Endogenous Credible Commitment and Party Competition Over Redistribution Under Alternative Electoral Institutions

Michael Becher

University of Konstanz

This version: February 2, 2015

Abstract

Democracies do not legally bind parties to their policy promises. Thus winning the power to set policy through elections requires making credible commitments to pivotal voters. This paper analyzes theoretically and empirically how the commitment problem affects partisan conflict over redistribution. A theoretical model shows that under majoritarian electoral rules parties’ efforts to achieve endogenous commitment using candidate selection to policies preferred by the middle class leads to different behavior and outcomes than suggested by existing theories that assume commitment or rule out endogenous commitment. Left parties may respond to rising inequality by moving to the right in majoritarian systems but not under proportional representation. The theory also unbundles the anti-left bias attributed to majoritarian systems. The empirical analysis finds evidence for key implications of this logic using panel data on party positions and by analyzing devolution in Britain as a natural experiment to compare candidates under alternative electoral rules.

*I am especially grateful to Nolan McCarty, Carles Boix, John Londregan, Francesc Amat, Lucy Barnes, Pablo Beramendi, Christian Breunig, Catherine de Vries, Michael Donnelly, John Huber, Iain McLean, Wolfgang C. Müller, Jonas Pontusson, Tom Romer, David Rueda, Ken Scheve, David Soskice, Daniel Stegmueller, and seminar participants at Barcelona (IPEG), Chicago (MPSA 2014), Florence (EUI), Konstanz, LSE, Washington, D.C. (APSA 2014) and Vienna. Excellent research assistance was provided by Florian Richter and Armand Zorn. The usual disclaimer applies.

†Graduate School of Decision Sciences and Department of Political Science, University of Konstanz, Germany. Email: michael.becher@uni-konstanz.de
Democracies do not legally bind individual politicians or political parties to their policy promises (Manin, Przeworski and Stokes, 1999). Thus winning the power to control policy through elections requires making credible policy appeals to pivotal voters. Doing so is challenging for parties with strong ideological or organizational links to core voters or groups whose interests diverge from that of the pivotal voters. In the absence of exogenous enforcement politicians that care about policy, for intrinsic or instrumental reasons, face a time inconsistency problem (Alesina, 1988). To win elections they want to commit to promises that, if elected, they have incentives to break. This may make them unelectable in the eyes of pivotal voters. This paper makes the point that the commitment problem affects party competition over redistribution in important ways. When and how do parties achieve endogenous commitment to policies preferred by pivotal voters? With what consequences for elections and redistribution? Do electoral systems influence parties’ responses to the commitment problem? These questions are central to debates in the literature on inequality and redistribution in democracies as well as the theoretical and empirical literature on party competition. Despite significant advances, discussed below, existing research does not fully answer these questions. I begin to address this shortcoming by providing a new theory of endogenous commitment and new evidence based on a times-series cross-section analysis of party positions and a natural experiment on party strategies under alternative electoral rules.

The problem of credible commitment in policymaking occupies a central place in modern political economy research. For instance, institutional reforms that introduce veto players as constraints on absolute rulers have been linked to good economic performance (North and Weingast, 1989; Stasavage, 2002). Scholars also argue that elites facing redistributive demands under the threat of revolution are limited in their ability to make credible policy concession but that the introduction of democracy can serve as a commitment device (Acemoglu and Robinson, 2006). Accounting for the absence of external enforcement of campaign promises in democracies, the innovative theory of Iversen and Soskice (2006) demonstrates
that lack of commitment in partisan conflict over redistribution can have profound consequences and explain why some democracies redistribute more than others. This is one of the enduring puzzles in the literature, as simple economic explanations fail to explain the large variation in redistributive policies across advanced industrialized democracies (e.g., Alesina and Glaeser, 2004; Lupu and Pontusson, 2011; Milanovic, 2000; Persson and Tabellini, 2003). The model of Iversen and Soskice (2006) compares the outcome of party competition over redistribution between majoritarian and proportional electoral systems when parties are unable to make binding promises. It considers a political economy with three groups (low-income, middle-income, high-income). Institutionally, proportional representation (PR) is characterized by coalition government between representative parties and majoritarian systems are characterized by competition between two leadership parties, with the winning party setting policy. In this setting, the middle income group is pivotal. The resulting implications are clear-cut. If the left party wins the election in majoritarian systems it may renege on its promised moderate policy and instead impose a policy of high taxes and little benefits for everybody but the poor. Right parties may also renege, but this only means low taxes and no benefits. As a consequence, middle class voters in majoritarian systems tend to strategically support right parties, leading to relatively low levels of redistribution. This stands in contrast to PR. Here coalition government between a left and centrist party provides a credible mechanism due to mutual veto power that protects the middle class from being soaked and elections tend to produce center-left coalitions resulting in high levels of redistribution. Thus, the model can explain the striking twin observation that governments are more left-leaning and redistribution is more pronounced under proportional than under majoritarian electoral institutions (Iversen and Soskice 2006; also see Funk and Gathmann, 2013; Persson and Tabellini 2003).1

