Online Appendix for article "Electoral Reform and Trade-Offs in Representation"

Michael Becher* Irene Menéndez González†

January 25, 2019

Contents

*Institute for Advanced Study in Toulouse and UT1. Email: michael.becher@iast.fr
†University of Mannheim. Email: menendez@uni-mannheim.de

S1 Theoretical model

S1.1 Proof of Proposition 1

We will show that the outcomes in Table 1 are the only equilibrium outcomes given the electoral geography. Proposition 1 directly follows from this. Recall that a perfectly coalition-proof equilibrium is robust to any coalitional deviation by voters and politicians that is self-enforcing. As is discussed by Bernheim, Peleg and Whinston (1987), it is natural to think of players being able to communicate about possible coalitions during the electoral campaign but not being able to commit to a particular voting or entry strategy. In the MR model, this means that each partisan group i in each district d and all 18 politicians can be treated as the relevant players. In the PR model, each group of voters i can be analyzed as a single player that may distribute its votes arbitrarily across lists.

Preliminaries Without loss of generality, let us normalize $u(|0|) = 0$ to reduce notation. The text states the assumption that voters prefer a parliament that implements their ideal policy to any parliament that implements the ideal policy of the next closest group in the policy space and includes one additional high-quality legislator. Formally, this minimal polarization condition requires that

$$-u(|x_L - x_M|) > \max\{(g(3) - g(2), g(2) - g(1), g(1) - g(0)\}.$$

We refer to it as assumption **A1** below. Moreover, given A1, high polarization is formally defined as

$$-u(|x_L - x_M|) > g(3) - g(1)$$

which we refer to as assumption **A2**.

For the PR game, let $\sigma_P = \{P^\omega, P^\omega, P^\omega\}$ denote party P's list of candidates, characterized

by fixed partisan type $P \in \{L, M, H\}$ and endogenous quality $\omega \in \{0, 1\}$.

Majority rule First, verify that the following two parliaments are equilibrium outcomes: $\{L_1^1, M_2^0, H_3^1\}$, $\{M_1^1, M_2^0, M_3^1\}$. Consider $\{L_1^1, M_2^0, H_3^1\}$. By A1, M voters in d=2 cannot commit to vote against M_2^0 as this would change policy to x_L or x_H. So M_2^0 enters, blocking the candidacy of M_2^1, and wins. M_2^0 prefers the parliament $\{L_1^1, M_2^0, H_3^1\}$ to $\{L_1^1, M_2^1, H_3^1\}$ as she obtains private benefits from office $\pi > c$ and disregards externalities of her behavior on g. In contrast, good types L_1^1 (H_3^1) in the other districts are able to run because L_1 (H_3) voters off the equilibrium path are willing to vote against bad types and support any M type instead, as this does not change $x^* = x_M$ given what everybody else is doing. There is no credible coalitional deviation to a Pareto-efficient parliament $g(3)$. Any such deviation has to ensure that M remains the median party in parliament so that M voters in d=2 are willing to vote against M_2^0. Clearly, either L or H voters will be better off reneging from the coalition to obtain their ideal policy. Next, consider $\{M_1^1, M_2^0, M_3^1\}$. This equilibrium requires that L voters in d=1 vote for any L_1^ω that declares candidacy in their district unless M_1^1 enters and no other district is electing a L-candidate. Symmetrically, H voters in d=3 vote for any H_3^ω that declares candidacy in their district unless M_3^1 enters and no other district is electing a H-candidate. Given these strategies, M voters in d=2 do not vote against M_2^0, as doing so would adversely change policy to x_L or x_H. Hence, politicians M_1^1, M_2^0, M_3^1 enter and get elected, and there are no self-enforcing coalitional deviations to induce a Pareto-efficient parliament $g(3)$.

Second, there are no other equilibria. There can be no equilibrium that entails $x^* \neq x_M$. Suppose otherwise, so that at least two L or two H legislators are elected. By A1, any of the above equilibria provide a credible coalitional deviation to such an outcome. For instance, consider $\{L_1^1, L_2^1, i_3^1\}$. Then M and H voters and candidates will jointly defect to induce $\{L_1^1, M_2^0, H_3^1\}$. Moreover, there is no equilibrium where $g^* < g(2)$ and $x^* = x_M$. Again, a

self-enforcing coalition of voters can at least induce parliament $\{L_1^1, M_2^0, H_3^1\}$. Furthermore, subgame perfection rules out the existence of a Pareto-efficient equilibrium where $g^* = g(3)$. Suppose otherwise and let $\{M_1^1, M_2^1, M_3^1\}$ be an equilibrium parliament. The entry of M_2^1 requires that, off the equilibrium path, M voters in $d = 2$ to vote against M_2^0 and for any L (or H) candidate instead. This is not credible, as the best-response of L voters in $d = 1$ dictates that they will drop their support for M_2^0 to secure a parliamentary majority for their preferred policy. The same logic applies to $\{L_1^1, M_2^1, M_3^1\}$ and $\{M_1^1, M_2^1, H_3^1\}$. Finally, it is easy to verify that there are no other equilibria that produce the same x^* and g^*.

Proportional representation There are two cases. First, suppose A2 holds (high polarization). The following two outcome-equivalent parliaments exist in equilibrium: $\{L^0, L^0, M^1\}$ or $\{L^0, L^0, H^1\}$ (we leave out subscripts for districts as there is only one polity-wide district under PR). The equilibrium party lists corresponding to the first parliament $\{L^0, L^0, M^1\}$ are as follows: $\sigma_L = \{L^0, L^0, L^0\}$, $\sigma_M = \{M^1, M^1, M^1\}$, and $\sigma_H = \emptyset$. By A2, the majority of L voters prefer the induced outcomes $x^* = x_L$ and $g^* = g(1)$ to any outcome $x^* \neq x_L$. Hence they cannot credibly commit to vote against list σ_L in which the two top spots are taken by bad types. As a consequence, all three bad types L^0 declare their candidacy, as the chance of winning office is sufficiently high by assumption ($\pi/3 > c$). The third parliamentary seat, by contrast, will be taken by a good type of party M. Concerning this last seat, M and H voters can commit to voting against a list with low-quality candidates as this will not change the median type in the legislature. As a result, gate keepers in party M allow the good types to declare candidacy, and by assumption $\pi/3 > c$ all of them will enter. The logic for outcome-equivalent parliament $\{L^0, L^0, H^1\}$ is symmetric, and the corresponding party lists are: $\sigma_L = \{L^0, L^0, L^0\}$, $\sigma_M = \emptyset$, and $\sigma_H = \{H^1, H^1, H^1\}$. There are no other equilibria. A2 rules out any parliament that leads to $x^* \neq x_L$. A self-enforcing coalition of voters and politicians can always achieve $\{L^0, L^0, M^1\}$ or $\{L^0, L^0, H^1\}$.

Second, suppose A2 does not hold (low polarization). Then the equilibrium parliament is $\{L^0, L^1, M^1\}$ or $\{L^0, L^1, H^1\}$. In equilibrium, party L runs the list $\sigma_L = \{L^0, L^1, L^1\}$, and one of the two remaining parties enters with a list full of good types (by $\pi/3 > c$) and the other stays out. (Mixed strategies are excluded.) By A1, left voters prefer σ_L, which leads to $x^* = x_L$ and $g^* = g(2)$, to any other parliament where L has no majority. As A2 does not hold, off the equilibrium path they vote against party L with two bad types on the top of the list, supporting the alternative list with high quality candidates instead. As a result, exactly one bad type L^0 declares candidacy, allowing two good types L^1 to enter and compete for the chance to win the second seat. The second party that enters has incentives to let the good types run as voters can credibly commit to vote for σ_L otherwise. It is straightforward to verify that there are no other equilibria. This completes the proof. □

S1.2 Alternative formalization

The model presented in the text assumes that politicians in a political party make costly entry decisions and low-quality types are gatekeepers that may block the entry of high-quality types in the election. An alternative approach to model political recruitment within parties is to assume that parties – whether a leadership or rank-and-file members – select candidates subject to the constraint that high quality candidates are costlier, reflecting foregone rents or opportunity costs to the party or better outside options of the candidates (Galasso and Nannicini, 2011, 2017). This approach focuses on the allocation decision of the party rather than the entry decisions of individual politicians and it does not give any special influence to low-quality politicians. In this section, we show that adopting this alternative formulation of party organization leads to the same institutional trade-off.

