swiss referendum ballots in 1970 and 1971. As explained above they are very similar content-wise, but women were only allowed to vote on the second date. The main independent variable in the analysis is thus the change in the number of voters between the two ballots and the dependent variable the change in the approving votes.

What makes the analysis more complicated is the fact that the ballots are not entirely identical. The main difference between the propositions is a time limit of ten years in the second proposition. I have noted before that propositions regarding the federal financial order including time limits have also been approved by the male voting population, like in 1954, 1958, and 1963. This means that not only women are expected to vote in favor of the proposition but also some men should change their minds and vote yes instead of no. To account for the less extreme content of the second proposition, I use voting results of the related vote in 1963 which also included a time limit and was accepted by a large margin (cf. Table 1). The true number of men who have switched from voting no to voting yes in municipality m, Δmen_m, is approximated by a variable $\Delta\widehat{men}_m$ calculated from the old voting results, and the error term ϵ_m. Then the true number of men changing their voting behavior can be written in the following way.

$$\begin{aligned} \Delta men_m &= \Delta\widehat{men}_m + \epsilon_m && (2) \\ &= yes_{m1963} * \frac{voters_{m1970}}{voters_{m1963}} - yes_{m1970} + \epsilon_m && (3) \end{aligned}$$

Δmen_m is the difference between the number of yes votes in the years 1963 and 1970, normalized by the growth in the number of voters during the seven year difference. In this manner, I can proxy the change in male approval rates when propositions include a time limit or not. The error term

ϵ_m reflects that this variable is only approximated, and voter preferences in municipalities might have slightly changed between 1963 and 1970.

The goal is to estimate the female acceptance rate for the second ballot proposition, $acceptance_f$. For notational ease, denote the year 1970 by $t = 1$, and 1971 by $t = 2$. Define the change in the number of yes votes between the ballots in municipality m as $\Delta yes_m \equiv yes_{m2} - yes_{m1}$, and the change in the number of voters as $\Delta voters_m \equiv voters_{m2} - voters_{m1}$.
The female acceptance rate can be written in the following form:

$$acceptance_f = \frac{\Delta yes_m}{\Delta voters_m} - \frac{\Delta men_m}{\Delta voters_m} \tag{4}$$

Female acceptance is the change in the number of yes votes relative to the change in the electorate, net of the change in male voting behavior where Δmen_m is defined in (3). Equation (4) can be easily transformed to

$$\Delta yes_m - \Delta men_m = acceptance_f * \Delta voters_m \tag{5}$$

In the last step, I use equation (3) to account for the approximation of men changing their voting behavior. This leads directly to the baseline estimation equation.

$$\Delta yes_m - \Delta \widehat{men}_m = \beta_1 \Delta voters_m + \epsilon_m \tag{6}$$

Under the exogeneity assumption $E(\Delta voters'_m \epsilon_m) = 0$, the coefficient β_1 of a linear regression thus identifies the female acceptance rate $acceptance_f$ which is the main object of interest here.

Note, that (6) is in fact a first difference equation. Therefore, by definition estimating (6) with least squares should be equivalent to estimating a fixed effects model of the following form where υ_m are municipal fixed effects.

$$yes_{mt} - men_{mt} = \beta_{FE} voters_{mt} + \upsilon_m + \epsilon_{mt}, \qquad t = 1, 2 \tag{7}$$

The equivalence of both coefficients $\beta_1 = \beta_{FE}$ only holds under the assumption of strict exogeneity, $E(\Delta voters'_m \epsilon_m) = 0$ (Wooldridge, 2010). Hence, I run both regression (6) and (7) to explore if exogeneity poses a problem.

Typically, for estimations using data from federal states, it is common to include canton fixed effects in the regression. I therefore run a specification including canton fixed effects ξ_c. They account for cantonal, time-invariant factors between both ballots like cultural differences, cantonal female voting rights, or compulsory voting rules.[13]

$$\Delta yes_m - \Delta \widehat{men}_m = \beta_2 \Delta voters_m + \xi_c + \epsilon_m \tag{8}$$

I further extend the analysis to refine the measure of the female acceptance rate. Intuitively, if there were no women in a municipality, female suffrage should not have any effect on the participation

[13] Note that female suffrage for cantonal votes is independent of federal regulation. Though some cantons introduced female voting close to the federal switch, many did not.

rate. Conversely, if a municipality was populated by women exclusively, the change in voters would reflect the change in acceptance with certainty. Given that foreigners have no federal voting rights in Switzerland, I calculate the share of Swiss adult women in a municipality, $\%women_m = (eligible_{m2} - eligible_{m1})/eligible_{m2}$, and multiply the change in voters $\Delta voters_m$ with this variable. The validity of the variable $\%women$ is implicitly based on the assumption that the Swiss population did not increase between November 1970 and June 1971. While this constitutes a simplification, the population size should not have grown a lot during 7 months, and the variable should be a good approximation. I run the following regression, and a further specification also including canton fixed effects.

$$\Delta yes_m - \Delta \widehat{men}_m = \beta_3 \%women_m \Delta voters_m + \epsilon_m \tag{9}$$

The female acceptance rate can then be recovered by calculating $\widehat{acceptance}_f = \widehat{\beta}_3 * \%women_m$. The population-weighted average share of women in the Swiss adult population amounts to 53.8 percent, but it varies between 0 and 72.2 percent. In total, 22 (11) municipalities have shares below (above) the band of 40 to 60 percent. Municipalities with low shares of Swiss adult women are located in 4 cantons Bern, Graubünden, Waadt, and Wallis, and have only 89 inhabitants on average.[14] Municipalities with high shares of Swiss women are located in six cantons and have 812 Swiss adults on average. Most of the shares are, however, close to 60% and thus not such big outliers.

I run different specifications excluding municipalities with extreme shares of female population. On average, 1,630 Swiss adults live in a municipality, and 95 percent of observations have less than 5,500 eligible citizens. In further specifications, I also exclude very large (above 10,000 Swiss adults) and very small (below 100 Swiss adults) municipalities. Excluding large municipalities technically means that data from three cantons with district data are excluded, as well as 32 relatively large municipalities. In this way, potential outliers can be accounted for.

3.3 Identifying Assumptions

The validity of the above estimations is based on two main identifying assumptions. They reflect that the estimations are based on aggregate data and not on individual observations. The first one is the following:

Assumption 1 *The change in the number of voters between ballots 1 and 2 stems from the female part of the population.*

Or put differently: men are not more likely to participate once women are enfranchised. This might be due to a decrease in the marginal benefit to vote when the electorate roughly doubles. Further evidence which confirms that men should be unlikely to increase their participation comes from comparing participation rates of the two similar ballots in 1953 and 1954, both with male suffrage

[14] Bern, Graubünden, Vaud, and Vallais are the four largest cantons in terms of area, and encompass some of the least densely inhabited regions in Switzerland.

only. For the first ballot in 1953 which did not include a time limit and was highly contested 60.27% of the male eligible population turned out. In contrast, the less contested proposition including a time limit of 4 years in 1954, drew only 46.77% of eligible men to the polls.[15] Hence, the additional participation is likely to be a lower bound for female participation.

I provide further evidence from cantonal votes that the change in the number of voters between the ballots reflects the number of female voters: on the day of the second vote, June 6 1971, 11 cantons have not yet introduced female suffrage for cantonal votes. This means that women in these cantons were allowed to vote on federal issues but not on cantonal ones. Of these 11 cantons, 5 held cantonal votes on 6 June 1971.[16] Consequently, the number of voters who voted on cantonal issues on this day were men with certainty. Owing to this constellation, the number of women voting can be calculated by taking the difference between the number of voters in the federal vote and the highest number of voters in the cantonal vote. I call this the true number of women. Next, I compare this number with my approximated number of female voters, $\Delta voters_c \equiv voters_{2c} - voters_{1c}$ for each of the five cantons c. On average, the true number of women exceeds my approximated number of female voters by 1.745 percentage points with a standard deviation of 7.794 percentage points. The standard deviation is relatively large because the values for two cantons are relatively far away from unity. This might be a result of the fact that whoever votes on cantonal issues does not necessarily have to vote on the federal issue and vice versa, so that there might be some roll-off between the votes. I compare federal and cantonal turnout in the five cantons for all ballot days between 1973 and 2010 which gives 287 cantonal ballot day observations after all 5 cantons introduced cantonal female voting rights. Data show that cantonal turnout for these 5 cantons is on average 1.520 percentage points higher than federal turnout with a standard deviation of 1.616. Therefore, the difference between the true and estimated number of female voters can partly be explained by selective abstention between cantonal and federal votes.

Ballot 2 is the first federal voting date after the introduction of female suffrage on which voters decided on two bills.[17] Intuitively, at the day of the introduction of universal suffrage not all women make use of their new rights immediately. It takes time until women grow accustomed to the possibility of voting (e.g., for the U.S. Lott and Kenny (1999) find that the full effect of female suffrage is revealed after 43 years of female voting rights). To visualize the effect of female suffrage on voter participation in Switzerland the number of voters normalized by the Swiss population of age for elections to the Swiss parliament between 1951 and 1991 is presented in Figure 2.[18] In this time period, parliament was elected every 4 years. The x-axis shows the election year t before and

[15] Turnout at a particular election day is influenced by all votes on the ballot list. Luckily, on the ballot day in 1954 there was no other federal vote, so turnout was truly for the vote under investigation. On the ballot day in 1953 there was one additional federal vote about the protection of waters. It received a narrowly smaller turnout than the other vote so it is safe to assume that the vote about the federal financial order was the main reason to turn out on that day.

[16] These were the cantons Bern, Graubünden, Schwyz, Thurgau, and Uri. Data on cantonal voting results is from the Centre for Research on Direct Democracy available online on www.c2d.ch.

[17] The other proposition was about the protection of humans and their environment.

[18] It is preferable to depict turnout for parliament than turnout for referendum votes since participation for referendums varies a lot which might be due to the importance of an issue or campaigning effects.

Figure 2: Participation Rate in Parliamentary Elections 1951-1991

Note: Participation rate ($voters_t$ / *Swiss adult population*$_t$) for parliamentary elections over time. Without (1951-1967) and with female suffrage (since 1971). Data from the Swiss Statistical Office.

after the introduction of female suffrage in 1971. The number of voters normalized by the Swiss adult population $participation_t = \frac{voters_t}{adults_t}$ is depicted on the y-axis. I take the total number of Swiss people above 20 years old from Swiss censuses in 1950, 1960, and 1970, and interpolate the numbers for the inter-census years. The data are from the Swiss Statistical Office. For the later years, $adults_t = eligible_t$ can be directly used from official election data. The fraction of voters as compared to the total adult population was steadily decreasing before the introduction of female suffrage. As expected, the participation rate jumps by more than 25 percentage points in the 1971 election with universal suffrage. However, afterwards the participation rate has a decreasing trend again. This contrasts with the observation of Lott and Kenny (1999) who show that the turnout rate in the U.S. continued increasing many years after the introduction of female suffrage. Thus on average women in Switzerland made use of their voting rights relatively quickly. This is not surprising since female suffrage was introduced relatively late in history. The timing coincides with higher education levels among women than in countries that enfranchised women around the first world war. Also, in some cantons women have received female voting rights for cantonal votes independently of federal voting rights such that they have gathered some voting experience even before 1971.