1Iversen and Soskice (2010) apply this logic to explain real exchange rates and competitiveness and Iversen and Soskice (2009) discuss its implication for inequality. Several other
Given the high stakes of the redistribution game it is puzzling, however, that parties competing under majoritarian rules systematically fail to solve the commitment problem in equilibrium. Left parties in particular, who otherwise have to endure significant policy losses, have clear incentives to devise a credible commitment mechanism. If successful, endogenous commitment may substantively change the redistributive game. To capture the centripetal incentives of election competition, Iversen and Soskice (2006) assume that parties in majoritarian systems are leadership parties. As their analysis reveals, however, this is not sufficient to achieve commitment. This paper builds on Iversen and Soskice (2006) and extends farther the idea that parties have incentives to achieve endogenous commitment. The main theoretical intuition is that under some conditions parties will make costly investments to credibly appeal to pivotal middle-class voters. More specifically, parties can achieve endogenous commitment if they do not only choose the party leader but also select candidates for the legislature, as is the case nearly all developed democracies (Ranney, 1981). By selecting a sufficient number of candidates from their moderate wing, parties can credibly tie their hands, at least probabilistically, in post-electoral policymaking. For the left, endogenous commitment can be good politics because centrist policy outcomes entail more redistribution to the poor than the policies pursued by a right government. Selecting moderates is costly, however, so whether candidate-based commitment is worthwhile depends on the behavior of competitors and features of the political economy. I develop a simple game theoretical model to clarify the logic of the argument and derive several novel observable implications, some of which are later confronted with data. The model considers a majoritarian electoral systems with multiple districts and parties that decide which type of candidate to select models analyze how electoral rules affect redistribution, though they to not aim to explain variation in government partisanship across institutions (e.g., see Persson and Tabellini 2000).
in each district. The analysis shows that parties’ strategic decision whether to commit to a centrist redistributive platform is endogenous to political competition and varies with income inequality and costs of recruiting moderates.

Endogenous commitment implies that left parties in majoritarian electoral systems need not suffer from an observable electoral bias. Rather, the existence of an electoral disadvantage is conditional on context. In that sense, a main result of Iversen and Soskice (2006, Proposition 1) is a special case. Comparing redistribution in the majoritarian setting with PR unbundles the anti-left bias of majoritarian systems observed in the literature. It reveals that differences in redistribution and government partisanship can occur because left parties are (i) less likely to win election under majoritarian than PR electoral rules (as in Iversen and Soskice (2006)) or (ii) because left parties are credibly more centrist in majoritarian systems (endogenous commitment). Existing empirical studies do not distinguish these mechanisms (Döring and Manow, 2013; Funk and Gathmann, 2013; Iversen and Soskice, 2006).

Another novel implication is that rising top inequality may cause left parties to move to the right in majoritarian but not in PR systems. This does not follow from conventional pivotal voter or partisan theories of redistribution. This perhaps counter-intuitive prediction for majoritarian systems is driven by the left’s response to the commitment problem. In particular, as the rich become richer relative to everyone else, middle-income voters will be more willing to support the left - but only if the left’s commitment to protect their interests is credible. As higher top incomes imply more gains from redistribution for the poor, the left has more incentive to make costly investments in moderate candidates. Because coalition government serves as a commitment devices under PR, the electoral incentives are different.²

²Barth, Finseraas and Moene (2014) probably provide the first model predicting that inequality can lead left parties to move to the right. My argument differs in important ways, including the commitment mechanism. There are two crucial differences regarding the predicted effect of inequality on parties. First, the impact of inequality varies by political
By the same logic, left parties should select different candidates under different electoral rules. The analysis also highlights that candidate selection in non-marginal districts matters in majoritarian settings. Parties may make costly investments in candidates beyond marginal districts, in contrast to existing theories (Galasso and Nannicini, 2011).