Specifically, let us suppose that each party $P \in \{L, M, H\}$ selects a slate of candidates for the parliament, denoted by $\sigma_P = \{P_1^\omega, P_2^\omega, P_3^\omega\}$, where $\omega = 0$ denotes a low-quality and $\omega = 1$ a high-quality type. Subscripts denote the electoral district under MR and the list

position under PR. As before, candidate selection takes place simultaneously in all three parties. Parties care about the policy outcome as well as office and they suffer a cost from selecting high-quality politicians, which need not be very large. Formally, a party's utility function is

$$U_P = u(|x^* - x_P|) + s\left(\sum_{d=1}^{3} P_d\right) - c\left(\sum_{d=1}^{3} P_d^1\right).$$

The first-term on the right-hand side captures policy motivations, the s-term captures the benefits of winning parliamentary seats beyond the ability to shape policy and the c-term captures the cost of recruiting high-quality politicians. The cost for nominating a bad politician is normalized to zero and there are positive marginal costs of selecting a high-quality type: $c(3) > c(2) > c(1) > c(0) = 0$. Consistent with evidence that parties do select high-quality politicians in some contests (Besley et al., 2017; Galasso and Nannicini, 2011), we assume that the costs are not prohibitive. In particular, assume that the gains of winning an additional parliamentary seat for sure outweigh the costs of selecting an additional high-quality candidate. As we will see, in equilibrium this can nonetheless lead to an undersupply of high-quality politicians.

Proposition 2 states that the equilibrium policy and valence outcome under majority rule are the same as with the assumption about party organization (i.e., gatekeeping) used in the main text.

Proposition 2. *Assuming the alternative model of party organization, the equilibrium policy and quality under majority rule are $x^* = x_M$ and quality is $g^* = g(2)$.*

Proof. First, verify that the following parliaments are equilibrium outcomes: $\{L_1^1, M_2^0, H_3^1\}$, $\{M_1^1, M_2^0, M_3^1\}$. Consider the first parliament, $\{L_1^1, M_2^0, H_3^1\}$. It is the result of the following party nomination decisions: $\sigma_L = \{L_1^1, L_2^0, L_3^0\}$, $\sigma_M = \{M_1^0, M_2^0, M_3^0\}$, and $\sigma_H = \{H_1^0, H_2^0, H_3^1\}$. The median voter in d=1 votes for L^1, the median voter in d=2 votes for M^0

and the median voter in d=3 votes for H^1. Given the nominated candidates, no group of voters can benefit from a deviation. By A1, M voters in d=2 cannot commit to vote against M_2^0 as this would change policy to x_L. The best response of party M is to select a low-quality candidate in the district. In the remaining two districts, however, party L or H nominates a high-quality candidate because voters are willing to vote against bad types of their party (off the equilibrium path) and support M^0 instead. This threat is subgame perfect as implementing it does not change the policy outcome given what everybody else is doing. Given the assumed benefits of office outweigh the cost of recruitment, party L (H) in d=1 (d=3) has incentives to select a high type. There is also no credible coalitional deviation to a Pareto-efficient parliament $g(3)$, because any feasible coalition is not self-enforcing. Next, consider the equilibrium generating parliament $\{M_1^1, M_2^0, M_3^1\}$. Voter best responses are identical to the model with endogenous entry and party nomination decisions are $\sigma_L = \{L_1^0, L_2^0, L_3^0\}$, $\sigma_M = \{M_1^1, M_2^0, M_3^1\}$, and $\sigma_H = \{H_1^0, H_2^0, H_3^0\}$. Given voters' strategies and what the other parties are doing, there are no beneficial deviations in the candidate selection stage. Finally, it can easily be verified that there are no other equilibria. The logic is nearly identical to the baseline model. □

Proposition 3 summarizes the outcome under PR under the alternative formalization if polarization is high. (There is no pure strategy equilibrium if A2 does not hold.) Taken together, Proposition 3 and Proposition 2 imply the same trade-off as in Proposition 1.

Proposition 3. *Suppose A2 holds. Assuming the alternative model of party organization, under proportional representation the equilibrium policy is* $x^* = x_M$ *and* $g^* = g(1)$.

Proof. Suppose A2 holds (high polarization). The following three outcome-equivalent parliaments exist in equilibrium: $\{L^0, L^0, M^1\}$, $\{L^0, L^0, H^1\}$, $\{L^0, L^0, L^1\}$. The equilibrium party lists corresponding to the first parliament $\{L^0, L^0, M^1\}$ are as follows: $\sigma_L = \{L^0, L^0, L^0\}$, $\sigma_M = \{M^1, M^0, M^0\}$, and $\sigma_H = \{H^0, H^0, H^0\}$. L voters vote for σ_L and the other voters support σ_M. By A2, the majority of L voters prefer the induced outcomes

$x^* = x_L$ and $g^* = g(1)$ to any outcome $x^* \neq x_L$. Because they cannot credibly commit to vote against list σ_L and given what the other voters are doing, party L has no incentive to nominate high-quality types. The third parliamentary seat, by contrast, will be taken by a good type of party M. Voters can commit to voting against a list $\sigma'_M = \{M^0, M^0, M^0\}$ as this will not change the median type in the legislature. The logic for outcome-equivalent parliaments $\{L^0, L^0, H^1\}$ and $\{L^0, L^0, L^1\}$ is symmetric. As in the baseline PR model, there are no other equilibria. □

S1.3 Majority rule with multi-member districts

In the main text we argue that the theory also applies to majority rule with multi-member electoral districts, such as the Swiss case we study in the empirical part of the paper. To illustrate how the institutional logic works in this situation, consider a 7-member parliament. As depicted in Table S1.1, let us assume that there are seven different municipalities of equal population size, denoted by letters a to g. In the majoritarian system, there are two multi-member districts (consisting of three and two municipalities, respectively) and two single-member districts (each comprising a single municipality). This means that a majority of seats under MR is allocated in multi-member districts. As in the baseline model, electoral geography is unequal such that the median in the population (L) is not the median in the median district. L voters are heavily concentrated in three-member district d = 1, H voters are concentrated in single-member districts d= 3, 4 and M voters are the median in the remaining two-member district. The specific distribution of voters within and across districts in Table S1.1 is for concreteness but not required for the argument. Under MR, a voter casts a vote for each seat to be filled in her district. There is no cumulation of votes and a candidate with at least the support of half the voters wins. If fewer candidates than there are seats obtain an absolute majority, the winner for any outstanding seat is determined in a second round using first-past-the post among the top candidates who did not obtain a seat

Table S1.1: Example with multi-member districts under majority rule

Municipality	Electoral district	Voters (fraction in each district) L	M	H
a	1	0.9	0.1	0.0
b	1	0.9	0.1	0.0
c	1	0.9	0.1	0.0
d	2	0.45	0.4	0.15
e	2	0.45	0.4	0.15
f	3	0.2	0.2	0.6
g	4	0.0	0.15	0.85
Population size		0.54	0.21	0.25

in the first round. Under PR, there is a polity-wide district. While not necessary for the argument, it simplifies the analysis to assume that voters who are indifferent over outcomes simply support the ideologically closest candidate(s) or party list.

Consider a situation with high polarization (A2).[1] Given the equilibrium concept, the outcome under majority rule is $x^* = x_M$ and $g^* = g(6)$ and the outcome under PR is $x^* = x_L$ and $g^* = g(3)$. Qualitatively, this implies the same institutional effect captured by the simpler model (Proposition 1): Compared to PR, majority rule is worse at representing the policy preferences of the population at large but better at selecting good politicians.

In the MR game, it is easy to verify that the following political behavior constitutes an equilibrium: in district 1, three type-L^1 politicians enter and win; in district 2, one M^0 and one M^1 candidate enters and wins; in districts 3 and 4, H^1 politicians enter and win. As a result, the median legislator will be from the M party, leading to $x = x_M$ and $g = g(6)$.