However, the second ballot took place only four months after female enfranchisement and was the first voting date that allowed all Swiss women to vote on federal issues. So while women might have already been more likely to vote in the federal elections at the end of 1971, they might have hesitated to participate in their very first voting opportunity. The participation decision is a selection into voting based on the citizens' utility from voting. The question is, thus, whether participation is a function of underlying variables which would render the participating female population unrepresentative for tax preferences of the female population. I overcome this problem

by carrying out the second part of my analysis with post-ballot data of votes that took place well after female suffrage was introduced.

The second identifying assumption relates to the change in the ballot propositions by including a time limit restriction in the second one.

Assumption 2 *Men who have approved of the first proposition should also be in favor of the second one which includes a time limit and is thus less radical. The inclusion of the time limit in the second ballot proposition makes some men switch from voting no to voting yes.*

Based on these assumptions, I have constructed a measure of the number of men changing their voting behavior from rejecting to approving, Δmen_m, as stated in equation (3) which uses voting data from 1963 and 1970 under the male-only voting regime. The validity of this proxy relies on the assumption that male preferences regarding government spending were relatively time constant between 1963 and 1970 and differences in acceptance are due to the inclusion of a time limit in the 1963 proposition and the lack of it in the latter. To substantiate this claim, I again provide evidence from the two comparable ballots on the federal financial order in 1953 and 1954. Recall, the first one had no time limit and was rejected, while the second one had a time limit and was approved by the entirely male electorate. The population-weighted average difference in approval rates for the two propositions was 27.7 percentage points which is substantial, and similar to the difference between 1963 and 1970 amounting to 29.8 percentage points. A t-test of both differences is highly significant. Because preferences between 1953 and 1954 can be assumed time constant, this significant difference in acceptance shows that the inclusion of a time limit is indeed responsible for higher shares of yes votes among the male population. Theoretically, some men might have radical preferences and vote against the second proposition even though they supported the first one to protest and signal dissatisfaction. However, based on the supporting evidence from past ballots that including a time limit on average increases voter support this should seldom be the case.

3.4 Data

For the estimation of the ATE, I collected a dataset of 2,143 Swiss municipalities with voting information for the relevant ballots on November 15, 1970, June 6, 1971, and December 8, 1963. Voting results include the number of yes and no votes, valid votes and eligible citizens. Data from the three cantons Aargau, Freiburg, and Ticino are not available at municipal level. Instead I include the data from voting districts which comprise several municipalities each for these three cantons adding 26 voting districts to the dataset.[19] For the canton Geneva, data are missing for the vote in 1963. Therefore, it is excluded from the analysis. All voting data come from the Political Atlas of Switzerland provided by the Swiss Statistical Office. In addition, I merge voting data with a set of demographic variables from the Swiss census of 1970 also published by the Swiss Statistical Office.

[19] I have contacted the cantonal archives of the three cantons in question. For only 20 municipalities in the canton Freiburg complete voting data required for the estimation exist, so using district data is the only way to include data from these cantons in the regressions.

Since voting data come from two ballots with a time difference of 7 months, municipal mutations need to be taken into account, because several municipalities merged during this time. Therefore, I adjust the voting data from ballot 1 such that they are comparable to ballot 2. I do the same for the census data which means that I sum the data from municipalities which have merged between 1970 and 1971. Information of municipal mergers comes from the online register of municipal mutations provided by the Swiss Statistical Office. There are two special cases in the cantons Bern and Thurgau. The first one in the canton Bern is that voting results from very small municipalities are counted and reported in some larger nearby municipality. Second, in Thurgau several municipalities which are available separately in the census data together form a political municipality with different administrative tasks. Voting data are reported for the latter only. I account for both special cases by adjusting the census data accordingly such that they are comparable.

3.5 Results: Average Treatment Effect

Table 2 shows the main regression results based on the votes in 1970 and 1971. With the exception of the fixed effects regression, the estimates are conducted using a weighted least squares estimator. Weights are proportional to the inverse of the total number of eligible voters in 1971 to account for heteroscedasticity of the standard error.[20] In all regressions I use clustered standard errors according to the 25 Swiss cantons to account for potential serial correlation of the error terms.[21] Cantons are the most natural cluster for Swiss municipalities.

All estimates are highly significant. In the baseline specification (1) with full sample and without canton fixed effects, the female acceptance rate amounts to 63.9%. Since the official voting result had an acceptance rate of 72.7%, this means that women were less likely to vote yes in the second ballot than men and the gender gap is negative. The second specification is run with a fixed effects estimator. Both first difference and fixed effects estimations lead to very similar regression coefficient. This means that violation of strict exogeneity is not a problem, and I can continue the rest of the analysis based on first-difference estimates only.

The remaining specifications include canton fixed effects. The estimated female acceptance rate slightly increases when canton fixed effects are accounted for. In other words, the gender preference gap for taxation becomes slightly smaller. In the remaining specifications (4) to (8), I restrict the sample according to either the number of Swiss adults in the municipality, or the share of women to account for potential outliers. The female acceptance rate is relatively stable over the various specifications. The most likely explanation is that estimates are weighted by the inverse of the eligible population. This way, big municipalities like Zurich or Bern which would have biased the results receive lower weights in the regression.

Table 3 reports the results when the change in voters is interacted with the percentage of Swiss

[20] Residual plots after ordinary least squares regressions reveal that the variance of the residual gets larger as $\Delta voters_m$ increases. Since this variable is negative for several observations, I preferably use the number of voters in 1971 for weighting. This variable is always positive.

[21] Note that today's 26^{th} canton Jura was still part of the bigger canton Bern in 1971.

adult women. The recovered female acceptance rate is in the first row of the table, and the estimated coefficient in the second one. Results have similar tendencies as in Table 2: including canton fixed effects results in higher estimated female acceptance rates. Again, excluding very large and very small municipalities as well as outliers regarding the share of women has only a small effect on the coefficient. However, in comparison to the baseline results when accounting for the share of women the female acceptance rate drops in all specifications (e.g., from 63.9% to 60.6% in the first specification).

For the calculation of the gender wage gap, I first calculate the male acceptance rate by taking the true voting result which has to be the gender turnout-weighted sum of the male and female acceptance rates. Taking the difference of the gender acceptance rates, yields a gender preference gap for taxation of -18.4 percentage points for baseline specification (1). For specification (3) including canton fixed effects the gender gap amounts to -12.9 percentage points, and -11.8 percentage points when excluding very large and very small municipalities in specification (6). The gender preference gap is consistently negative over all specifications. To calculate standard errors, I apply a bootstrap procedure with 1999 repetitions (standard errors reported in brackets). Based on a one-sided test, I can reject the hypothesis suggested by the literature that the gender preference gap should be positive. The alternative hypothesis of a negative gender preference gap can be accepted at significance levels of 5 and 10% for specifications (1)-(2),(7)-10), (12) and (14). In the remaining specifications the gender gap is not significantly different from zero.

At first sight, this is a surprising result. At second glance, however, there might exist several explanations for this finding. The most important one is that the ballot propositions at questions concerned government expenditure as an aggregate and did not distinguish separate spending issues. Literature examining gender preference gaps cited above suggested that women should be more likely to favor government spending on items like redistribution or health. Both are items which concern women more directly, either via lower employment rates and wages, or via the care for their children (e.g., Funk & Gathmann, 2012; Miller, 2008).

The estimation strategy exploits the institutional change in female voting rights but relies on several assumptions about male voting behavior. E.g., I assume that men turn out with the same probability as before. But potentially they altered their turnout behavior as a reaction to female voting. Since the estimated gender preference gap is negative, however, note that overestimating male participation is unproblematic as this would even underestimate the negative preference gap.

Since the analysis is based on data from Switzerland which has a strong federal structure, a substitution effect from preferences for federal spending to cantonal spending due to female suffrage is of concern. However, the results of Stutzer and Kienast (2005) suggest that cantonal spending decreased with female suffrage while federal spending increased (Abrams & Settle, 1999). This is evidence against a substitution effect. Hence, preferences captured in my analysis are not specific for federal government expenditure but for government expenditure in general.

Table 2: Estimates of the Female Acceptance Rate from Ballot Propositions in 1970 and 1971

Sample	(1) Full sample	(2) Full sample	(3) Full sample	(4) <10,000 Swiss
Female Acceptance	0.639***	0.640***	0.666***	0.667***
	(0.023)	(0.002)	(0.019)	(0.027)
Gender Preference Gap	- 0.184**	- 0.181**	-0.129	-0.128
	(0.107)	(0.086)	(0.101)	(0.112)
Canton FE	No	No	Yes	Yes
Estimation Method	FD	FE	FD	FD
Adjusted R^2	0.962	0.985	0.972	0.962
Observations	2,143	4,286	2,143	2,095

Sample	(5) >100 Swiss	(6) <10,000 and >100 Swiss	(7) <60 and >40% women	(8) <55 and >45% women
Female Acceptance	0.669***	0.672***	0.666***	0.660***
	(0.017)	(0.024)	(0.019)	(0.022)
Gender Preference Gap	-0.124	-0.118	-0.130*	-0.142*
	(0.101)	(0.112)	(0.101)	(0.111)
Canton FE Yes	Yes	Yes	Yes	Yes
Estimation Method	FD	FD	FD	FD
Adjusted R^2	0.975	0.965	0.972	0.968
Observations	1,868	1,820	2,110	1,816

Note: *** p<0.01, ** p<0.05, * p<0.1. First difference (FD) with weighted least squares or fixed effects (FE) estimates. Dependent variable is change in the number of yes votes (FD), or number of yes votes (FE). Clustered standard errors at cantonal level in brackets. Standard errors for the gender preference gap are bootstrapped with 1999 repetitions. For the fixed effects estimator, within R^2 is reported.

Table 3: Estimates of the Female Acceptance Rate from Ballot Propositions in 1970 and 1971

Sample	(9) Full sample	(10) Full sample	(11) <10,000 Swiss	(12) >100 Swiss
Female Acceptance	0.606	0.621	0.635	0.626
ΔVoters $*$ %Women	1.198***	1.228***	1.257***	1.228***
	(0.038)	(0.031)	(0.041)	(0.030)
Gender Preference Gap	-0.244*	-0.215*	-0.188	-0.204*
	(0.149)	(0.145)	(0.148)	(0.145)
Canton FE	No	Yes	Yes	Yes
Estimation Method	FD	FD	FD	FD
Adjusted R^2	0.965	0.973	0.964	0.975
Observations	2,143	2,143	2,095	1,868

Sample	(13) <10,000 and >100 Swiss	(14) <60 and >40% women	(15) <55 and >45% women
Female Acceptance	0.641	0.623	0.633
ΔVoters $*$ %Women	1.261***	1.228***	1.253***
	(0.038)	(0.031)	(0.039)
Gender Preference Gap	-0.173	-0.213*	-0.188
	(0.150)	(0.145)	(0.153)
Canton FE	Yes	Yes	Yes
Estimation Method	FD	FD	FD
Adjusted R^2	0.967	0.973	0.969
Observations	1,820	2,110	1,816

Note: *** p<0.01, ** p<0.05, * p<0.1. First difference (FD) with weighted least squares. Dependent variable is change in the number of yes votes. Clustered standard errors at cantonal level in brackets. Standard errors for the gender preference gap are bootstrapped with 1999 repetitions.