I use two distinct data sets to evaluate two critical implications of the theory. First, a times-series cross-section analysis covering 16 affluent democracies between 1950 and 2010 finds that there is an institution-varying impact of income inequality on the position of the mainstream left party. Consistent with the theory, left parties in majoritarian systems move to the right in response to rising inequality whereas left parties in PR do not. This result also contributes to the large empirical literature on the determinants of party positions, especially recent work on the consequences of income inequality (Barth, Finseraas and Moene, 2014; McCarty, Poole and Rosenthal, 2006; Pontusson and Rueda, 2010). The few existing studies do not consider the conditioning influence of electoral institutions and examine a smaller number of elections. Second, I provide some evidence on the candidate selection mechanism. To do so, I analyze the creation of the new Scottish Parliament in Britain as a natural experiment. The institutional setting created by decentralization enables me to compare the selection of candidates by the main center-left parties in the same districts but under alternative electoral rules. The research design accounts for unobserved district heterogeneity and changes in the median voter at large across elections. The main finding is that, in line with the logic of endogenous commitment, Labour disproportionally increases its share of leftist candidates when competition moves to PR.

Second, even a rise in inequality that increases the tax base can induce the left to commit to a less generous platform.
A Solution of the Commitment Problem

The political economy literature spells out the importance of credibility in political competition, but it has paid relatively little attention to what makes policy promises by political parties credible. Most theoretical accounts either simply assume that parties’ policy promises are binding or that policy promises are not binding and parties can take no action to endogenously increase their credibility (Persson and Tabellini, 2000, ch. 3-5). One may argue that parties easily achieve commitment merely based on re-election incentives. Modeling party competition as an infinitely repeated game demonstrates that there can be an equilibrium with endogeneous commitment to the median voter’s ideal point by policy-motivated parties purely based on reputation. But the seminal analysis of Alesina (1988) also shows that the existence of this type of equilibrium depends many factors, such as voters and politicians having long time horizons and use particular punishment strategies, that need not be present.\footnote{Taking a different approach, Harrington (1993) analyzes the impact of re-election pressures on the link between promises and policy in a model with uncertainty and divergent priors about the mapping from policies to outcomes.} The perfect commitment implied by the long-run view is also difficult to reconcile with evidence on partisan effects (Lee, Moretti and Butler, 2004) and cross-national variation in government partisanship across countries (Iversen and Soskice, 2006). Thus, scholars have started to analyze additional commitment mechanisms.

Candidate selection is a commitment device available to parties even when time horizons are limited. This is illustrated clearly by the British Labour Party. The creation of ‘new’ Labour was to a significant part about solving the commitment problem that lies at the heart of the paper and candidate selection was an important part of the solution. After four successive electoral defeats to the Conservative Party between 1979 and 1992, commitment
to middle-class friendly tax policy was perceived as a real problem by the party. Surveys and focus groups indicated that pivotal groups of voters did not trust the moderation of the parties’ redistributive program (Gould, 1998). In this context, reforming the candidate selection process to reduce the influence of labor unions was a deliberate strategy to back up the moderation of the redistributive platform with more moderate candidates and thus an important step on the way to ‘new’ Labour and three successive electoral victories (Russell, 2005). My theoretical model captures and elucidates the importance of candidate selection as an endogenous commitment device. It is related to the citizen candidate framework pioneered by Besley and Coate (1997) and Osborne and Slivinski (1996). The common analytical ground concerns the importance of strategic delegation. However, the model departs from the citizen-candidate framework in important ways. It focuses on the selection of candidates within parties, rather than entry decisions by individual candidates. It also incorporates uncertainty about candidates’ motivation.