[1]Given the increased size of the parliament compared to the baseline model, the equivalent high polarization assumption is that $-u(|x_L - x_M|) > g(7) - g(3)$.

As in the baseline model, M voters in d=2 cannot credibly commit to vote against two M^0-candidates as this would change policy to x_L or x_H given what voters in the other districts are doing. However, they can coordinate to credibly vote for one high quality candidate of party H as this will not change the legislative median but improve policy, thus providing incentives for the entry of one M^1-candidate. In the other districts, good types are selected because voters can vote on quality without changing the partisan identity of the median legislator. Entry decisions follow and all MPs are elected in the first round. As in the baseline model, there is no credible coalitional deviation to a Pareto-efficient parliament $g(7)$. Other outcomes $x \neq x_M$ or $g < g(6)$ cannot occur in equilibrium, as a self-enforcing coalition of voters can always induce $x = x_M$ and $g = g(6)$.

In the PR game, the equilibrium parliament will consist of four L^0 MPs and three high-quality MPs from at least one other party. If indifferent voters support the list of the ideologically closest party and given group sizes in Table S1.1, the unique equilibrium parliament is $\{L^0, L^0, L^0, L^0, M^1, H_1, H_1\}$. A2 implies that L voters prefer the induced outcomes $x^* = x_L$ and $g^* = g(3)$ to any outcome $x^* \neq x_L$ and so politicians of the L party have no incentives to compete on good politicians, in contrast to the other parties.

S2 Data

This appendix provides additional information on the data used in the Swiss case. Table S2.2 reports summary statistics and sources for the variables used in the main analysis reported in Table 3. Recall that the unit of analysis varies with the dependent variable. It is MP-vote in columns 1-3, district-vote in column 4, and MP-parliament in columns 5-7. Descriptive statistics in Table S2.2 are based on the MP-parliament dataset except for the congruence variables.

Note that descriptives for reform intensity cover both the pre-reform and the post-reform parliament. However, between districts in the post-reform parliament median reform intensity is 1.1 (mentioned in the text and underlying Figure 2) and mean reform intensity is 1.07.

Table S2.3 lists the cantonal referendums matched to roll-call votes for the analysis of congruence, including a short description of each issue and summary statistics for district-level referendum outcomes.[2] Referendum results were retrieved from the cantonal database on election and referendums available at `http://www.wahlen-abstimmungen.zh.ch/internet/justiz_inneres/wahlen-abstimmungen/de/abstimmungen/abstimmungsarchiv.html` Parliamentary votes are coded from the parliamentary records (Kantonsrat, 1917, 1920). It is noteworthy that the analysis includes key economic issues (tax reform or working time regulation) and constitutional issues (electoral reform or reform of legislative institutions) before and after the reform. The mean level of support varies considerably across policies. Moreover, there is large variation in policy preferences across districts, which is a strong indication of political polarization. The cross-sectional range in the yes-vote share is always larger than 30 percentage points and on several key issues it is twice as large. As indicated in the table, there are large differences between rural and urban districts. But there also is considerable

[2]The mean support across districts does not correspond to the overall yes-vote (%) in the population because districts, the unit of analysis, vary in size.

variation between urban districts (e.g., there is 31-point gap between Zürich 1 and Zürich 2 on electoral reform or a 16-point gap between Winterthur and Zürich 2 on working time regulation).

Table S2.4 compares referenda with matched to roll-call votes, listed in Table S2.3, with referenda for which there is no roll-call vote. This reveals that referenda with corresponding roll-call votes are fairly similar on key observational features – whether the referendum is mandatory, turnout and the yes-vote share – to those without roll-call votes. Matched referenda are somewhat more contested, though the difference is not statistically significant, and have a slightly higher (by three percentage points) turnout.

Finally, Figure S2.1 plots the municipality-level vote share in the referendum to adopt PR. At a more fine-grained level, it illustrates the large variation in mass support for the reform.

Table S2.2: Descriptive statistics and sources for variables used in analysis in Table 3

	Mean	SD	Min	Max	Source
District					
Reform intensity	0.55	0.65	0.00	1.89	Official compilation of laws (*Offizielle Sammlung*) of canton Zürich Vol. 30
Support reform (share yes vote)	0.53	0.19	0.06	0.93	Cantonal referendum database[1], Statistical Yearbook City of Zürich 1916
Language fractionalization	0.11	0.06	0.01	0.27	Decennial federal census (1910), Statistical Yearbook City of Zürich 1910-11
Religious fractionalization	0.34	0.13	0.04	0.55	Same as above
Foreign population (share)	0.19	0.11	0.02	0.38	Same as above
Employment in industry and crafts	0.52	0.12	0.16	0.75	Decennial federal census (1920)
Members of parliament					
Congruence (MP-canton median)	0.60	0.49	0.00	1.00	Calculated from referendum database and parliamentary records; see text
Congruence (majority of MPs)	0.61	0.49	0.00	1.00	Same as above
Attendance	0.86	0.14	0.18	1.00	Parliamentary records (Kantonsrat, 1917, 1920)
Age (in years)	52.04	10.46	25	79	Same as above
Death	0.04	0.20	0.00	1.00	Same as above
Worker	0.12	0.33	0.00	1.00	Same as above
Farmer	0.09	0.29	0.00	1.00	Same as above
Nationalrat	0.06	0.24	0.00	1.00	Same as above
Left Party	0.29	0.46	0.00	1.00	Parliamentary records, newspapers (*Der Grütlianer*; *Neue Züricher Zeitung*; *Das Volksrecht*)
Farmer's Party	0.07	0.26	0.00	1.00	Same as above

Notes: [1] *Retrieved from* `http://www.wahlen-abstimmungen.zh.ch/internet/justiz_inneres/wahlen-abstimmungen/de/abstimmungen/abstimmungsarchiv.html`

Table S2.3: Cantonal referendums matched to roll-call votes in parliament

Vote ID	Issue	Referendum yes-vote (%)			
		Mean	SD	Min (district)	Max (district)
17427	Introduction of PR for cantonal parliament.	49.9	18.2	24.5 (Dielsdorf)	85.5 (Zürich 2)
17460	Comprehensive tax reform.	55.2	14.9	26.6 (Dielsdorf)	82.2 (Winterthur)
18431	Territorial reform: merger of municipality of Winterthur with 5 suburbs.	82.7	9.5	65.0 (Affoltern)	96.9 (Zürich 2)
18511	Law concerning organization and rules for cantonal parliament.	44.0	10.9	24.4 (Bülach)	66.3 (Zürich 4)
18615	Working time regulation (8-hour day and 48-hour week).	26.1	8.5	12.7 (Dielsdorf)	44.0 (Winterthur)

Notes: Each observation is a post-reform electoral district. Given variation in district size, the mean across districts does not correspond to the overall yes-vote (%) in population.