4 Direct and Indirect Gender Effect on Government Spending

4.1 Mediation Framework

The second question focuses on the underlying causes of the gender preference gap: what part of the gap is due to being female, and what part can be explained by other factors like socioeconomic differences like employment status and education that distinguish women from men? The question about the correct definition of the counterfactual arises. In the subsequent analysis, I propose a framework that allows to identify the size and causes of gender preference gaps for government spending.

Gender affects the outcome support of spending Y via two different channels. First, being female has an immediate impact on the probability of supporting larger governments which is henceforth referred to as the "direct" effect. Second, gender has a mediated effect through other variables (mediators) that thereafter affect the outcome variable which is commonly termed the "indirect" effect of the group variable. Write these mediating variables as a vector of observables M where $M_k,\ k \in \{1, ..., K\}$ denotes the kth element of the vector. $M(g)$ is a function of gender. In the jargon of the mediation literature $M(g)$ "lies on the causal path" of gender to support behavior, where gender marks the start of the causal chain (Baron & Kenny, 1986; Imai & Yamamoto, 2013). For example, employment status is an important factor influencing preferences for government spending. However, employment itself is a function of gender since women are less likely to work and more likely to stay at home to care for children than men. Thus, part of the effect of gender on support for government spending is mediated by employment status. The outcome variable has to be written as a function of both gender and the mediating variable, $Y(g, M(g))$ such that the total effect τ is:

$$\tau = E[Y(1)] - E[Y(0] = E[Y(1, M(1))] - E[Y(0, M(0))] \tag{10}$$

Further, the indirect $\eta(g)$ and direct $\delta(g)$ effects are defined as:

$$\eta(g) = E[Y(g, M(1))] - E[Y(g, M(0))] \tag{11}$$

$$\delta(g) = E[Y(1, M(g))] - E[Y(0, M(g))] \tag{12}$$

The indirect effect (11) reflects the difference in expected values when evaluating mediators for both groups while keeping gender constant at g, and (12) the direct effect which shows the effect of varying gender on the difference in expected outcomes when mediators are kept at their values for g. Note that outcome $Y(g, M(1-g))$ is never observed because each individual can only be observed in either one of the groups G but never how the individual would have acted if he was of the other gender. Decomposition of the total effect is thus based on the potential outcome framework (e.g., Rubin, 2004). Since individual treatment effects cannot be estimated, the analysis relies on population averages: the above equations (11) and (12) already denote population average indirect and direct effects. Also note, that after a simple manipulation the total effect can be written as

the sum of indirect and direct effects:

$$\begin{aligned}
\tau &= E[Y(1,M(1))] - E[Y(1,M(0))] + E[Y(1,M(0))] - E[Y(0,M(0))] \\
&= \eta(1) + \delta(0) \qquad (13) \\
&= E[Y(0,M(1))] - E[Y(0,M(0))] + E[Y(1,M(1))] - E[Y(0,M(1))] \\
&= \eta(0) + \delta(1) \qquad (14)
\end{aligned}$$

For identification, it is required that G is independent (i.e., assignment into groups G is random), and mediators M are exogenous when conditioning on G (Huber, 2014). While it can be argued that assignment into gender is random or non-manipulable by other factors (at least in the context of this paper), independence of mediators is a relatively strong assumption. Conditional on G, the error term is not allowed to impact mediators M and the outcome Y at the same time. This assumption is easily violated: for example, being of pension age affects government spending preferences but at the same time also reduces the probability of being employed which is a gender mediator as argued above.

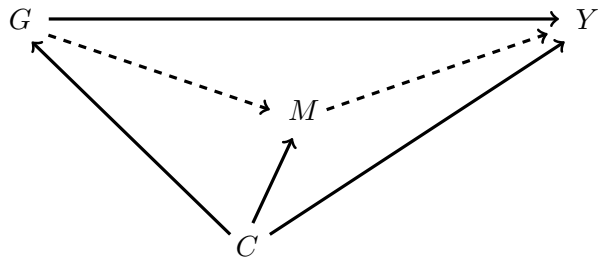
The solution is to replace the independence assumption by a set of conditional independence assumptions. $C = [C_1, ..., C_J]$, $j \in \{1., , , .J\}$ is a vector of j observables. They have a confounding effect on G, M, and Y which means that they influence some or all of the three variables. Pension age would be such a confounding factor by the above argumentation. $\eta(g)$ and $\delta(g)$ are then correctly identified under a sequential ignorability assumption (e.g., Huber, 2013; Imai, Keele & Yamamoto, 2010).

Assumption 3 (Sequential Ignorability)
1.1 $\{Y(g', m), M(g)\} \perp G|C \; \forall \; g', g \in \{0, 1\}$
1.2 $\{Y(g', m)\} \perp M|G = g, C = c \; \forall \; g', g \in \{0, 1\}$
1.3 $P(G = g|M = m, C = c) > 0 \; \forall \; g \in \{0, 1\}$

Assumption 1.1 implies that once conditioning on confounders C no other confounders exist which would either impact gender G and mediators M at the same time, or gender and outcome Y, or both. Ignorability thus means that besides confounders C all other variables can be ignored, or that all confounders must be observed. According to Assumption 1.2, after conditioning on gender G and confounders C, no variables should have an effect on the mediators and the outcome Y. In more detail, the assumption demands that how mediators impact the outcome variable is not confounded once controlling for gender and confounders. For example, if pension age was truly the only confounding variable, the effect of employment status on supporting government spending is unconfounded after conditioning on gender and pension age. The last Assumption 1.3 is a common support assumption demanding enough comparable observations for both groups $g = 1$ and $g = 0$ in order to have comparable units across both groups. There should be enough individuals in the sample that are similar regarding all mediators and confounders but differ by gender. If employment status and pension age were the only mediators and confounders, all feasible combinations of the two variable should be observed in the data for men as well as women (e.g., employed and below

pension age, unemployed and below pension age, etc.). A graphical representation of the mediation framework is depicted in Figure 3. The solid lines represent direct effects while the dashed lines visualize the indirect gender effect.

Figure 3: Mediation Framework



Note: G are groups, M are mediators, C are confounders, and Y the outcome. Solid lines represent direct effects, and dashed ones indirect or mediated effects.

4.2 Estimation Strategy: Identifying the Direct and Indirect Gender Effect

The standard way of estimating $\eta(g)$ and $\delta(g)$ is through a set of linear equations (e.g., Baron & Kenny, 1986; Judd & Kenny, 1981, or Blinder (1973) or Oaxaca (1973) for linear wage decompositions, but without confounding factors). However, linearity imposes a relatively strong functional form assumption which might be overly restrictive given that my outcome variable is bivariate (Hicks & Tingley, 2011). Huber (2013, 2014) proposes to use a nonparametric model of the following form instead:

$$M_k = \chi_k(G, C, \upsilon), \ for\ k \in 1, ..., K \tag{15}$$

$$Y = \phi(G, M, C, \epsilon) \tag{16}$$

This approach is more flexible and appropriate for my analysis since χ and ϕ are functions that do not need to be specified more precisely. Under the sequential ignorability assumptions 1.1 to 1.3, Huber (2013) shows that the direct and indirect effects are non-parametrically identified. The identification relies on a reweighing mechanism according to propensity scores $P(G = 1|M, C)$ and $P(G = 1|C)$. From this the direct and indirect effect can be calculated by using sample moments and propensity score estimates.

$$\eta = E\left(\frac{Y \cdot G}{P(G=1|C)}\right) - E\left(\frac{Y \cdot G}{P(G=1|M,C)} \cdot \frac{1 - P(G=1|M,C)}{1 - P(G=1|C)}\right) \tag{17}$$

$$\delta = E\left(\frac{Y \cdot G}{P(G=1|M,C)} \cdot \frac{1 - P(G=1|M,C)}{1 - P(G=1|C)}\right) - E\left(\frac{Y \cdot (1-G)}{1 - P(G=1|C)}\right) \tag{18}$$

For estimation, I use the normalized variants of (17) and (18) as in Huber (2013) and suggested by

Imbens (2004). Their exact form can be found in the appendix. Both propensity scores $p(M, C)$ and $p(C)$ are estimated with probit regressions.

The ATE can be estimated with a probit regression of gender on approval while conditioning on confounders which gives $\widehat{\tau}$. From this it is straightforward to calculate the mediated indirect effect for women and men:

$$\widehat{\eta}(1) = \widehat{\tau} - \widehat{\delta}(0) \tag{19}$$

$$\widehat{\eta}(0) = \widehat{\tau} - \widehat{\delta}(1) \tag{20}$$

For a detailed derivation of the estimator and its normalized version, I refer the reader to Huber (2013).

4.3 Data

Post-ballot surveys are conducted shortly after all referendum and initiative ballots at national level in Switzerland since 1981. The project is called VOXit, and the data are published by the Swiss foundation for research in social sciences.[22] Randomly chosen respondents answer a questionnaire by telephone. Among the information included are the voting behavior and various socioeconomic controls as well as contextual information. The advantage of these polls is that voters as well as eligible citizens who did not go to the polls answer the questions. Importantly, until the end of 1999 they include the hypothetical answer of the nonparticipating respondents to the question of how they would have decided if they had voted. This allows me to conduct an analysis of voters and nonvoters by gender.

The mediation analysis is based on all three votes regarding the federal financial order between 1981 and 1999. These are the ballots voted on 29 November 1981, 2 June 1991, and 28 November 1993 which is the last ballot concerning the federal financial order before 1999, and thus containing the answers of non-voters. Though tax rates and deductions have of course changed since the ballot propositions in 1971, the matter is in fact identical to the propositions analyzed above. The propositions of 1981 and 1993 include time limits for the federal financial order until 1994 (Federal Announcement 1981 II, p.561) and 2006 (Federal Announcement 2003, p.1540) respectively. The 1991 proposition does not have a time limit. While it might be of concern that women have not yet grown accustomed to their voting rights in 1971 and might have hesitated to participate, for the later ballots female voting rights were already well established. Moreover, any potentially strategic male voting behavior stemming from the introduction of female suffrage should have ceased to exist by then.

I use all observations for which the participation and voting decision are available. All observations where according to the survey the respondent submitted an empty vote are dropped in order to follow the official rule to calculate voting results. I drop observations where the respondent claims to have turned out for the vote but there is an answer for the voting behavior of non-participants

[22] Data are available online on the following homepage: http://nesstar.sidos.ch/webview/index.jsp

in the data set (and vice versa for non-participants with information on voting for that individual). These are only 11 and 17 cases respectively and most likely the result of data mistakes.