**Theoretical Model**

In this section, I propose a stylized model to analyze when and with what consequences parties competing under majoritarian electoral rules make costly investments in moderate candidates to achieve endogenous commitment. Building on Iversen and Soskice (2006), the model considers a political economy with three economic groups, low-income ($L$), middle-

---

4 Citizens across developed democracies regularly express low levels of trust in parties (Dalton and Weldon, 2005).

5 Examples of other commitment devices discussed in the literature include patronage (Keefer and Vlaicu, 2008) or party formation (Levy, 2004; Morelli, 2004).
income (M), and high-income (H), of equal size and indexed by J (also see Acemoglu and Robinson, 2006; Persson and Tabellini, 2000). Each group cares about its disposable income and political conflict is over the choice of a possibly group-specific tax rate and transfer payment.6

Preferences and Redistributive Policy

J’s indirect utility function is $V^J(p^J) = y^J + p^J$, where $y^J$ is the exogenous gross income and $p^J$ is the net government transfer received by J. Each group has a maximum taxable capacity $T^J$ that is less than its exogenous income, capturing incentive effects of taxation. Hence the net transfer to a group is constrained below: $p^J \geq -\bar{T}^J$. Taxable capacity increases in exogenous income, $0 = \bar{T}^L < \bar{T}^M < \bar{T}^H$, where the taxable capacity of the low-income group is normalized to zero. The total revenue available for redistribution is $\bar{T} \equiv \bar{T}^M + \bar{T}^H$. The budget is balanced, $\sum_J p^J = 0$. The final assumption is that net transfers are non-regressive, $p^L \geq p^M \geq p^H$. The policy setup is more general than the more common assumption of a linear income tax and a general lump sum payment. In sum, given these assumptions the most preferred transfer policies for group J, denoted by the vector $\hat{p}^J = (p^L, p^M, p^R)$, are as follows:

$$\begin{align*}
\hat{p}^L &= (\bar{T}, -\bar{T}^M, -\bar{T}^H) \\
\hat{p}^M &= (\frac{\bar{T}^H}{2}, \frac{\bar{T}^H}{2}, -\bar{T}^H) \\
\hat{p}^H &= (0, 0, 0)
\end{align*}$$ (1)

Assuming equal group size is without loss of generality. What matters is that no group constitutes a majority.

6Assuming equal group size is without loss of generality. What matters is that no group constitutes a majority.
Accordingly, low-income citizens would most prefer to tax the middle- and high-income group at their full capacity and receive all the transfers. Middle-income citizens would tax the high-income group at their full capacity and keep as much to themselves as allowed by non-regressivity. High-income citizens prefer no income redistribution.

**Parties, Candidates, and Institutions**

Redistributive policy is determined in a multimember legislature by politicians elected under plurality rule (first-past-the-post) in single-member districts. Politicians are policy motivated. They care about the well-being of a particular group. Corresponding to the three types of voters, there are three types of politicians \( (L, M, H) \). Two political parties compete by selecting a candidate for each district. Policy promises made before the election are not binding. But as we will see, the degree of credible commitment to the policy preferred by the pivotal middle-income voter is endogenous to candidate selection. After the election, the agenda setter is randomly drawn from the majority party’s legislators and makes a proposal that becomes policy only if it is supported by a majority of legislators. This setting can be interpreted as a Westminster-style parliamentary democracy, which is characterized by majoritarian electoral rules and a fusion of executive-legislative relations.

Suppose there are two parties \( P \in \{LM, MH\} \) that have a chance of winning the election, a center-left party, \( LM \), and a center-right party, \( MH \). This assumption captures Duverger’s Law, according to which majoritarian electoral rules lead to the emergence of a two-party system (Cox, 1997; Duverger, 1954). \( LM \) consists of a majority faction of \( L \) politicians and a minority faction of \( M \) politicians. \( MH \) consists of a majority faction of \( H \) politicians and a minority faction of \( M \) politicians. The key departure from Iversen and Soskice (2006) is that there are multiple districts and parties select candidates for each district. Specifically, consider three heterogeneous districts (e.g., see Persson and Tabellini, 2000, ch. 8). In the marginal district, middle-income voters \( M \) are pivotal. There also is a district where \( L \) and
$H$ voters, respectively, are in the majority. I will refer to the latter as safe districts for LM and MH because their candidate wins for sure.

In terms of party organization, the model captures that candidate selection in established democracies, with the notable exception of open primaries in several U.S. states, usually takes place within political parties and successful parliamentary candidates need to be supported by a majority of members in national or local party bodies. While parties’ internal rules on candidate selection vary, the party leadership typically does not have the formal power to unilaterally select candidates (Bille, 2001; Ranney, 1981). Thus, the model assumes that the majority faction in the party selects the candidates: the $L$-majority select candidates in party LM the $H$-majority select candidates in party MH. While the logic of endogenous commitment also applies if the party leadership controls candidate selection, the model emphasizes the important point that party centralization is not necessary for commitment.