Table S2.4: Comparing matched to non-matched referendums

	No roll call Mean	Roll call Mean	Difference (p value)
Mandatory referendum	0.74	0.80	0.06 (0.80)
Turnout referendum	0.72	0.75	-0.03 (0.04)
Yes vote referendum	0.68	0.54	0.14 (0.22)
N	27	5	

Notes: Referendum data are from the Abstimmungsarchiv *of the canton Zürich and information on roll calls is from the parliamentary records of the Kantonsrat.*

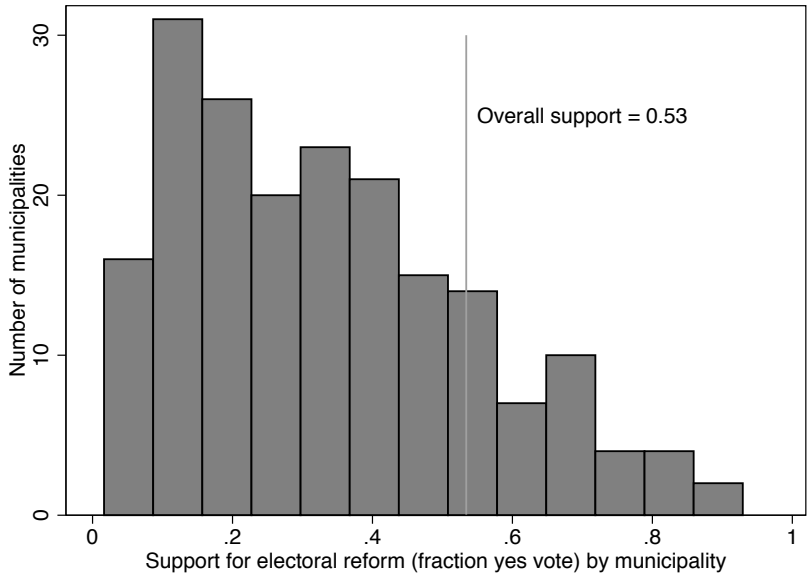


Figure S2.1: Referendum on introduction of PR in December 1916

Notes: Each observation is a municipality (*politische Gemeinde*). There are 187 municipalities in the canton. They vary in population size. The city of Zürich, which is one municipality, is split up into its contemporaneous electoral districts. Data are from the cantonal referendum database (*Abstimmungsarchiv*) and the Statistical Yearbook of the city of Zürich for 1916.

S3 Additional empirical results

S3.1 Measuring the quality of politicians

In the paper, we explain why regular attendance taps into the quality of politicians concerning their integrity or internal motivation (also see Fisman et al., 2015; Gagliarducci, Nannicini and Naticchioni, 2011; Høyland, Hobolt and Hix, 2017). In this section, we provide supporting evidence for this claim showing that attendance is positively related to MPs' reelection rate as well their speechmaking. Moreover, we report results using speeches and education as the dependent variable. The latter taps into a separate aspect of quality relating to competence. These analyses confirm the findings based on attendance. Finally, we consider local political experience as an attribute that features prominently in personal vote theories of electoral institutions.

Reelection. As previewed in the measurement section of the paper, Appendix Table S3.1 shows that MPs' attendance is a statistically and substantively relevant predictor of whether they are reelected. This auxiliary analysis focuses on the two pre-reform parliaments elected under majority rule (1911-1914 and 1914-1917) to avoid confounding the relationship with the subsequent electoral reform. The dependent variable is a dummy equal to 1 if an incumbent MP from the 1911-1914 parliament (Kantonsrat) is reelected in the 1914 election and 0 otherwise, and we estimate linear probability models. The sample includes all MPs that did not exit parliament before the end of the term. For ease of interpretation, attendance has been z-standardized (i.e., mean 0 and unit standard deviation).

The specifications reported in Appendix Table S3.1 start with a simple regression model that only includes attendance and subsequently adds electoral district fixed effects for the 53 electoral districts in the majoritarian system and a vector of MP characteristics (local political experience, education, member of national parliament, age, worker, farmer). As in

the main analysis, adding MP characteristics leads to a small decline in sample size due to missing biographical information for replacement MPs. The sign and size of the coefficient on attendance is robust across specifications and statistically significant at the five percent level except in model 2, where $p = 0.05$. Substantively, the estimate from column (3) in Table S3.1 suggests that one standard deviation increase in attendance is related to a 10 percentage point increase, on average, in the probability of reelection. This corresponds to a 12 percent increase relative to the mean reelection rate of 0.83. This finding is robust to excluding politicians aged 68 and above (model 4), who may be more likely to attend less and seek reelection due to health reasons or retirement.

These results support the claim that attendance is a signal of valence that mattered to parties at the time. They are consistent with partisan selection based on quality, though from the data we cannot verify this directly. It can also be that MPs planning to exit parliament slack off systematically. Either way, attendance is clearly linked to parliamentary careers.

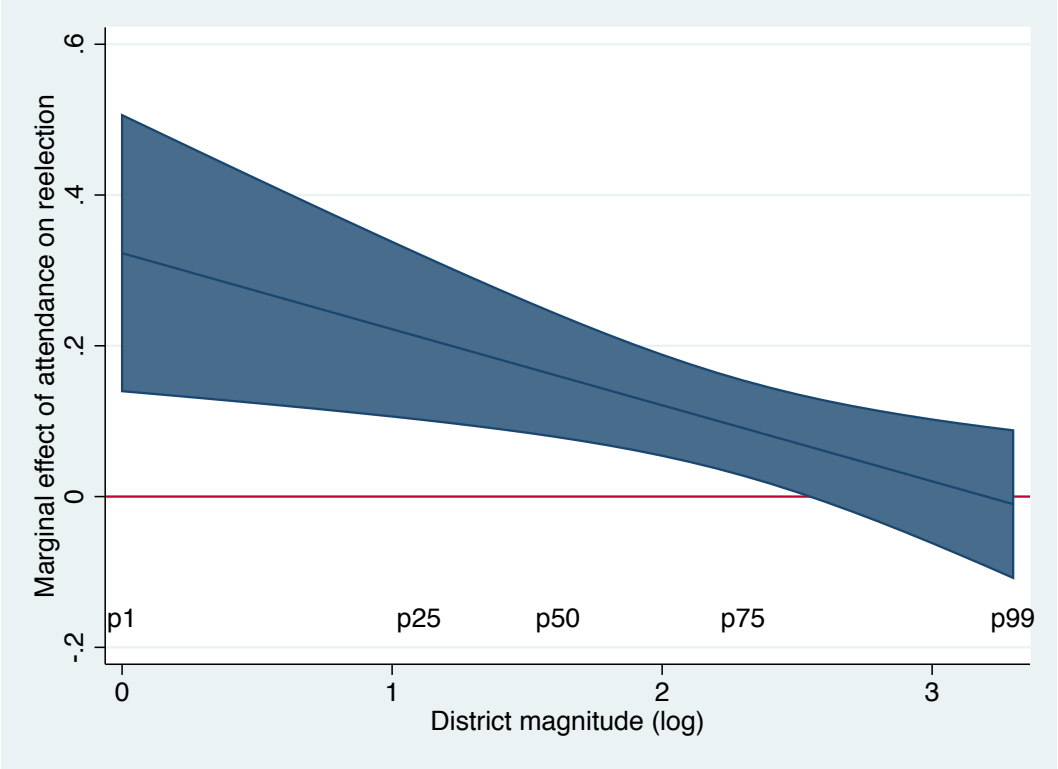
It is also noteworthy that the positive relationship between attendance and reelection is not restricted to single-member or low-magnitude districts. Model (5) excludes single-member and binomial districts. It produces a virtually identical coefficient estimate. Furthermore, Model 6 includes a multiplicative interaction term between attendance and (log of) district magnitude. Estimates based on this interactive specification indicate that while the relationship is most pronounced in small districts and becomes weaker as district magnitude increases, the slope of the interaction term is not very steep. As is illustrated by the marginal effects plot in Figure S3.1, there is a positive and statistically significant relationship between attendance and reelection even in large districts with a magnitude up to about 14. To put this in perspective, in the post-reform districted PR system the median district magnitude is 10. We interpret this as indirect evidence that informational issues do not rule out selection based on quality in multi-member districts.

Table S3.1: Parliamentary participation and reelection

	(1)	(2)	(3)	(4)	(5)	(6)
Attendance	0.07	0.07	0.10	0.10	0.10	0.32
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.09)
Attendance ×						-0.10
District magn. (log)						(0.04)
District FE		✓	✓	✓	✓	✓
MP characteristics			✓	✓	✓	✓
Observations	221	221	208	190	177	208

Notes: Dependent variable: a dummy equal to 1 if an MP from 1911-1914 cantonal parliament is reelected in 1914 election and 0 otherwise. Estimation is by OLS. Standard errors in parentheses are clustered by electoral district (53 clusters). The sample includes all MPs that did not exit parliament before the end of the term. Mean reelection rate: 0.83. For ease of interpretation, attendance has been z-standardized. Model (4) excludes MPs aged 67 and above. Model (5) excludes all districts with a district magnitude of 1 or 2. Note that in the interactive specification of model (6), the constituent term for district magnitude is absorbed by the district fixed effects. MP characteristics: local political experience, education, member of national parliament, age, worker, farmer.