A typical concern about using surveys to elicit voter preferences is potential survey bias: either respondents misrepresent their voting behavior, or they choose not to participate in the survey conditional on their characteristics. To account for this issue, Funk (2012) proposes to compare the official voting results with the share of survey respondents claiming they have voted "yes". Subtracting the official results from the survey results based on the voting population only, yields a difference of 10.37 (1981), -1.84 (1991), and 1.64 (1993) percentage points between the two. The 1991 and 1993 values confirm Funk's (2012) result that on average no significant survey bias occurs in votes concerning federal finances. However, the first value suggests the existence of a survey bias and points to problems of accuracy with the survey results from 1981.

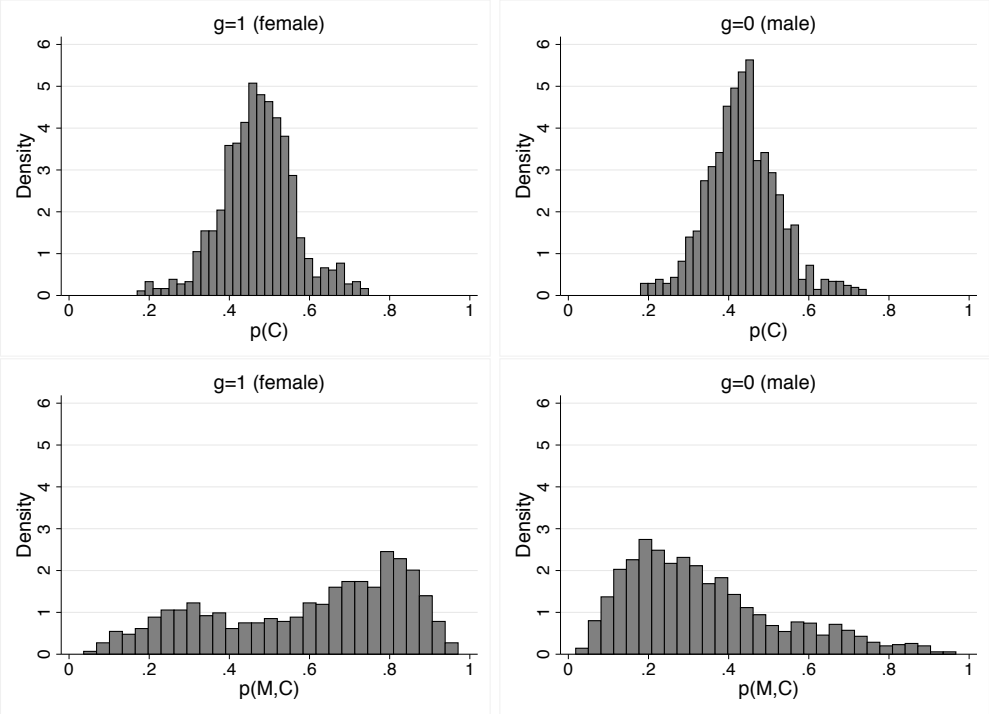
4.4 Empirical Specification

For the estimations, I use the pooled set of the responses after three different votes of both voters and non-voters who hypothetically state their voting decision. I also run regressions for each ballot separately and for combinations of two ballots. The dummy Y_i takes the value 1 if individual i voted yes or would have voted yes, and the value 0 else. The main variable of interest is the gender dummy G_i which becomes 1 for women, and 0 for men.

In addition to voting behavior, the post-ballot surveys contain a wealth of socioeconomic and contextual variables which can be used as mediators and confounders. Some of the variables are available for all three votes and are denoted by M_{basic} and C_{basic}. Some variables have only been collected for a subset of the three votes and are therefore denoted by $M_{extended}$ and $C_{extended}$ respectively. As basic mediators of gender on accepting the voting propositions I consider the following variables: *work:full-time* is a dummy if the respondent is full-time employed. A second dummy *work:part-time* denotes if the respondent has a part-time job. Being neither full-time or part-time employed means that the respondent is either unemployed, in pension, in education, or stays at home. The dummy *education:high school* becomes one if the respondent has a high-school degree or higher. The dummy *education:vocational* is one if the respondent received vocational training. Additional mediators available for a subset of the votes are the following: *life standard* is a dummy with value one if a respondent rated his or her life standard as high or medium high as compared to medium low and low (available for the votes in 1981 and 1991). Two dummies denote the respondent's occupation in case he or she is not working full-time: *work:pension* denotes if a respondent has retired, and *work:household* reflects if a respondent stays at home to keep house (available for the votes in 1991 and 1993). I also include four *income* dummies (available for the vote in 1993). They reflect household income on a scale from 0 to 4 (3,001-5,000; 5,001-7,000; 7,001-9,000; >9,000 Swiss Francs per month). The category (<3,001) is left out as reference category.

The following variables are used as confounders potentially influencing both the outcome and mediators at the same time: *age* denotes the respondent's age. The dummies *status:married* and *status:single* denote the respective marital status. If the respondent is *Catholic* this dummy takes

Figure 4: Histograms of Propensity Scores by Gender



Note: Based on data from VOX-surveys no. 161, 421, and 511. Propensity scores from probit estimates. $p(M, C)$ propensity scores include mediators and confounders, $p(C)$ only the latter.

the value one.[23] *urban* is one if the respondent lives in a city and not in a rural area. To account for cultural differences between the geographical and linguistic areas in Switzerland, I include the region dummies *region:West*, *region:Center*, *region:Center-West*, and *region:Center-East*. The Southern region is left out as reference group. For the votes in 1981 and 1991 I include dummies for the size of the respondent's household: *housesize:2* denotes a two-person household, *housesize:3,4* three or four persons, and *housesize:5* households with five or more inhabitants.

To summarize, the vector of basic gender mediators is $M_{basic} = \{$*work:full-time, work:part-time, education:high school, education:vocational*$\}$ which is extended by some of the following mediators for some of the votes $M_{extended} = \{$*life standard, work:pension, work:household, income1, income2, income3, income4*$\}$. The vector of basic confounders is $C_{basic} = \{$*age, status:married, status:single, Catholic, urban, region:West, region:Center, region:Center-West, region:Center-East*$\}$ which is extended by $C_{extended} = \{$*housesize:2, housesize:3,4, housesize:5*$\}$. Summary statistics of the variables by gender, the mean difference and t-statistic are reported in Table 4.

I provide the propensity scores regression results of $P(G = 1|M, C)$ and $P(G = 1|C)$ conducted with probit estimates in the Appendix (Tables 6 and 7). Figure 4 shows exemplary histograms of propensity scores for both men and women in the total sample with basic mediators and confounders to validate the common support assumption 1.3 requiring enough comparable observations across

[23] The majority of the Swiss population is either Roman Catholic (46.2% in 1980), or Protestant (45.3% in 1980) (data are from the website of the Swiss Statistical Office www.bfs.admin.ch).

men and women.[24] They reveal that the common support assumption is violated for specifications based on the joint sample of votes from 1991 and 1993 when adding the extended vector of mediators including dummies for retirement, house keeping, and income dummies. Including the measures for house keeping and income evidently reduces the number of comparable observations across gender because men tend to earn more on average (6.2 vs. 10.5% are in the >9,000 income category), and hardly any men stay at home for house keeping whereas a large share of women does (47.5 vs. 1.2%). For this reason, the extended vector of mediators and confounders cannot be used with the ballots from 1991 and 1993.

Table 4: Descriptives of Post-Ballot Surveys by Gender

Variable	Mean (women)	Mean (men)	Difference	t-statistic	p-value
yes	0.566	0.622	-0.056	2.555	0.011
MEDIATORS					
work:full-time	0.288	0.688	-0.401	19.532	0.000
work:part-time	0.249	0.054	0.195	-12.984	0.000
education:high school	0.222	0.308	-0.086	4.369	0.000
education:vocational	0.496	0.528	-0.032	1.445	0.149
life standard	0.360	0.389	-0.029	1.033	0.302
work:pension	0.141	0.201	-0.060	3.039	0.002
work:household	0.475	0.012	0.463	-25.795	0.000
income:1	0.342	0.332	0.011	-0.300	0.764
income:2	0.227	0.242	-0.015	0.472	0.637
income:3	0.142	0.174	-0.032	1.176	0.240
income:4	0.062	0.105	-0.043	2.084	0.038
CONFOUNDERS					
age	44.973	47.196	-2.223	2.932	0.003
status:married	0.629	0.631	-0.002	0.072	0.943
status:single	0.190	0.256	-0.066	3.523	0.000
Catholic	0.419	0.448	-0.029	1.296	0.195
urban	0.622	0.578	0.044	-2.019	0.044
region:West	0.221	0.214	0.007	-0.355	0.723
region:Center	0.251	0.257	-0.006	0.312	0.755
region:Center-West	0.254	0.243	0.011	-0.547	0.584
region:Center-East	0.255	0.250	0.004	-0.228	0.819
housesize:2	0.277	0.343	-0.066	2.4389	0.0149
housesize:3	0.433	0.369	0.064	-2.2355	0.0256
housesize:4	0.114	0.117	-0.003	0.1684	0.8663

Note: T-tests based on data from VOX-surveys no. 161, 421, and 511. Data are available online on http://nesstar.sidos.ch/webview/index.jsp.

[24] Propensity score histograms by gender for all other main specifications are attached in the Appendix.

4.5 Results: Direct and Indirect Effects

The main mediation analysis results based on individual post-ballot responses from the votes in 1981, 1991, and 1993 are reported in Table 5. The first column shows the results for the complete sample, while in specifications (2)-(8) the results are reported for a subset of the three ballots. With the exception of columns (6) and (8) I use the set of mediators and confounders available for all three ballots (M_{basic} and C_{basic}). The estimates are based on all survey respondents. The sample is reduced to respondents who reported to vote in column (9).

The total effect τ in the first row is the marginal effect from a probit regression of acceptance on gender while controlling for all confounders. Being a woman decreases the probability of voting yes by 5.4 percentage points in the total sample, and varies between -3.2 and -9.7 percentage points in subsamples. The estimates are significant at conventional significance levels in most specifications. Only in the subsamples (2) and (5) using data exclusively from the 1981 ballot (as well as in combination with the 1991 ballot) and from reported voters in (9), there is no significant difference in the acceptance behavior of men and women. Throughout all specifications, a positive gender gap can be rejected at the 1% significance level. Comparing the total effect to the results from the 1970/1971 referendums in the previous section, suggests qualitatively similar results. But the size of the gender preference gap is considerably reduced. There are several explanations. First, the post-ballot surveys are based on data from votes that took place at a later point in time. Any kind of strategic voting incentives or hesitation to vote which might have been present in 1970/1971, have most likely disappeared. Second, though similar, the ballot propositions are not identical which also might explain part of the variation over time. Further, there is a survey bias in the 1981 survey suggesting problems with the representativeness of the data from that year.

The direct and indirect effects evaluated for women, $\delta(1)$ and $\eta(1)$, are in the second and third rows of the results table. The respective values for men are reported in the last two rows for completeness. Standard errors are computed using a bootstrap procedure with 1999 iterations. Again, with the exception of estimates including the survey from 1981 in (2), (5) and (6), all marginal effects are highly significant. The direct effect of being female on the probability of voting yes in the referendum is negative in all estimations. Thus, when evaluating mediators at their expected values for women, being female reduces the probability of accepting the fiscal financial order. The effect is seizable: in the significant specifications it varies between -9.4 percentage points in the complete sample and -13.2 percentage points.