The pool of candidates contains politicians with different exogenously given reputations, mirroring the two factions of each party: candidates that have a reputation to be a moderate type ($M$) and candidates that have a reputation to be a type representing the party’s core constituency ($L$ in party LM and $H$ in party MH). The reputation of candidates is an informative but imperfect predictor of their true type. In other words, for a given reputation there is uncertainty about the true type of candidates. As a consequence, commitment based on candidate selection is probabilistic rather than deterministic.\footnote{This also preserves the asymmetry of the Iversen-Soskice model. Given non-regressivity, equilibrium behavior shares some similarities with valence models of party competition.} Formally, for party LM (MH) a candidate with a reputation of being a $M$-type is denoted by $\beta_{LM}$ ($\beta_{MH}$) and a candidate with the reputation of being a $L$ ($H$) type is denoted by $\alpha_{LM}$ ($\alpha_{MH}$), where $\beta_{LM}$ ($\beta_{MH}$) candidates are more likely to be $M$ than $L$ ($H$) types and $\alpha_{LM}$ ($\alpha_{MH}$) candidates are...
less likely to be $M$ than $L$ ($H$) types. In particular, I assume $Pr(M|\beta_{LM}) = Pr(M|\beta_{MH}) = Pr(L|\alpha_{LM}) = Pr(H|\alpha_{MH}) = \pi > 1/2$ and $Pr(L|\beta_{LM}) = Pr(H|\beta_{MH}) = Pr(M|\alpha_{LM}) = Pr(M|\alpha_{MH}) = 1 - \pi$.

For simplicity, assume that in the electoral stage the majority faction in each party chooses between full moderation ($x_P = 1$), which entails nominating candidate with the reputation of being $M$-types in both the marginal and the party’s safe district, and partial moderation ($x_P = 0$), which entails nominating a moderate candidate type in the marginal but not their safe district. Partial moderation is consistent with the majoritarian model of Iversen and Soskice (2006). As parties may use mixed strategies, let $\sigma_P$ denote the probability that party $P$ chooses full moderation. Full moderation entails an exogenous cost $c_P > 0$ for the majority faction. This is a reduced form to capture motivations other than policy outcomes. A natural interpretation is that each type of party member is also ideologically committed to a particular policy (Roemer, 1999). Full moderation means diluting the party’s traditional ideology. Thus, members do not like to select moderates unless they are moderates themselves.$^8$

After parties’ candidates are revealed, an election is held and voters decide which candidate to support in their district. The middle-income group $M$ in the marginal district decides which party wins a majority of seats. The simple post-electoral bargaining protocol captures that the majority party controls the legislative agenda and there are no exogenous restrictions on legislative voting. The agenda setter, which one can think of as the chief executive or responsible cabinet minister, is drawn at random from the elected members of the majority party and gets to make a take-it-or-leave it policy proposal. The policy passes

$^8$Another interpretation is that incumbents in safe districts have office motivations to resist replacement by candidates with a more centrist reputation.
if supported by a majority, otherwise there is no redistribution.

To summarize, the sequence of events is as follows:

1. The majority factions in parties LM and MH simultaneously choose their candidate selection strategies, \( \sigma_{LM} \) and \( \sigma_{MH} \), consisting of the probability of pursuing full moderation.

2. The realized values of \( \sigma_{LM} \) and \( \sigma_{MH} \) are observed and a general election is held. In each district, the candidate with a plurality of votes wins.

3. Nature randomly draws the agenda setter from the members of parliament of the majority party and reveals the types (\( L, M, \) or \( H \)) of the elected politicians.

4. The agenda setter proposes a net transfer policy, \( \mathbf{b} = (p^L, p^M, p^H) \) subject to the economic feasibility constraints.

5. If a majority of legislators supports \( \mathbf{b} \), it becomes policy; otherwise there is no redistribution.

The equilibrium concept is subgame perfect Nash equilibrium. As usual, equilibria where individuals use weakly dominated strategies (politicians voting for policies that result in strictly lower payoffs as other politicians are doing so as well) are excluded.