Figure S3.1: Parliamentary attendance and reelection



Note: The figure displays the marginal effect of (standardized) parliamentary attendance on the reelection probability conditional on logged district magnitude from Table S3.1, model 3, with 95% confidence intervals. In the bottom, markers p1,p25,...,p99 indicate percentiles of the empirical distribution of the conditioning variable.

Speeches. Appendix Table S3.2 reports additional results about the conditional relationship between parliamentary attendance and speeches. A negative count model finds a positive and significant relationship. This is consistent with evidence from other settings cited in the paper and further bolsters the argument that attendance is a useful proxy for MPs' parliamentary effort.

Appendix Table S3.3 presents results from an analysis using the number of legislative speeches in key debates as the outcome variable. Key debates concern the votes used in the congruence analysis. While speeches have been used before in the literature as a proxy for legislative effort, one may be concerned that agenda control limits their usefulness (e.g., Proksch and Slapin, 2012; Schwarz, Traber and Benoit, 2017). However, this concern is minimized in the cantonal parliament we study. The speaking agenda is open. Any member of parliament may take the floor in a particular debate. The parliamentary rules guarantee that even if a majority votes to end a debate, any member who has not yet spoken on the issue has the right to take the floor. Given the nature of the dependent variable, we estimate a negative-binomial count model that allows for overdispersion in the speech counts. To capture heterogeneity across debates, the model allows for random variation in the dispersion parameter by debate. In addition to the usual co-variates, the model includes a dummy for the rapporteur in a given debate. The estimation results indicate a significant negative effect of reform intensity.

Table S3.2: Parliamentary attendance and speeches

	(1)	(2)	(3)
Parliamentary attendance	2.17	2.88	4.02
	(0.97)	(1.13)	(1.37)
Education			0.86
			(0.60)
District controls	✓	✓	✓
MP characteristics		✓	✓
Observations	204	203	203

Notes: Dependent variable: number of speeches given by MP in key debates in the pre-reform parliamentary term (Kantonsrat 1914-1917). Key debates concern the votes used in the congruence analysis. The table shows the results from a negative-binomial count model, which allows for overdispersion, estimating the relationship between parliamentary attendance and parliamentary speeches. Standard errors are in parentheses. MPs exiting early or entering late and the president of parliament, who does not participate in debates, are excluded. District controls and MP characteristics are the same as in the main analysis. Education is a dummy for PhD.

Table S3.3: Electoral reform and parliamentary speeches

	(1)	(2)	(3)
Reform intensity	-0.44	-0.48	-0.50
	(0.21)	(0.22)	(0.23)
District controls	✓	✓	✓
District FE		✓	✓
MP characteristics			✓
Observations	1,053	1,053	1,051

Notes: Dependent variable: total number of speeches given by MP in a key debate. The table shows the results from a negative-binomial count model that allows for random variation in the dispersion parameter by debate. Standard errors are in parentheses. MPs exiting early or entering late are excluded. District controls and MP characteristics are the same as in the main analysis plus a dummy for the rapporteur. In addition, model (3) includes a dummy for the rapporteur.

Education. Appendix Table S3.4 repeats the analysis with educational attainment as the dependent variable. While our focus is on politicians' motivation and integrity, education taps into the competence dimension of politicians' quality that has been the subject of several existing studies of institutions and selection (Besley and Reynal-Querol, 2011; Galasso and Nannicini, 2011). Ultimately, both aspects of quality are important for representation. Theoretical models typically assume that quality is one dimensional and they can be interpreted as either competence or integrity, though in the real-world these aspects may not go hand in hand and the effects of electoral institutions may vary across different dimensions of quality. One concern with education in the literature is that it confounds social background with competence, which may be especially relevant before the mass expansion of tertiary education after World War II in Europe, and the assumption that politicians with more formal education are more competent leaders is subject to empirical controversy (Carnes and Lupu, 2016). However, the recruitment of highly educated politicians was not a strategy exclusive to established parties. Social Democrats also recruited politicians with high formal education. As a result, one can argue that education is a meaningful proxy in this historical setting as well that can shed light on the logic of political recruitment under alternative institutions.

Given our biographical data, we code whether an MP has a doctoral degree (most frequently in law or medicine, but also in arts and sciences). This is the case for 14.7% of MPs in the pre-reform parliament. Using this binary variable for educational attainment as the dependent variable, we find evidence that the electoral reform tends to be linked to decline in the selection of highly educated MPs, though the effect is not statistically significant in all specifications. To the extent that these types have higher opportunity costs, this is broadly consistent with the theoretical logic.

Table S3.4: Electoral reform and education of MPs

	(1)	(2)	(3)	(4)	(5)
Reform intensity	-0.06	-0.04	-0.06	-0.04	-0.05
	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)
District controls	✓	✓	✓	✓	✓
District FE		✓		✓	✓
Varying trends					✓
Observations	491	491	723	723	723

Notes: Dependent variable: a dummy variable indicating if MP has a doctoral degree (14.7 % in pre-reform parliament). Models 3-5 adds the 1911-1914 parliamentary term. Estimation is by OLS. Standard errors in parentheses are clustered at level of post-reform electoral districts. District controls are the same as in main specification.

Local experience. Building on earlier personal vote theories of electoral institutions, recent contributions in this literature focus on how electoral rules shape the political selection of candidates with local attributes through strategic decisions by parties and candidates. Previous local political experience or birthplace have been used in recent studies (Nemoto and Shugart, 2013; Shugart, Valdini and Suominen, 2005). As discussed in the main text, we think of these theories as complementary to our theoretical argument. Quality construed in our framework as motivation to contribute to broader parliamentary activities does not inherently favor local over national projects. Our main behavioral measures reflect this focus. It is nonetheless useful to examine local attributes of MPs. While data on birthplace is not available for most of our MPs, we have calculated a dummy variable indicating if an MP has political experience at the local level (e.g., council member or elected administrative office). The theoretical expectation here is more ambiguous than that for attendance, speechmaking or congruence. Our model does not predict that reform intensity is related to the localness of MPs. Theories of the personal vote suggest that a higher dosage of PR should lead to reduced incentives to rely on local attributes if voters do not choose between candidates of the same party (Carey and Shugart, 1995; Shugart, Valdini and Suominen, 2005). De facto, most but not all voters cast straight party votes in the open-list PR election we study. However, there are some preference voters, which muddies the prediction - but from this perspective the same should be true for our other outcome variables. Table S3.5 reports the estimation results. They show that while the sign of reform intensity is consistently negative, the effect is imprecisely estimated (i.e., never significant at the 5 percent level). This strengthens the interpretation that our findings do not simply reflect the local vs. national trade-off discussed in the personal vote literature.

Table S3.5: Electoral reform and local political experience

	(1)	(2)	(3)
Reform intensity	-0.06	-0.05	-0.06
	(0.04)	(0.04)	(0.03)
District controls	✓	✓	✓
District FE		✓	✓
MP characteristics			✓
Observations	491	491	462

Notes: Dependent variable: a dummy variable indicating if MP has local-level political experience. Estimation is by OLS. Standard errors in parentheses are clustered at level of post-reform electoral districts. District controls are the same as in main specification.

S3.2 Further robustness checks

As discussed in the robustness section of the paper, Appendix Table S3.6 shows that the results reported in the main text are robust to alternative ways of operationalizing the outcome variables or electoral institutions as well as to relaxing the parallel trends assumption. Concerning the analysis of MP-voter congruence, one may ask how accounting for abstentions and absent MPs changes the results. So far, the analysis has excluded these cases. One may be concerned that the positive effect of reform intensity masks strategic non-decisions by MPs in districts that were exposed to a larger increase in district magnitude, thus overstating policy responsiveness. To assess this possibility in a straightforward way, column 1 codes all MPs that abstained or were absent on the day of the vote as having cast a vote dissonant with the popular vote. If the problem is relevant, this re-coding should produce a significantly diminished effect. To also allow for the opposite possibility, column 2 codes abstainers or absentees as having cast a congruent vote. In either case, the procedure increases the number of observations, but it does not meaningfully alter the results. The coefficient on reform intensity remains large and significant in each specification.