Also the mediated effect, i.e. the gender difference in voting behavior induced by socioeconomic gender differences, is negative but only significant in three out of nine specifications. The mediated effect amounts to -4.1 percentage points in the complete sample (1), and is -8.1 and -10.2 percentage points in the other two significant specifications.

For robustness, I account for potentially serially correlated standard errors within ballots, and rerun specifications (1), (6) and (7) with ballot clustered standard errors (cf. Table 8 in the Appendix). The direct gender effect now turns significant in all specifications while the indirect effect remains unaffected.

Table 5: Direct and Indirect Effects

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Total effect:	τ	-0.054**	-0.032	-0.097***	-0.060*	-0.064	-0.062**	-0.082**	-0.082**	-0.043
		(0.022)	(0.038)	(0.037)	(0.033)	(0.029)	(0.029)	(0.034)	(0.034)	(0.027)
Women										
Direct effect:	$\delta(1)$	-0.094***	-0.018	-0.102**	-0.112***	-0.061	-0.066	-0.132***	-0.129***	-0.106***
		(0.029)	(0.063)	(0.052)	(0.039)	(0.042)	(0.045)	(0.041)	(0.043)	(0.034)
Indirect effect:	$\eta(1)$	-0.041**	-0.024	-0.102**	-0.002	-0.017	-0.081***	0.001	-0.001	-0.016
		(0.020)	(0.053)	(0.045)	(0.028)	(0.029)	(0.030)	(0.031)	(0.034)	(0.027)
Men										
Direct effect:	$\delta(0)$	-0.013	-0.008	0.005	-0.057	-0.047	0.019	-0.082	-0.081*	-0.027
		(0.029)	(0.064)	(0.059)	(0.043)	(0.040)	(0.040)	(0.045)	(0.049)	(0.038)
Indirect effect:	$\eta(0)$	0.040*	-0.014	0.006	0.053**	-0.003	0.003	0.050**	0.048	0.063***
		(0.021)	(0.051)	(0.041)	(0.026)	(0.031)	(0.034)	(0.028)	(0.029)	(0.022)
Mediators, Confounders		basic	basic	basic	basic	basic	extended	basic	extended	basic
Ballots		all	1981	1991	1993	1981,1991	1981,1991	1993	1993	all
Observations		2,018	535	686	797	1,190	1,190	719	719	1,258

Note: *** p<0.01, ** p<0.05, * p<0.1. Based on data from VOX-surveys no. 161, 421, and 511. Inverse propensity score weighted results. The binary dependent variable is 1 if the respondent voted yes, 0 if no. Standard errors in brackets. Standard errors for direct (δ) and indirect effects (η) are based on 1999 bootstrap iterations. Standard errors of total effect τ are from probit estimates. Specifications (1) to (5), (7) and (9) include mediators and confounders available for all three votes. In columns (6) and (8) the extended set of mediators and confounder is used. In (5) and (7) the number of observations is reduced to those for which the extended set is available. Column (9) is based on respondents who reported to vote only.

Observations from the surveys are not independently drawn in so far that respondents live in various cantons. Cantons are not reported for 195 observations. First, I rerun the estimates (1) to (8) with the reduced set of observations for which information on cantons is available (reported in Table 9 in the Appendix). I also exclude the regional dummies from the set of confounders because they correspond to groups of cantons. That way, I make sure that the significance of the results is not affected by the reduced sample size. Indeed, coefficients of direct and indirect effects change only slightly. Second, I cluster the error terms at cantonal level (reported in Table 10). Even though significance is slightly reduced as compared to coefficients without canton clusters, results remain qualitatively and quantitatively similar. Hence, accounting for the cantonal structure of the data does not have a big effect on the results.

The size and significance of the direct effect are considerable and robust, whereas there is little evidence for mediated effects through employment or education. A large part of the effect of being female on preferences for government spending runs through intrinsically female factors, and directly affects voting behavior.

In sum, evidence points to the importance of direct gender effects when explaining the gender gap in acceptance of the fiscal order. Though the effect is called direct, it not only refers to observable differences between men and women. Potential interpretations of the strong direct gender effect are unobserved mediators. Research based on experimental techniques examines gender gaps other than socioeconomic differences which might explain why women could have different preferences for government spending than men (cf. Croson and Gneezy (2009) as well as Shapiro and Mahajan (1986) for literature reviews). Literature documents that women are more risk averse (e.g., Holt & Laurya, 2002, 2005; Schubert et al., 1999) and dislike competition (Gneezy, Niederle & Rustichini, 2003; Niederle & Vesterlund, 2007). Experimental evidence suggests that women are more altruistic, and dislike inequality (Andreoni & Vesterlund, 2001; Selten & Ockenfels, 1998). This evidence might partly reflect the strong direct gender effect.

5 Concluding Remarks

The aim of this paper is to provide direct evidence for gender preferences for government expenditurse from ballot analysis. This method is preferable to analyzing indirect links between the electorate, politicians, and their subsequent choice of budgets and policies since the relation between preferences and subsequent voting behavior is much clearer. The analysis of the main effect is based on aggregate voting data around the introduction of female suffrage in Switzerland such that individual voting behavior remains unobserved. However, I argue extensively that my preference measures are likely to reflect gender preference gaps, and also provide evidence from post-ballot surveys of comparable votes.

I find that approval for government spending is lower among women than men. My findings contrast with the standard interpretation that female preferences for large governments led to an increase in government spending after their enfranchisement. Most likely, the timing of female

voting and the takeoff of government spending are correlated. But potentially the causation runs through other factors like policies fostered towards the enlarged electorate, or the election of more spending-friendly politicians.

While the results seem surprising at first, they are compatible with other results from literature pointing towards the existence of gender preference gaps but only for certain spending categories. The scope of government might matter more in the analysis of gender preference gaps than the size of government itself. For example, Abrams and Settle (1999) find particularly strong effects of female suffrage on welfare spending in Switzerland. Similarly, Aidt et al. (2006) show increases in health, welfare, and education spending which are categories typically relevant to women. Also Funk and Gathmann (2012) discover gender preference gaps from individual data for health, environmental issues, defense and welfare spending.

This paper also shows that most of the negative gender preference gap is due to intrinsic or unobserved gender differences. Socioeconomic gender gaps in employment or education play a minor role in explaining why men and women vote differently on fiscal issues. Nevertheless, they should be taken into account in similar analyses to correctly attribute the causes of preference gaps to gender and to other observable differences.

References

Abrams, Burton A. and Russell F. Settle, "Women's Suffrage and the Growth of the Welfare State," *Public Choice*, 1999, *100* (3/4), 289–300.

Aidt, Toke S. and Bianca Dallal, "Female voting power: The contribution of women's suffrage to the growth of social spending in Western Europe (1869-1960)," *Public Choice*, 2008, *134* (3/4), 391–417.

_ **, Jayasri Dutta, and Elena Loukoianova**, "Democracy Comes to Europe: Franchise Extension and Fiscal Outcomes 1830-1938," *European Economic Review*, February 2006, *50* (2), 249–283.

Andreoni, James and Lise Vesterlund, "Which Is the Fair Sex? Gender Differences in Altruism," *Quarterly Journal of Economics*, 2001, (February), 293–312.

Baron, Reuben M. and David A. Kenny, "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 1986, *51* (6), 1173–1182.

Becker, Gary, "A Theory of Marriage," in Theodore W. Schultz, ed., *Economics of the Family: Marriage, Children, and Human Capital*, Chicago: University of Chicago Press, 1974, pp. 299 – 351.

Bertocchi, Graziella, "The Enfranchisement of Women and the Welfare State," *European Economic Review*, May 2011, *55* (4), 535–553.

Blinder, Alan S., "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, 1973, *8*, 436–455.

Bolliger, Christian, "Die bürgerlichen Seiten eine spürare Steuerentlastung durch," in Wolf Linder, Christian Bolliger, and Yvan Rielle, eds., *Handbuch der eidgenössischen Volksabstimmungen 1848-2007*, Bern: Haupt, 2010, pp. 261 – 262.

Croson, Rachel and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature*, May 2009, *47* (2), 448–474.

Edlund, Lena and Rohini Pande, "Why Have Women Become Left-Wing? The Political Gender Gap and the Decline in Marriage," *Quarterly Journal of Economics*, 2002, (August), 917–961.

Funk, Patricia, "How Accurate are Surveyed Preferences for Public Policies? Evidence from a Unique Institutional Setup," 2012. Barcelona GSE Working Paper 657.

_ **and Christina Gathmann**, "Gender Gaps in Policy Making: Evidence from Direct Democracy," 2012.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini, "Performance in Competitive Environments: Gender Differences," *Quarterly Journal of Economics*, 2003, (August), 1049–1074.

Grütter, Alfred, *Die Eidgenössische Wehrsteuer, ihre Entwicklung und Bedeutung*, Zürich: Juris Druck + Verlag, 1968.

Hicks, Raymond and Dustin Tingley, "Causal Mediation Analysis," *Stata Journal*, 2011, *11* (4), 1–15.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, December 2002, *92* (5), 1644–1655.

_ **and** _ , "Risk Aversion and Incentive Effects : New Data Without Order Effects," *American Economic Review*, 2005, *95* (3), 902–904.

Huber, Martin, "Identifying Causal Mechanisms (Primarily) Based on Inverse Probability Weighting," *Journal of Applied Econometrics*, 2013.

_ , "Causal Pitfalls in the Decomposition of Wage Gaps," *Journal of Business and Economic Statistics, Forthcoming*, 2014, pp. 1–25.

Husted, Thomas A. and Lawrence W. Kenny, "The Effect of the Expansion of the Voting Franchise on the Size of Government," 1997, *105* (1), 54–82.

Imai, Kosuke and Teppei Yamamoto, "Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments," *Political Analysis*, 2013, *21*, 141–171.

_ **, Luke Keele, and Teppei Yamamoto**, "Identification, Inferences and Sensitivity Analysis for Causal Mediation Effects," *Statistical Science*, 2010, *25*, 51–71.

Imbens, Guido W., "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Eeview," *Review of Economics and Statistics*, 2004, *86* (1), 4–29.

Judd, Charles M. and David A. Kenny, "Process Analysis: Estimating Mediation in Treatment Evaluations," *Evaluation Review*, 1981, *5*, 602–619.

Linder, Wolf, "Direct Democracy," in Ulrich Kloeti, Peter Knoepfel, Hanspeter Kriesi, Wolf Linder, Yannis Papadopoulos, and Pascal Sciarini, eds., *Handbook of Swiss Politics*, Zürich: Neue Zürcher Zeitung, 2007, pp. 101 – 120.

Lott, John R. and Lawrence W. Kenny, "Did Women's Suffrage Change the Size and Scope of Government?," *Journal of Political Economy*, 1999, *107* (6), 1163–1198.

Meltzer, Allan H. and Scott F. Richard, "A Rational Theory of the Size of Government," *Journal of Political Economy*, 1981, *89* (5), 914–927.

Miller, Grant, "Women's Suffrage, Political Responsiveness, and Child Survival in American History," *Quarterly Journal of Economics*, 2008, (August), 1287–1327.