**Equilibrium and Implications for Redistributive Politics**

As illustrated by Figure 1, equilibrium outcomes of partisan competition over redistribution vary by income inequality and parties' cost of moderation. A formal statement and proof are in online Appendix S1 (Proposition 1). First, suppose the difference in taxable income between middle-income and high-income citizens is relatively small or the cost of moderation is sufficiently high for both parties (region I in Figure 1). This means that either
middle-income voters prefer party MH even when party LM makes more efforts to moderate (because commitment is only probabilistic and the structure of inequality makes the perils of left government more pronounced) or LM’s cost of moderation is simply too high. As a consequence, neither party has an incentive to pursue a costly commitment strategy. Hence, LM suffers from an observable electoral disadvantage and the policy outcome entails little redistribution to the poor in expectation. Second, a different political equilibrium emerges if the income difference between the middle and the top are sufficiently large and there are asymmetric cost with sufficiently low cost only for LM (region II in Figure 1). In this scenario, LM makes a greater effort to commit by recruiting moderates (including some safe districts) than MH and, as a consequence, has better electoral prospects. Redistribution is higher than in the previous case without costly commitment. The center-left party does not suffer from an electoral disadvantage as it has credibly moved to the center while its competitor has not. Thus, there is a positional but not an electoral anti-left bias of partisan competition. Third, consider the scenario where income differences between the middle and the high-income groups are sufficiently large and cost are sufficiently low for both parties (region III in Figure 1). In this situation, the equilibrium is in mixed strategies where both parties choose a moderate candidate selection strategy with positive probability (but strictly less than one). Electoral outcomes are uncertain ex ante and there is no generic electoral disadvantage for LM. Expected redistribution is higher than in the case without costly commitment but lower than in the case with asymmetric commitment only by the left.

Effects of Endogenous Commitment

The theory has several novel implications. First, endogenous commitment implies that center-left parties need not suffer from an observable electoral disadvantage in majoritarian electoral systems. Several prominent theories in the literature maintain that majoritarian institutions entail an electoral bias against the left. This means that even if there is a
Figure 1: Equilibria in the majoritarian model. The figure shows how parties’ electoral strategies vary by top income inequality (i.e., the ratio of taxable income of high and middle-income voters) and the cost of recruiting moderate candidates. The probability that the center-left LM (center-right MH) party pursues a costly strategy of recruiting middle-income types to commit to a moderate redistributive platform is denoted by $\sigma_{LM}$ ($\sigma_{MH}$).
relatively balanced distribution of policy preferences, the electoral system induces electoral outcomes that favor the right. As discussed before, the model of Iversen and Soskice (2006, Proposition 1) implies that, given the absence of commitment, strategic voting by the middle class favors the right as the least-bad alternative to a left government that may break its promises and make the middle class a net loser of the redistributive game.\footnote{A related theory developed by Rodden (2012) is centered on political geography.} Hence, on average, center-left party loose elections under majoritarianism. Iversen and Soskice (2006, p. 170) point out that their analysis suggests that left parties may have incentives to find a commitment device. The model develop here captures and further explores that intuition. It shows that endogenous commitment using parliamentary candidates implies that center-left parties may escape their electoral dilemma. This holds even if center-left parties are disadvantaged a priori by the absence of external commitment and policy constraints. Furthermore, the existence of an electoral disadvantage is conditional on context. When income inequality between the middle and the top is sufficiently high and parties’ moderation cost are sufficiently low, center-left parties will have incentives to pursue a strategy of costly commitment. As a consequence, they tend to be electorally competitive. When income inequality between the middle and the top is low or parties’ cost are sufficiently high, costly commitment is not an equilibrium strategy and center-left parties will systematically lose elections. This logic can account for why center-left parties in majoritarian systems may exhibit relatively long spells of electoral success as well as long spells in opposition, which does not follow straightforwardly from the theories discussed above or Downsian’ models of party competition that assume commitment and imply (near-)convergence.

Second, an increase in the taxable income of the rich relative to the middle class increases the probability that the center-left party in majoritarian systems moves toward the center.\footnote{Probability refers to the larger set of parameters for which the center-lefty party pursues
The moderation effect seems counterintuitive as partisan and Downsian’ models assuming
commitment suggest the opposite. The standard partisan approach (Hibbs, 1977) suggests
that left parties will respond to rising inequality by moving toward the left, re-emphasizing
egalitarian policies favored by their core constituencies. If middle class voters would benefit
from higher taxes on the rich, then canonical models of party competition that assume
commitment also suggest that all parities should respond to higher top inequality by moving
to the left, without necessarily converging to the same position (Calvert, 1985). Given the
commitment problem, however, the right-ward move follows rationally from the center-left’s
incentives to achieve endogenous commitment. An increase in the distance of middle to
top incomes increases the willingness of the middle income group to support the center-left
party. But this is conditional on the party’s ability to appease middle class fear’s of being
soaked at the expense of the poor. With higher inequality between the rich and there merely
well-off, the redistributive gains of a moderate strategy increase for the left, and so it will
be more willing to select candidate with a moderate reputation. Note that if top inequality
increases because the high income group becomes richer relative to everybody else, then the
left’s moderation can occur despite a rise in the tax base for redistribution (different from
Barth, Finseraas and Moene 2014).