Furthermore, column 3 takes an alternative approach to capture the institutional environment faced by a particular MP. In this specification, the reform intensity variable is dropped. Instead, the model includes the (log) district magnitude in a given district d and electoral term t. Given the inclusion of post-reform district fixed effects, the main variation also comes from varying exposure to the electoral reform across districts, and hence the interpretation is similar to our main specification. A difference is that this alternative specification measures the pre-reform district magnitude at a lower level of pre-reform districts rather than using the average at the level of post-reform districts. A drawback is that it makes a separability assumption concerning the effect of district magnitude and the introduction of the PR voting rule, and the latter cannot be distinguished from a general time trend. Reassuringly, the results from this alternative specification confirm our main results.

Turning to the analysis of parliamentary attendance, one issue is that both selection based on integrity and reelection incentives may shape MPs' attendance behavior. If reelection motives are dominant, the effect of reform intensity should be less pronounced when the dependent variable excludes the latter part of the term leading up to the next election. Studies of retrospective voting based on the economy suggest that voters' evaluation are heavily skewed toward the last 2-4 quarters before the election (Achen and Bartels 2016, ch. 6; Healy and Lenz 2014). Following this logic, column 4 excludes the last year before the election in the calculation of MPs' attendance rate. Qualitatively, the results are unchanged. Though the effect of reform intensity is about one-third larger. This bolsters the interpretation that parliamentary attendance captures variation in quality of MPs rather than reelection incentives. Column 5 employs the alternative institutional measure, (log) district magnitude, in the attendance regression. Again, this check confirms the main results.

Finally, column 6 includes attendance data from the 1911-14 parliamentary term and so the analysis covers three terms. This allows us to control for varying time trends and to test for the existence of pre-treatment trends (Angrist and Pischke, 2009), and doing so confirms our previous results. (Recall that data limitations restrict this test to parliamentary attendance.) The specification includes time trends that vary by the subsequent exposure to the reform. Specifically, four dummy variables based on the ratio of post-reform to pre-reform district magnitude, approximately corresponding to the four quartiles of the distribution, are interacted with a linear time trend. In addition, column 7 adds a variable that switches on reform intensity during the 1911-14 parliamentary term and sets it to zero otherwise. This is a natural way to assess the existence of pre-treatments trends. The resulting coefficient is small and not statistically significant. This bolsters the validity of our empirical strategy. In contrast, the reform intensity coefficient is substantively and statistically significant. It is about one-third larger than in the baseline specification.

Table S3.6: Effect of electoral reform in alternative specifications

	Congruence			Attendance		
	Abstaining or absent: dissonant	Abstaining or absent: congruent	Alternative reform measure	Excluding 4 pre-election quarters	Alternative reform measure	All three terms, varying trends
	(1)	(2)	(3)	(4)	(5)	(6)
Reform intensity	0.171	0.195		-0.059		-0.071
	(0.057)	(0.049)		(0.024)		(0.024)
Log(District magnitude)			0.132		-0.037	
			(0.051)		(0.014)	
Reform intensity, pre-treatment						-0.009
						(0.023)
Observations	1,104	1,104	898	463	471	699

Notes: The table shows results from alternative specifications. All models include district fixed effects and the same (time-varying) district controls included in Table 3. Models with congruence as the dependent variable also include vote fixed effects. In model 1 (model 2), all abstainers and absentees are coded as casting a dissonant (congruent) vote. In model 4, the measure of attendance excludes the last year before the election. Models 3 and 5 replace the reform intensity variable with the natural log of the district magnitude in a given electoral district. Model 6 adds data on the 1911-14 parliamentary term. It includes time trends varying by treatment exposure and tests for pre-treatment trends by including a reform intensity variable that is switched on in the 1911-14 term and is zero otherwise. Estimation is by OLS. Standard errors in parentheses are clustered at level of post-reform electoral districts.

Additional aggregate level results. Going beyond the district-level results reported in column 4 of Table 3 in the main text, Table S3.7 reports additional aggregate-level results base on other district-level measures of policy representation. They confirm the results on the congruence dimension of representation. Columns 1-3 of Table S3.7 use the average congruence of MPs in the district on a given issue. This is simply the aggregate version of the dependent variable in the micro-level specification. This serves to show that the main findings from the individual-level analysis are not sensitive to the level of analysis. Following a suggestion from an anonymous reviewer, specifications 4-6 use an alternative measure defined as the average popular support minus the average legislative support on a given issue in a district. We take the absolute value of this difference, creating something approximating a distance measure. Given the scale of the dependent variable, we should observe a negative effect of the reform (i.e., a declining gap between voters and politicians on average). This is what we find and the reform intensity coefficient is significant at the 10 percent level in column 6 ($p = 0.068$). Note that a previous version of this table mistakenly reported results in columns 4-6 that measured % Voters - % MPs without taking the absolute value.

Table S3.7: Additional aggregate-level results: effect of reform intensity on other district-level measures of policy representation

	Average Congruence			Absolute difference: % Voters - % MPs		
	(1)	(2)	(3)	(4)	(5)	(6)
Reform intensity	0.22 0.084	0.22 (0.080)	0.17 (0.099)	-0.02 (0.022)	-0.022 (0.020)	-0.065 (0.033)
District FE		✓	✓		✓	✓
District controls	✓		✓	✓		✓
Observations	90	90	90	90	90	90

Notes: All models include vote fixed effects. District controls are the same as in Table 3. Estimation is by OLS. Standard errors in parentheses are clustered at level of post-reform electoral districts.

S3.3 Heterogeneity

Excluding the city of Zürich. At the local level, the city of Zürich had adopted PR in 1913, four years before it was introduced at the cantonal level. This means that some voters and some politicians in the canton were more familiar with the new system than others. (During the campaign to introduce PR at the cantonal level, the recent city-level reform was cited by proponents as proof that PR works.) Moreover, the city of Zürich was also characterized by the highest population density in the canton (1469 per square km compared to the canton average of 291). While district fixed effects pick up such heterogeneity across *Bezirke* in the main analysis, one may ask whether the results are driven by the city with its (slightly) longer experience with PR or its dense population. While this considerably reduces the size of the sample, we can exclude the city of Zürich from the analysis to address this point. Results are reported in Table S3.8. They show that the estimates are very similar to the ones using the full sample.

Pre-reform district magnitude. Table S3.9 examines whether the effect of reform intensity varies by the pre-reform level of district magnitude. Following canonical arguments about the declining marginal effect of district magnitude on seats-votes disproportionality (e.g., Rae, 1967), one may conjecture that the effects of the reform on representation are less pronounced for districts with a larger pre-reform district magnitude (mostly urban areas in this case). To assess this possibility, Appendix Table S3.9 reports results from OLS models that interact reform intensity with the (average) level of pre-reform district magnitude. The results are mixed. For MP-voter congruence as the dependent variable, there is some evidence that the impact of reform intensity is less pronounced where district magnitude was already quite high. In all models, the interaction term is significant at the 10 percent level. For attendance as the dependent variable, there is no evidence of a varying effect of reform intensity. The slope on the multiplicative interaction term has the "wrong" sign, is

substantively small and not significant at any conventional level.

As an additional check, columns (4) and (8) exclude the smallest electoral districts, which does not change the results. This precludes the possibility that the effects are driven only by the low-magnitude districts.

One explanation for these patterns may be that while there is considerable variation in reform intensity, district magnitude increased by at least six seats in the smallest pre-reform electoral districts, effectively smoothing over the steepest part of the seats-votes curve.

Table S3.8: Results excluding the city of Zürich

	Congruence MP-median voter				Parliamentary attendance		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Reform intensity	0.17	0.17	0.16	0.48	-0.033	-0.032	-0.050
	(0.064)	(0.060)	(0.054)	(0.086)	(0.022)	(0.019)	(0.016)
District controls	✓	✓	✓	✓	✓	✓	✓
District FE		✓	✓	✓		✓	✓
MP characteristics			✓				✓
Vote FE	✓	✓	✓	✓	n/a	n/a	n/a
Observations	564	564	543	60	287	287	276

Notes: All specifications are the same as in Table 3 but excluding all MPs from districts within the city of Zürich.