Niederle, Muriel and Lise Vesterlund, "Do Women Shy Away from Competition? Do Men Compete Too Much?," *Quarterly Journal of Economics*, 2007, (August), 1067–1101.

Oaxaca, Ronald, "Male-Female Wage Differences in Urban Labour Markets," *International Economic Review*, 1973, *14*, 693–709.

Oechslin, Hanspeter, *Die Entwicklung des Bundessteuersystems der Schweiz von 1848 bis 1966*, Einsiedeln: Etzel-Druck AG, 1967.

Rubin, Donald B., "Direct and Indirect Causal Effects Via Potential Outcomes," *Scandinavian Journal of Statistics*, 2004, *31*, 161–170.

Ruckstuhl, Lotti, *Frauen sprengen Fesseln*, Bonstetten: Interfeminas, 1986.

Schubert, Renate, Martin Brown, Matthias Gysler, and Hans Wolfgang Brachinger, "Financial Decision-Making: Are Women Really More Risk-Averse?," *American Economic Review*, 1999, *89* (2), 381–385.

Selten, Reinhard and Axel Ockenfels, "An Experimental Solidarity Game," *Journal of Economic Behavior & Organization*, March 1998, *34* (4), 517–539.

Shapiro, Robert Y. and Harpreet Mahajan, "Gender Differences in Policy Preferences: A Summary of Trends from the 1960s to the 1980s," *Public Opinion Quarterly*, 1986, *50* (1), 42–61.

Stutzer, Alois and Lukas Kienast, "Demokratische Beteiligung und Staatsausgaben: Die Auswirkungen des Frauenstimmrechts," *Swiss Journal of Economics and Statistics,*, 2005, *141* (4), 617–650.

Swiss Statistical Office, *Finanzen und Steuern von Bund, Kantonen und Gemeinden 1971*, Bern: Eidgenössisches Statistisches Amt Publikationsdienst, 1973.

_ , *Öffentliche Finanzen der Schweiz 1972*, Bern: Eidgenössisches Statistisches Amt Publikationsdienst, 1974.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data: Second Edition*, Cambridge, MA: MIT Press, 2010.

Appendix

A Federal Announcements / Bundesblätter

The federal announcements are accessible online via *http : //www.amts druckschriften.bar.admin.ch.*

- Federal Announcement 1962 I, pp. 997-1014. Botschaft des Bundesrates an die Bundesversammlung über die Weiterführung der Finanzordnung des Bundes.
- Federal Announcement 1969 II, pp. 749-807. Botschaft des Bundesrates and die Bundesversammlung über die Änderung der Finanzordnung des Bundes.
- Federal Announcement 1970 II, pp. 1-5. Bundesbeschluss über die Änderung der Finanzordnung des Bundes.
- Federal Announcement 1970 II, pp. 1581-1608. Botschaft des Bundesrates an die Bundesversammlung über die Weiterführung der Finanzordnung des Bundes.
- Federal Announcement 1971 I, pp. 486-491. Bundesbeschluss über die Weiterführung der Finanzordnung des Bundes.
- Federal Announcement 2003 I, pp. 1531-1565. Botschaft über die neue Finanzordnung.
- Information about mutations of the municipalities are taken from the historical municipality register of the Swiss Statistical Office available online *http : //www.bfs.admin.ch/bfs/ portal/de/index/infothek/ nomenklaturen/blank/blank/ gem_liste/*02*.html*
- The Année Politique Suisse (2012) is accessible online (*http : //www. anneepolitique.ch/de/aps − online.php*) and provides additional background information on ballots.
- Number of voters for cantonal votes is available online from the Centre for research on direct democracy on www.c2d.ch.
- Information about municipalities counting votes together in the canton Bern, and political municipalities in the canton Thurgau were received by email from the Swiss Statistical Office. They are available on request.
- Data used from Swiss census (1970): total population
- Voting data are from the Political Atlas of Switzerland of the Swiss Statistical Office. They were retrieved for the following ballots:

- Bundesbeschluss vom 27.09.1963 über die Weiterführung der Finanzordnung des Bundes (Verlängerung der Geltungsdauer von Art.41ter BV und Ermässigung der Wehrsteuer). Ballot on 8 December 1963.
- Bundesbeschluss vom 24.06.1970 über die Änderung der Finanzordnung des Bundes. Ballot on 15 November 1970.
- Bundesbeschluss vom 09.10.1970 über die Einführung des Frauen- stimm- und Wahlrechts in eidgenössischen Angelegenheiten. Ballot on 7 February 1971.
- Bundesbeschluss vom 11.03.1971 über die Weiterführung der Finanzordnung des Bundes. Ballot on 6 June 1971.

B Estimators of Direct and Indirect Effects

To estimate the direct and indirect effects, I estimate normalized versions of (17) and (18) (Huber, 2013). For the normalization weights are adjusted such that they add up to one for both men and women. For simplification, write $p(M,C) \equiv P(G=1|M,C)$ and $p(C) \equiv P(G=1|C)$, with their their estimated counterparts $\widehat{p}(M,C)$ and $\widehat{p}(C)$. Let i denote the index for each of the N observations. Then the direct effect evaluated at $g=1$ and $g=0$ respectively is identified by the following equations:

$$\widehat{\delta}(1) = \left[\sum_{i=1}^{N} Y_i W_i^1\right]\left[\sum_{i=1}^{N} Y_i W_i^1\right]^{-1} - \left[\sum_{i=1}^{N} Y_i W_i^2\right]\left[\sum_{i=1}^{N} Y_i W_i^2\right]^{-1} \tag{21}$$

$$\widehat{\delta}(0) = \left[\sum_{i=1}^{N} Y_i W_i^3\right]\left[\sum_{i=1}^{N} Y_i W_i^3\right]^{-1} - \left[\sum_{i=1}^{N} Y_i W_i^4\right]\left[\sum_{i=1}^{N} Y_i W_i^4\right]^{-1} \tag{22}$$

The four weights are defined as:

$$\begin{aligned} W_i^1 &\equiv \frac{G_i}{\widehat{p}(C_i)} \\ W_i^2 &\equiv \frac{(1-G_i)\widehat{p}(M_i,C_i)}{(1-\widehat{p}(M_i,C_i))\widehat{p}(C_i)} \\ W_i^3 &\equiv \frac{G_i(1-\widehat{p}(M_i,C_i))}{\widehat{p}(M_i,C_i)(1-\widehat{p}(C_i))} \\ W_i^4 &\equiv \frac{1-G_i}{1-\widehat{p}(C_i)} \end{aligned}$$

$\widehat{p}(M_i,C_i)$ and $\widehat{p}(C_i)$ are estimated with probit regressions, and the rest of the estimator is based on sample moments from which it is straightforward to calculate both $\widehat{\delta}(1)$ and $\widehat{\delta}(0)$.

C Propensity Score Histograms

Propensity Score Histograms of Table 5, No Clustering

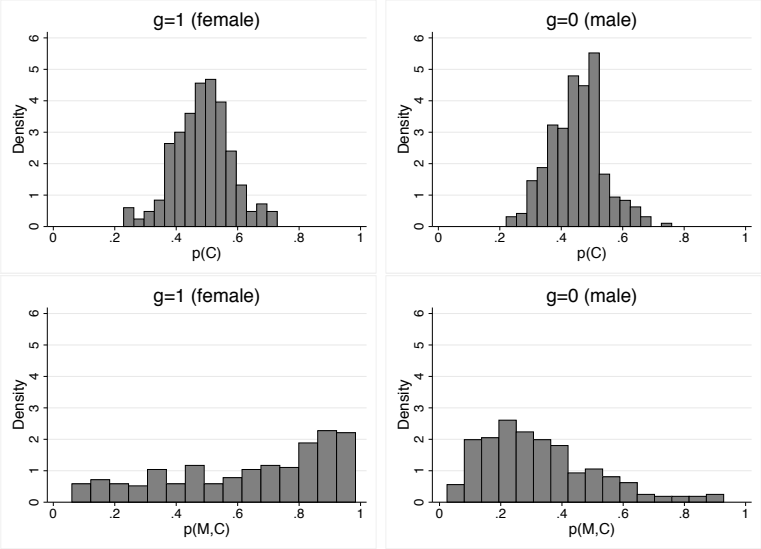


Figure 5: Histograms of Propensity Scores (M_{basic} and C_{basic}, Vote 1981)

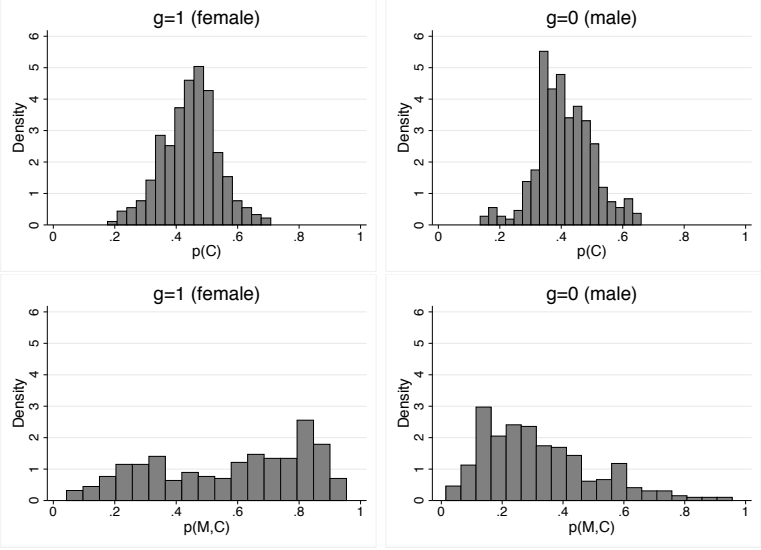


Figure 6: Histograms of Propensity Scores (M_{basic} and C_{basic}, Vote 1991)

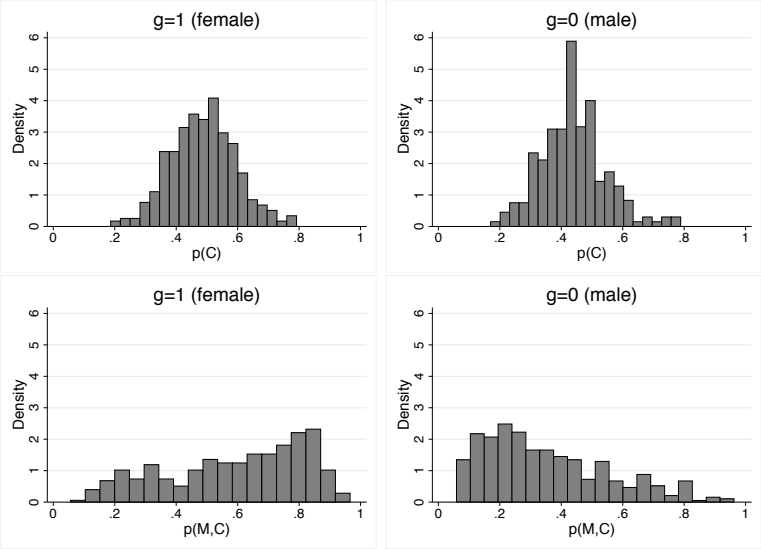


Figure 7: Histograms of Propensity Scores (M_{basic} and C_{basic}, Vote 1993)