Third, what parties do in non-marginal districts matters for the outcome in marginal
(swing) districts. Indeed, parties may have incentives to nominate moderates even in safe
districts to achieve more commitment in marginal districts. This does not follow from stan-
dard political economy models of majoritarian systems in general (Persson and Tabellini,
2000) or models of costly candidate selection by parties in particular (Galasso and Nan-
icini, 2011), where parties tend to concentrate costly (high quality) candidates in marginal

a strategy of costly moderation.
districts. The importance of candidate choice in non-marginal districts is driven by a form of strategic voting that is induced by the commitment problem: the vote choice of the pivotal voter depends on the overall candidate selection by the main parties.

Comparison with PR

The outcome of redistributive conflict under majoritarian institutions analyzed so far can also be compared to redistribution under PR. A clear benchmark is provided by the PR model in Iversen and Soskice (2006), which is based on the same general assumptions about preferences and policy. Key features of PR in their model are that there are group-based parties and, as no party can expect to win a majority, policy is determined through a process of coalition bargaining between parties. Crucially, coalition government implies that left parties in government are subjected to a veto by a coalition partner. This makes center-left coalition governments attractive to middle-income voters. Coalition government, put simply, results in endogenous commitment due to mutual vetoes (Stasavage 2002; Tsebelis 2002). Comparing redistributive politics across the majoritarian and the PR system, the analysis concurs with that of Iversen and Soskice (2006) in that government partisanship is a key mechanism though which electoral rules affect economic policy. There tends to be an anti-left bias in majoritarian compared to proportional electoral rules that leads to lower redistribution (for a formal statement, see Proposition 2 in Appendix S2). Extending their framework to allow for endogenous commitment highlights that the anti-left bias may take two distinct forms and is conditional on features of the political economy. The anti-left bias has a positional as well as an electoral component. Observed differences in government partisanship and redistribution across electoral rules can arise because left parties competing under majoritarian rules are (i) less electorally successful compared to the left parties competing under PR (the mechanism emphasized by Iversen and Soskice 2006) or (ii) are credibly more centrist (the endogenous commitment mechanism). Studies using center of gravity measures of govern-
ment partisanship have documented a robust difference in partisanship across electoral rules (Iversen and Soskice 2006; Döring and Manow 2013; also see Funk and Gathmann, 2013). Interpreting this observed difference as the effect of electoral bias most likely overstates the electoral component as it does not account for endogenous commitment, which entails a positional bias but not necessarily an electoral bias.

The comparison also implies that there should be observable differences in candidate selection across electoral rules. Holding other things equal, left parties have more incentives to nominate moderate candidates in majoritarian than in proportional electoral systems, where coalition government provides endogenous commitment. By the same logic, the effect of income inequality on the position of the center-left party, discussed above for the majoritarian setting, should vary by electoral system. While an increase in inequality between the rich and the middle classes increases the incentives for the left to moderate under majoritarianism, this is not be the case under PR.

Extension

An inter-temporal issue ignored until now is that current redistribution influences the future distribution of pre-fisc income and thereby also shapes future conflict over redistribution. To explore how such dynamic considerations shape endogenous commitment, I have analyzed a dynamic version of the static model, in the spirit of models like that of Besley and Coate (1998) (see online Appendix S2). There are two electoral periods. After the first period, there is a shock to the distribution of incomes that depends on the size of the redistributive budget in the first period. The economic assumption is that higher income redistribution in the first period leads to lower income growth and a more compressed pre-fisc income distribution in the second period. Reassuringly, the results from the static model still go through in this repeated elections framework. But accounting for dynamics generates additional insights.

The commitment problem does not go away even even if policymakers in the first period
put a large weight on second-period utility and anticipate how current redistributive policy affects future income distribution and voting behavior. Consider the scenario where the center-left party LM wins the election in the first period and there is a majority $L$-types in the legislature. While policymakers are forward-looking and understand the economic and political cost of redistribution, a $L$-majority will generally choose more redistribution to the low-income group than preferred by the pivotal middle-income voter $M$ (Proposition 3 in S2).