Table S3.9: Heterogeneity: pre-reform district magnitude and the effect of reform intensity

	Congruence				Attendance			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reform intensity	0.244	0.258	0.236	0.189	-0.039	-0.045	-0.060	-0.053
	(0.061)	(0.066)	(0.085)	(0.050)	(0.020)	(0.021)	(0.018)	(0.025)
Reform intensity × Pre-reform district magnitude	-0.020	-0.025	-0.009	-0.014	-0.001	-0.001	0.001	-0.000
	(0.009)	(0.011)	(0.016)	(0.008)	(0.004)	(0.003)	(0.004)	(0.004)
District controls	✓	✓	✓	✓	✓	✓	✓	✓
District FE		✓	✓	✓		✓	✓	✓
MP characteristics			✓				✓	
Vote FE	✓	✓	✓	✓	n/a	n/a	n/a	n/a
Observations	898	898	864	829	471	471	450	439

Notes: District controls are the same as in Table 3. In models without district fixed effects, the constituent term Pre-reform district magnitude is included. Otherwise, it is absorbed by the fixed effects. Models (4) and (8) exclude districts with pre-reform district magnitude < 2. Estimation is by OLS. Standard errors in parentheses are clustered at level of post-reform electoral districts.

S3.4 Exploring mechanisms

The analysis reported in Appendix Table S3.10 shows that the effects are robust to controlling for political experience. We control for two aspects of MPs' prior political experience: (i) A dummy variable indicating if an MP has political experience at the local level (e.g., council member or elected administrative office). (ii) The number of years the MP has served in the cantonal parliament until the beginning of the current term (seniority). Clearly, adding either or both variables does not affect the results. The results are also robust to allowing seniority to have a curve-linear effect (not shown). Consistent with the finding that higher reform intensity does not lead to a consistent decline in the selection of MPs with previous political experience, these findings rule out that the effect of the reform is mechanically driven by supply-constrained parties that cannot find enough high-quality candidates.

Combined with the use of party fixed effects in extended specifications in the paper, these additional results also rule out the explanation that the reform effect on attendance is driven by fixed differences across parties, such as lower attendance of farmers due their occupational demands during harvest. To bolster this point descriptively, Figure S3.2 provides boxplots for parliamentary attendance by political party for the 1917-1920 term. It shows that average attendance is very similar for members of bourgeois parties, Social Democrats or the Farmer's Party. Mean attendance is highest among the Farmer's Party, though this small difference is not statistically significant at the five percent level.

Beyond ruling out these alternative channels, note that the evidence is broadly consistent with the implication of the argument that political selection of politicians based on partisanship and quality constitute an important channel through which electoral institutions influence representation. Given the geographic concentration of left voters in industrialized areas, the introduction of PR went hand in hand with a nearly twofold increase of the seats won by Social Democrats from 20.7% to 38.1%. The left saw a moderate increase in its overall vote share from 34.2% to 39.8%, which stems in part from an extension of competi-

tion to districts previously not contested. Overall, the increase in seats was mainly due to a more proportional translation of votes into seats. Both channels occur in the model. With the reform, the left's votes-seats-ratio drops from 1.65 to 1.04 – changing from significant underrepresentation to near-perfect proportionality. This is in line with the theoretical logic and it reflects the hopes and fears of contemporary politicians. For instance, a leading Social Democratic party newspaper saw in the disproportional votes-seats translation in the 1914 election a "cry for proportional representation," and it calculated that, even holding the number of candidates and distribution of votes constant, the party would increase its seat share by more than 50%.[3] A regression analysis reported in Appendix Table S3.11 shows that left MPs are more congruent, on average, with the cantonal median voter than other MPs before and after the reform. This suggests that changing the partisan color of parliament matters substantively, not just descriptively. In addition to the increasing representation of the Social Democrats, the new Farmers' Party won 11% of the seats.

Appendix Table S3.11 reports the results of an additional implication of the model concerning the institution-varying relationship between political parties and the quality of their MPs. The theory posits that the quality of politicians is not inherently different across political parties. Electoral institutions shape the incentives of parties to select good politicians. The argument implies that there is an institution-varying correlation between the partisan color and quality of politicians. Under MR elected L types have higher quality than the average. Recall that relevant equilibrium parliament is $\{L_1^1, M_2^0, H_3^1\}$. Under PR, this is no longer the case and the reverse can occur. The reason is not limited supply of good politicians. All parties could in principle run a full list of good types. Rather, it reflects the varying abilities of voters to credibly commit to vote against bad politicians of their most-preferred party. In line with this auxiliary implication we find that, on average, left MPs have significantly better attendance records than other MPs before the reform, controlling

[3] *Grütlianer* (Zürich), April 27, 1914, p. 1.

for MPs' socio-demographic characteristics. This attendance premium disappears after the reform.

Table S3.10: Effect of electoral reform on political representation controlling for local political experience and seniority in cantonal parliament

	Congruence			Attendance		
	(1)	(2)	(3)	(4)	(5)	(6)
Reform intensity	0.21	0.21	0.21	-0.057	-0.057	-0.057
	(0.066)	(0.065)	(0.066)	(0.015)	(0.015)	(0.015)
Local political experience	-0.34		-0.04	-0.002		-0.002
	(0.021)		(0.020)	(0.021)		(0.021)
Parl. seniority (years)		0.001	0.002		-0.000	-0.000
		(0.001)	(0.001)		(0.001)	(0.001)
District controls	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
MP char.	✓	✓	✓	✓	✓	✓
Observations	864	864	864	450	450	450

Notes: Dependent variable: congruence between MP and cantonal median voter (models 1-3); parliamentary attendance (models 4-6). Estimation is by OLS. Standard errors in parentheses are clustered at level of post-reform electoral districts. All congruence models include vote fixed effects. District controls and MP characteristics are the same as in main specification (Table 3).

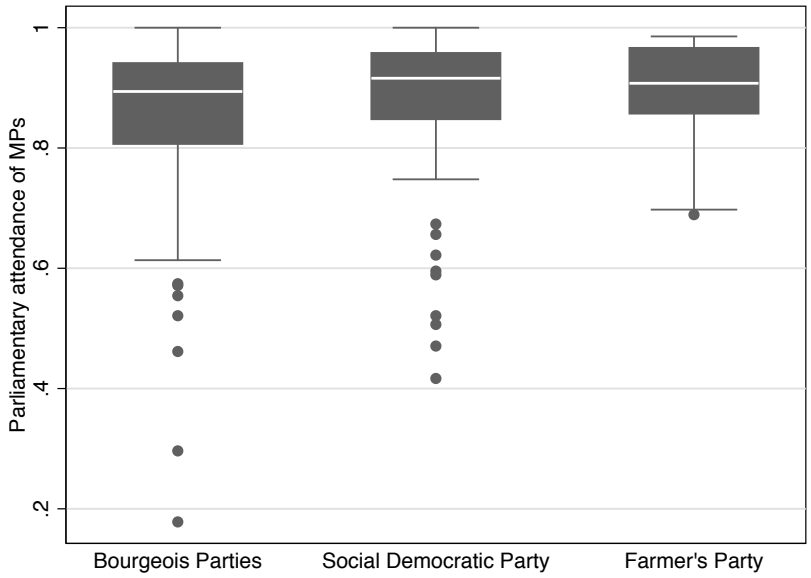


Figure S3.2: Attendance in 1917-1920 parliament by party affiliation

Notes: Boxplots by party. Mean attendance is 0.87.