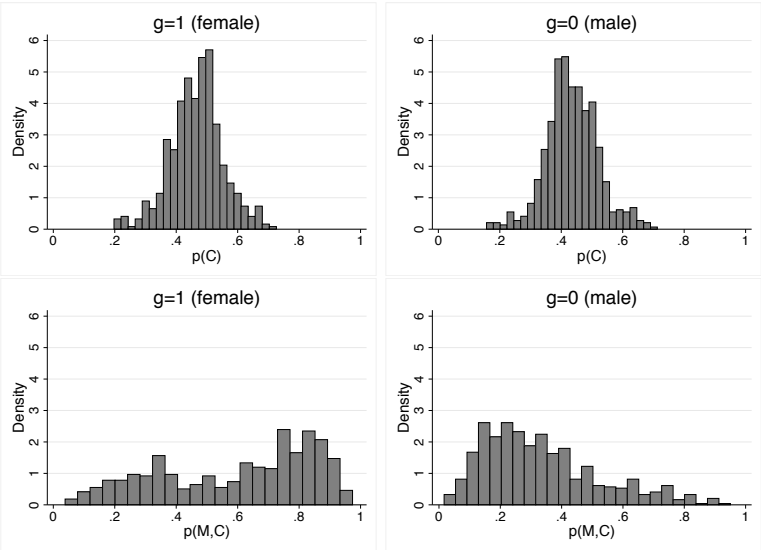


Figure 8: Histograms of Propensity Scores (M_{basic} and C_{basic}, Votes 1981, 1991)

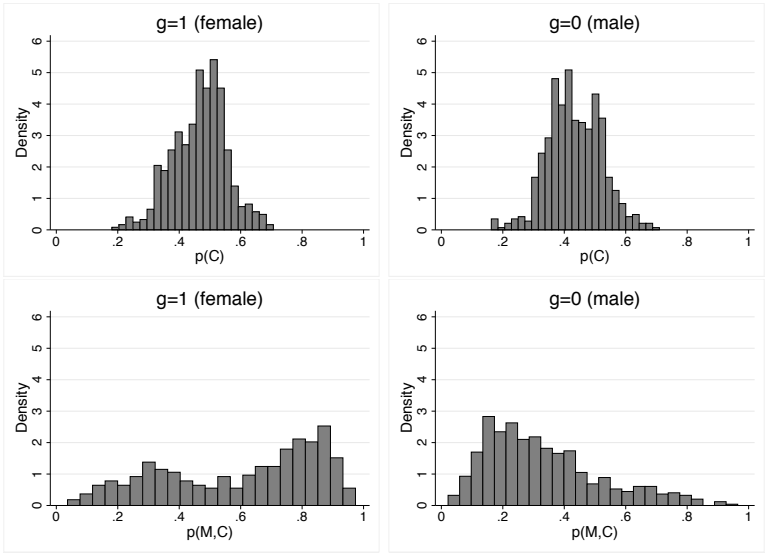


Figure 9: Histograms of Propensity Scores ($M_{extended}$ and $C_{extended}$, Votes 1981, 1991)

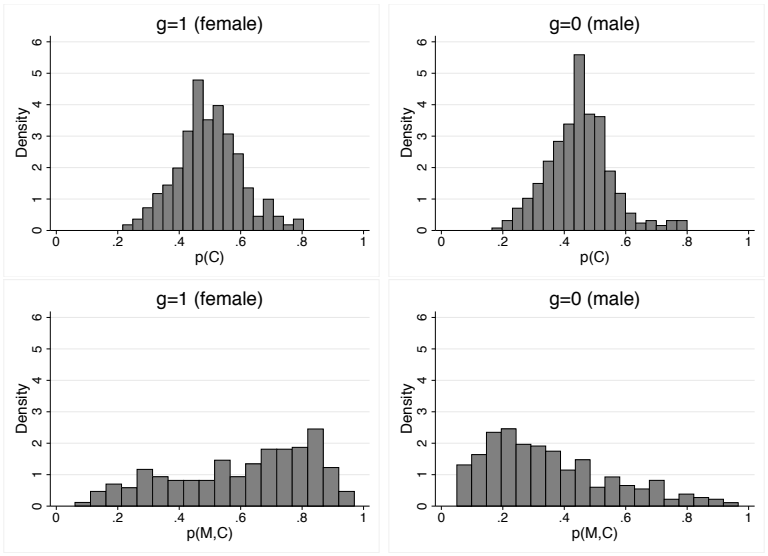


Figure 10: Histograms of Propensity Scores (M_{basic} and C_{basic}, Vote 1993)

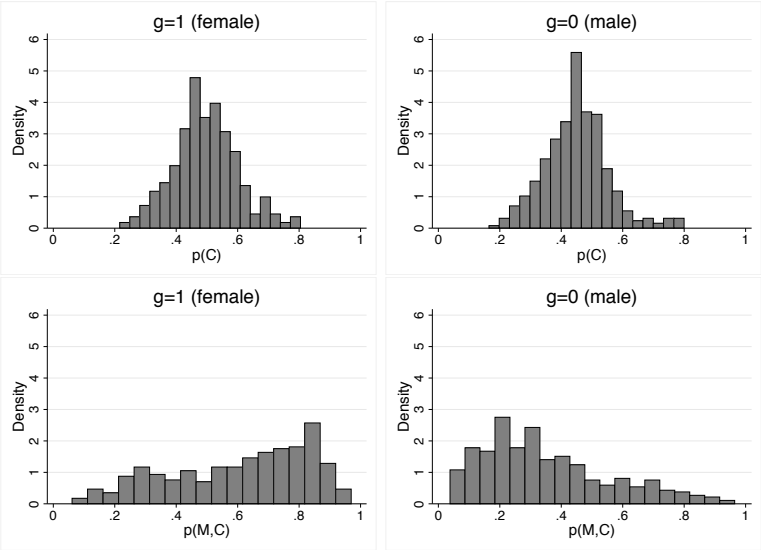


Figure 11: Histograms of Propensity Scores ($M_{extended}$ and $C_{extended}$, Vote 1993)

Propensity Score Histograms of Votes 1991 and 1993 (not used in the paper), No Clustering

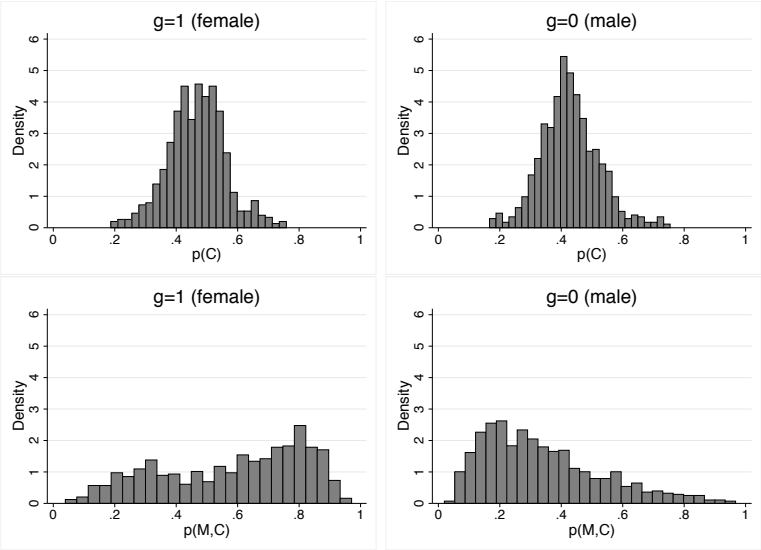


Figure 12: Histograms of Propensity Scores (M_{basic} and C_{basic}, Votes 1991, 1993)

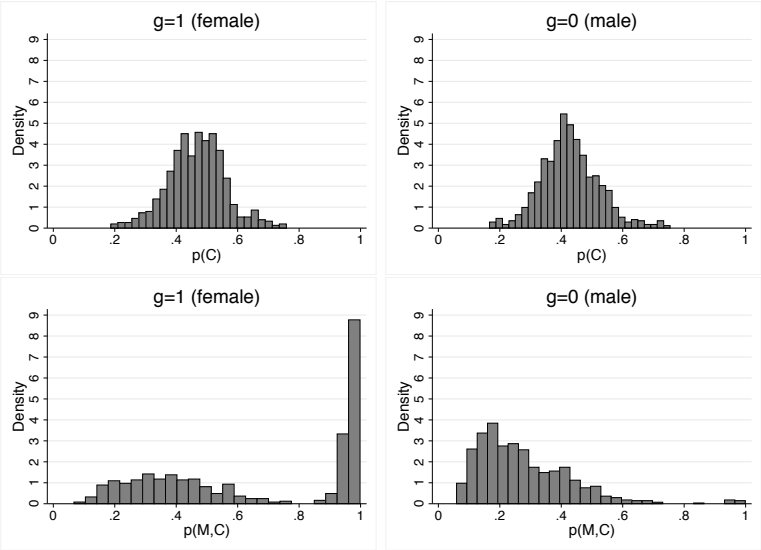


Figure 13: Histograms of Propensity Scores ($M_{extended}$ and $C_{extended}$, Votes 1991, 1993)

D Propensity Score Estimates

Table 6: Propensity Score Estimates with M_{basic} and C_{basic}

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MEDIATORS								
work:		-1.208***		-1.338***		-1.097***		-1.045***
full-time		(0.076)		(0.104)		(0.087)		(0.125)
work:		0.328***		0.275*		0.420***		0.447***
part-time		(0.104)		(0.145)		(0.116)		(0.162)
education:		-0.699***		-0.739***		-0.735***		-0.709***
high school		(0.094)		(0.127)		(0.113)		(0.162)
education:		-0.341***		-0.361***		-0.323***		-0.337**
vocational		(0.083)		(0.104)		(0.102)		(0.150)
CONFOUNDERS								
status:married	-0.400***	-0.409***	-0.386***	-0.418***	-0.407***	-0.448***	-0.481***	-0.478***
	(0.085)	(0.090)	(0.112)	(0.121)	(0.100)	(0.105)	(0.143)	(0.150)
status:single	-0.792***	-0.747***	-0.764***	-0.714***	-0.837***	-0.804***	-0.863***	-0.776***
	(0.109)	(0.115)	(0.144)	(0.153)	(0.128)	(0.133)	(0.183)	(0.190)
Catholic	-0.031	-0.033	-0.045	-0.090	-0.020	-0.000	0.078	0.143
	(0.061)	(0.066)	(0.079)	(0.086)	(0.071)	(0.077)	(0.102)	(0.112)
region:West	0.414**	0.244	0.471*	0.295	0.384**	0.249	0.255	0.116
	(0.190)	(0.203)	(0.282)	(0.313)	(0.195)	(0.208)	(0.279)	(0.293)
region:Center	0.418**	0.292	0.573**	0.392	0.357*	0.281	0.156	0.102
	(0.189)	(0.201)	(0.282)	(0.312)	(0.193)	(0.205)	(0.277)	(0.290)
region:Center	0.402**	0.237	0.467*	0.333	0.416**	0.235	0.309	0.162
-West	(0.190)	(0.202)	(0.282)	(0.312)	(0.194)	(0.206)	(0.279)	(0.292)
region:Center	0.412**	0.289	0.485*	0.356	0.400**	0.301	0.319	0.178
-East	(0.189)	(0.200)	(0.281)	(0.311)	(0.193)	(0.204)	(0.278)	(0.290)
age	-0.012***	-0.027***	-0.010***	-0.030***	-0.013***	-0.025***	-0.015***	-0.025***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)
urban	0.112*	0.137**	0.073	0.053	0.066	0.130	0.086	0.197*
	(0.061)	(0.067)	(0.079)	(0.086)	(0.073)	(0.080)	(0.106)	(0.116)
constant	0.401*	2.149***	0.232	2.379***	0.487*	1.994***	0.783**	2.041***
	(0.234)	(0.280)	(0.338)	(0.412)	(0.255)	(0.308)	(0.358)	(0.427)
Adjusted R^2	0.026	0.199	0.023	0.209	0.028	0.192	0.034	0.199
Observations	2018	2018	1190	1190	1483	1483	719	719
Ballots	all	all	'81, '91	'81, '91	'91, '93	'91, '93	'93	'93

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in brackets. Based on data from VOX-surveys no. 161, 421, and 511. Propensity score estimates based on vectors M_{basic} and C_{basic} exclusively. Dependent variable is dummy being 1 for female, 0 for men. Columns (1), (3), (5) and (7) are $P(G = 1|C)$. Columns (2), (4), (6), and (8) are $P(G = 1|M, C)$.

Table 7: Propensity Score Estimates with $M_{extended}$ and $C_{extended}$

	(1)	(2)	(3)	(4)	(5)	(6)
MEDIATORS						
work:full-time		-1.361***		-0.376***		-0.218
		(0.106)		(0.141)		(0.202)
work:part-time		0.265*		0.112		0.092
		(0.147)		(0.153)		(0.210)
education:high school		-0.810***		-0.624***		-0.425**
		(0.133)		(0.124)		(0.190)
education:vocational		-0.385***		-0.352***		-0.227
		(0.106)		(0.112)		(0.173)
life standard		0.161*				
		(0.089)				
work:pension				0.105		0.107
				(0.204)		(0.312)
work:household				2.308***		2.367***
				(0.181)		(0.243)
income:1 (3.001-5.000)						-0.268
						(0.173)
income:2 (5.001-7.000)						-0.454**
						(0.199)
income:3 (7.001-9.000)						-0.370*
						(0.222)
income:4 (>9.001)						-0.576**
						(0.257)
CONFOUNDERS						
status:married	-0.365***	-0.403***	-0.407***	-0.720***	-0.481***	-0.688***
	(0.133)	(0.142)	(0.100)	(0.112)	(0.143)	(0.169)
status:single	-0.732***	-0.668***	-0.837***	-0.438***	-0.863***	-0.265
	(0.146)	(0.155)	(0.128)	(0.138)	(0.183)	(0.200)
Catholic	-0.042	-0.087	-0.020	-0.076	0.078	-0.008
	(0.079)	(0.087)	(0.071)	(0.084)	(0.102)	(0.125)
region:West	0.466*	0.327	0.384**	0.198	0.255	0.057
	(0.283)	(0.315)	(0.195)	(0.217)	(0.279)	(0.313)
region:Center	0.571**	0.415	0.357*	0.169	0.156	0.008
	(0.282)	(0.313)	(0.193)	(0.214)	(0.277)	(0.310)
region:Center-West	0.465*	0.356	0.416**	0.150	0.309	0.163
	(0.283)	(0.313)	(0.194)	(0.216)	(0.279)	(0.312)
region:Center-East	0.485*	0.381	0.400**	0.138	0.319	-0.017
	(0.281)	(0.312)	(0.193)	(0.214)	(0.278)	(0.312)
age	-0.009***	-0.029***	-0.013***	-0.012***	-0.015***	-0.011*
	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.006)
urban	0.076	0.061	0.066	0.227***	0.086	0.309**
	(0.079)	(0.087)	(0.073)	(0.087)	(0.106)	(0.131)
housesize:2	-0.113	-0.139				
	(0.133)	(0.140)				
housesize:3,4	0.033	0.008				
	(0.142)	(0.150)				
housesize:5	-0.053	-0.015				
	(0.171)	(0.185)				
constant	0.184	2.309***	0.487*	0.804**	0.783**	0.809*
	(0.355)	(0.429)	(0.255)	(0.335)	(0.358)	(0.478)
Adjusted R^2	0.024	0.212	0.028	0.322	0.034	0.349
Observations	1190	1190	1483	1483	719	719
Ballots	1981, 1991	1981, 1991	1991, 1993	1991, 1993	1993	1993

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in brackets. Based on VOX-surveys no. 161, 421, and 511. Propensity score estimates based on vectors $M_{extended}$ and $C_{extended}$ and vectors M_{basic} and C_{basic}. Dependent variable is dummy being 1 for female, 0 for men. Columns (1), (3), and (5) are $P(G = 1|C)$. Columns (2), (4), and (6) are $P(G = 1|M, C)$.

E Direct and Indirect Effects

Table 8: Direct and Indirect Effects: Ballot Clustering

		(1)	(2)	(3)
Total effect:	τ	-0.054***	-0.064	-0.062
		(0.014)	(0.043)	(0.041)
Women				
Direct effect:	$\delta(1)$	-0.094***	-0.061**	-0.066***
		(0.026)	(0.030)	(0.025)
Indirect effect:	$\eta(1)$	-0.041	-0.017	-0.081**
		(0.025)	(0.032)	(0.035)
Men				
Direct effect:	$\delta(0)$	-0.013	-0.047***	0.019
		(0.021)	(0.011)	(0.015)
Indirect effect:	$\eta(0)$	0.040**	-0.003	0.003
		(0.020)	(0.008)	(0.012)
Mediators, Confounders		basic	basic	extended
Ballots		all	1981,1991	1981,1991
Observations		2,018	1,190	1,190

Note: *** p<0.01, ** p<0.05, * p<0.1. Based on data from VOX-surveys no. 161, 421, and 511. Inverse propensity score weighted results. The binary dependent variable is 1 if the respondent voted yes, 0 if no. Ballot clustered standard errors in brackets. Standard errors for direct (δ) and indirect effects (η) are based on 1999 bootstrap iterations. Standard errors of total effect τ are from probit estimates. Specifications (1) and (2) include mediators and confounders available for all three votes. In column (3) the extended set of mediators and confounder is used.

Table 9: Direct and Indirect Effects: Cantons Observable

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total effect:	τ	-0.056**	-0.024	-0.099***	-0.053	-0.071**	-0.070**	-0.074**	-0.074**
		(0.023)	(0.047)	(0.037)	(0.033)	(0.032)	(0.032)	(0.034)	(0.034)
Women									
Direct effect:	$\delta(1)$	-0.095***	0.083	-0.110**	-0.109***	-0.059	-0.056	-0.129***	-0.127***
		(0.030)	(0.079)	(0.054)	(0.040)	(0.044)	(0.046)	(0.043)	(0.043)
Indirect effect:	$\eta(1)$	-0.044**	-0.040	-0.106**	0.001	-0.099***	-0.107***	0.006	0.006
		(0.021)	(0.057)	(0.045)	(0.029)	(0.033)	(0.032)	(0.031)	(0.034)
Men									
Direct effect:	$\delta(0)$	-0.012	0.016	0.007	-0.055	0.029	0.036	-0.080*	-0.079*
		(0.030)	(0.071)	(0.057)	(0.045)	(0.043)	(0.043)	(0.046)	(0.047)
Indirect effect:	$\eta(0)$	0.039*	-0.107*	0.011	0.055**	-0.011	-0.015	0.055*	0.053*
		(0.022)	(0.064)	(0.042)	(0.026)	(0.033)	(0.035)	(0.029)	(0.030)
Mediators, Confounders		basic	basic	basic	basic	basic	extended	basic	extended
Ballots		all	1981	1991	1993	1981,1991	1981,1991	1993	1993
Observations		1825	342	686	797	998	998	719	719

Note: *** p<0.01, ** p<0.05, * p<0.1. Based on data from VOX-surveys no. 161, 421, and 511. Only observations with information about the respondent's canton of residence. Inverse propensity score weighted results. The binary dependent variable is 1 if the respondent voted yes, 0 if no. Standard errors in brackets. Standard errors for direct (δ) and indirect effects (η) are based on 1999 bootstrap iterations. Standard errors of total effect τ are from probit estimates. Specifications (1) to (5) and (7) include mediators and confounders available for all three votes. In columns (6) and (8) the extended set of mediators and confounder is used. In (5) and (7) the number of observations is reduced to those for which the extended set is available.

Table 10: Direct and Indirect Effects: Canton Clustering

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total effect:	τ	-0.056**	-0.024	-0.099***	-0.053	-0.071**	-0.070**	-0.074**	-0.074**
		(0.023)	(0.030)	(0.032)	(0.034)	(0.029)	(0.029)	(0.035)	(0.035)
Women									
Direct effect:	$\delta(1)$	-0.095***	0.083	-0.110**	-0.109**	-0.059	-0.056	-0.129***	-0.127***
		(0.029)	(0.088)	(0.054)	(0.044)	(0.046)	(0.048)	(0.046)	(0.046)
Indirect effect:	$\eta(1)$	-0.044**	-0.040	-0.106***	0.001	-0.099***	-0.107***	0.006	0.006
		(0.018)	(0.036)	(0.036)	(0.037)	(0.031)	(0.034)	(0.044)	(0.046)
Men									
Direct effect:	$\delta(0)$	-0.012	0.016	0.007	-0.055	0.029	0.036	-0.080	-0.079
		(0.028)	(0.054)	(0.046)	(0.051)	(0.039)	(0.036)	(0.058)	(0.059)
Indirect effect:	$\eta(0)$	0.039*	-0.107	0.011	0.055*	-0.011	-0.015	0.055*	0.053
		(0.021)	(0.076)	(0.039)	(0.033)	(0.033)	(0.036)	(0.030)	(0.033)
Mediators, Confounders		basic	basic	basic	basic	basic	extended	basic	extended
Ballots		all	1981	1991	1993	1981,1991	1981,1991	1993	1993
Observations		1825	342	686	797	998	998	719	719

Note: *** p<0.01, ** p<0.05, * p<0.1. Based on data from VOX-surveys no. 161, 421, and 511. Only observations with information about the respondent's canton of residence. Inverse propensity score weighted results. The binary dependent variable is 1 if the respondent voted yes, 0 if no. Canton clustered standard errors in brackets. Standard errors for direct (δ) and indirect effects (η) are based on 1999 bootstrap iterations. Standard errors of total effect τ are from probit estimates. Specifications (1) to (5) and (7) include mediators and confounders available for all three votes. In columns (6) and (8) the extended set of mediators and confounder is used. In (5) and (7) the number of observations is reduced to those for which the extended set is available.