Table S3.11: Left MPs, policy responsiveness and parliamentary attendance

	Congruence MP-median voter				Parliamentary attendance			
	Pre-reform (1914-17)	Pre-reform (1914-17)	Post-reform (1917-20)	Post-reform (1917-20)	Pre-reform (1914-17)	Pre-reform (1914-17)	Post-reform (1917-20)	Post-reform (1917-20)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Left MP	0.26	0.25	0.40	0.37	0.08	0.10	0.02	0.03
	(0.08)	(0.07)	(0.03)	(0.05)	(0.03)	(0.03)	(0.02)	(0.02)
District control	✓	✓	✓	✓	✓	✓	✓	✓
MP characteristics		✓		✓		✓		✓
Observations	386	375	512	489	228	222	243	228

Notes: Dependent variable: congruence (models 1-4) is a dummy variable indicating whether an MP's parliamentary vote on an issue is congruent with the canton majority in the corresponding referendum (available for parliaments elected 1914 and 1917); parliamentary attendance (models 5-8) is an MP's average attendance rate in a given parliament (parliaments elected 1914 and 1917). Estimation is by OLS. Standard errors in parentheses are clustered at level of electoral districts. District controls: voter support for electoral reform (incl. second-order polynomial), language fractionalization, religious fractionalization, foreign population, employment share in industry. MP characteristics: age, worker, farmer, member of national parliament, exit due to death.

S4 Contemporary Analysis

Table S4.1 summarizes the variables and their sources for the analysis of electoral reform in the European Parliament. As noted in the text, the analysis focuses on the last parliament elected under the old rules (1994-1999) and the first post-reform parliament (1999-2004). The reason is that in 1994 many British districts were redrawn and in 2004 enlargement significantly altered the composition of the European Parliament. Note that 91% of all pre-reform districts are perfectly nested in post-reform districts, and the remaining cases are allocated mostly to one post-reform district, with the exception of Staffordshire East and Derby, which is dropped from the analysis.

Electoral rules are coded based on the European Parliamentary Elections Act of 1999 (UK) and comparative reports (European Parliament Directorate General for Research, 1997, 1999). As explained in the paper and listed in Table S4.1, all three outcome variables are calculated from Hix, Noury and Roland (2007), which provide data on all recorded legislative votes. Their database does not include identifiers for electoral districts, which are needed for our analysis. We matched MEPs to their districts based on election data.

Table S4.1: Descriptive statistics for analysis of electoral reform in European Parliament

	Mean	SD	Min	Max	Source
Reform intensity	0.13	0.50	0	2.40	Coded based on European Parliamentary Elections Act of 1999 (UK), available at `http://www.legislation.gov.uk/ukpga/1999/1/pdfs/ukpga_19990001_en.pdf`, and comparative reports from European Parliament Directorate General for Research (1997, 1999).
Attendance	0.69	0.21	0.00	1.00	Calculated from data compiled by Hix, Noury and Roland (2007), available at `http://personal.lse.ac.uk/hix/HixNouryRolandEPdata.HTM`
Shirking	0.08	0.09	0.00	0.67	Same as above.
NOMINATE (first dimension)	0.10	0.50	-0.92	0.93	Same as above.
Mean citizen ideology	5.14	0.28	4.72	5.98	European Election Study 1994 and 1999 (Schmitt et al., 1997; Eijk et al., 1999).

Notes: includes MEPs in the 4th and 5th European Parliament (N=1,426). In the 4th term, there is no survey data on citizen ideology for Austria, Finland and Sweden (102 cases). When estimating NOMINATE scores, Hix, Noury and Roland (2007) dropped MEPs who participated in less than 20 roll-call votes (25 cases, 3 without survey data).

Appendix References

Achen, Christopher H. and Larry M. Bartels. 2016. *Democracy for Realists: Why Elections Do Not Produce Responsive Government.* Princeton: Princeon University Press.

Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics.* Princeton and Oxford: Princeton University Press.

Bernheim, Douglas B., Bezalel Peleg and Michael D. Whinston. 1987. "Coalition-Proof Nash Equilibria I. Concepts." *Journal of Economic Theory* 42:1–12.

Besley, Timothy and Marta Reynal-Querol. 2011. "Do Democracies Select More Educated Leaders?" *American Political Science Review* 105(3):552–566.

Besley, Timothy, Olle Folke, Torsten Persson and Johanna Rickne. 2017. "Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden." *American Economic Review,* 107(8):2204–42.

Carey, John M. and Matthew Soberg Shugart. 1995. "Incentives to Cultivate a Personal Vote: A Rank Ordering of Electoral Formulas." *Electoral Studies* 14(4):417–439.

Carnes, Nicholas and Noam Lupu. 2016. "What Good Is a College Degree? Education and Leader Quality Reconsidered." *Journal of Politics* 78(1):35–49.

Eijk, C., M. Franklin, K. Schoenbach, H. Schmitt and H. Semetko. 1999. "European Election Study 1999." `https://doi.org/10.17026/dans-z9j-vy6m`. International Research Group "European Election Studies", IPSOS, Hamburg, Germany (primary investigator). Data file available from DANS.

European Parliament Directorate General for Research. 1997. "Legislation Governing Elections to the European Parliament." `http://www.europarl.europa.eu/workingpapers/`

poli/w13/country_en.htm. European Parliament, Luxembourg: Working Paper Political series Working Paper Political series W 13.

European Parliament Directorate General for Research. 1999. "Legislation Governing Elections to the European Parliament." http://www.europarl.europa.eu/RegData/etudes/etudes/join/1999/166617/IPOL-AFCO_ET(1999)166617_EN.pdf. European Parliament, Luxembourg. Working Document Political Series 01.1999.

Fisman, Raymond, Nikolaj A. Harmon, Emir Kamenica and Inger Munk. 2015. "Labor Supply of Politicians." *Journal of the European Economic Association* 13(5):871–905.

Gagliarducci, Stefano, Tommaso Nannicini and Paolo Naticchioni. 2011. "Electoral Rules and Politicians' Behavior: A Micro Test." *American Economic Journal: Economic Policy* 3(3):144–74.

Galasso, Vincenzo and Tommaso Nannicini. 2011. "Competing on Good Politicians." *American Political Science Review* 105(1):79–99.

Galasso, Vincenzo and Tommaso Nannicini. 2017. "Political Selection Under Alterenative Electoral Rules." *Public Choice* 171(3-4):257–281.

Healy, Andrew and Gabriel S. Lenz. 2014. "Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy." *American Journal of Political Science* 58(1):31–47.

Hix, Simon, Abdul G. Noury and Gérard Roland. 2007. *Democratic Politics in the European Parliament*. Cambridge, U.K.: Cambridge University Press.

Høyland, Bjørn, Sara B. Hobolt and Simon Hix. 2017. "Career Ambitions and Legislative Participation: The Moderating Effect of Electoral Institutions." *British Journal of Political Science* Forthcoming.

Kantonsrat. 1917. *Protokoll des Kantonsrates für die Amtsperiode 1914-1917*. Zürich: Genossenschaftsbuchdruckerei.

Kantonsrat. 1920. *Protokoll des Kantonsrates für die Amtsperiode 1917-1920*. Zürich: Genossenschaft Buchdruckerei des Schweiz. Grütlivereins.

Nemoto, Kuniaki and Matthew S. Shugart. 2013. "Localism and coordination under three different electoral systems: The national district of the Japanese House of Councillors." *Electoral Studies* 32(1):1–12.

Proksch, Sven-Oliver and Jonathan B. Slapin. 2012. "Institutional Foundations of Legislative Speech." *American Journal of Political Science* 56(3):520–537.

Rae, Douglas W. 1967. *The Political Consequences of Electoral Laws*. Yale University Press.

Schmitt, H., C. Eijk, E. Scholz and M. Klein. 1997. "European Election Study 1994." `doi:10.4232/1.2865`. European Commission [Principal investigator]. Available from GESIS Data Archive, Cologne. ZA2865 Data file Version 1.0.0.

Schwarz, Daniel, Denise Traber and Kenneth Benoit. 2017. "Estimating Intra-Party Preferences: Comparing Speeches to Votes." *Political Science Research and Methods* 5(2):379–396.

Shugart, Matthew Søberg, Melody Ellis Valdini and Kati Suominen. 2005. "Looking for Locals: Voter Information Demands and Personal Vote-Earning Attributes of Legislators under Proportional Representation." *American Journal of Political Science* 49(2):437–449.