Intuitively, first-period policy in the dynamic framework can be more moderate than in the static model if policymakers put a high weight on the second period. Politicians understand that high redistribution in the first period can undermine economic efficiency and political support among the middle class in the second period. In response, the $L$-majority will choose to moderate current redistribution if it cares sufficiently about the future - though not enough to eliminate the commitment problem. The reason is that $M$ voters cannot credibly threaten to punish all deviations from their most preferred policy.

Dynamic considerations can actually increase the commitment problem faced by LM - because forward-looking middle-income voters ($M$) now have an additional motivation to fear a left government: not just the current loss imposed by redistribution to the poor, but also the reduction in future income as a consequence of high redistribution. As a consequence, they can be more reluctant to support the LM party in the dynamic compared to the static setting. Appendix S2 describes such an equilibrium (Proposition 4). The severity of the commitment problem is indicated by the relative income threshold below which middle-class voters have a dominant strategy to vote for the center-right party (MH). In this equilibrium of the dynamic game, the endogenous threshold is to the right of the threshold of the static game, $\frac{2(1-x)^2}{\pi-1/2}$, depicted in Figure 1. Up to a point, commitment becomes more of a problem for the left party as voters attach more weight to the future.
Evidence

The empirical section confronts two novel implications of the theoretical model with evidence from two distinct data sets.

Analysis of Party Positions

The first analysis focuses on how within-country changes in top income inequality affect party positions and whether this relationship depends on electoral rules. The basic statistical specification is as follows:

\[
Y_{it} = \alpha_i + \beta_1 TI_{it} + \beta_2 PR_{it} + \beta_3 TI_{it} \times PR_{it} + \beta_4 X_{it} + \epsilon_{it}
\]  

(2)

where \(Y_{it}\) is the economic left-right position of the main left party (lower values indicate a more leftist position) in country \(i\) at election year \(t\), \(\alpha_i\) is a country specific intercept (capturing possibly unobserved time-invariant confounders), \(TI_{it}\) is a measure of top income inequality, \(PR_{it}\) is an indicator for proportional representation, \(X_{it}\) stands for various control variables, and \(\epsilon_{it}\) is the error term. The interaction term \(TI_{it} \times PR_{it}\) allows the impact of inequality to vary by electoral system. The prediction of the theory is that an increase in top income inequality leads the left to move to the right in majoritarian systems (\(\beta_1 > 0\)). Under PR, on the other hand, inequality does not lead to a right-ward move; in terms of coefficients, this means, first, that there is a significant negative interaction term (\(\beta_3 < 0\)) and that the marginal effect of inequality in PR systems is at least weakly negative (\(\beta_1 + \beta_3 \leq 0\)).

The analysis covers 16 advanced industrial parliamentary democracies between 1950 and 2010 for which the inequality data are available: Australia, Canada, Denmark, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. In line with the theory as well as the study
of Pontusson and Rueda (2010), the focus is on the position of the main left party.\footnote{See Appendix S3 for a list of the parties, sources for all variables, and descriptive statistics.}

**Measurement**

The dependent variable is measured using new estimates of the publicly stated economic left-right position based on party manifestos from Elff (2013), where larger values indicate a more rightist (i.e., less supportive of redistribution and state intervention) position. The positions were estimated by Elff (2013) using data coded by Comparative Manifesto Project (CMP, see Budge et al. 2001) for a large set of countries since World War II. As has been extensively described elsewhere (Budge et al., 2001), CMP has manually coded each quasi-sentence in a given electoral program into one of 56 policy categories. The left-right summary indicator provided by CMP, which is based on the relative frequency of sentences in left versus right policy categories, has been widely used in the literature on party positions and partisan effects on policy. Its time-coverage is unique and it correlates fairly highly with alternative measures (such as expert judgements) available for a subset of country-years. The measure developed by Elff (2013) deals with important limitations of the CMP index. His scaling procedure, which is related to the estimation of ideal points in legislatures, disentangles policy positions from the relative frequency or salience of an issue. It also relaxes the assumption that all coding categories are a priori determined to be either left or right. The variables does not attempt to measure whether a party’s stated position is perceived to be credible by voters. But it allows researchers to examine whether left parties make more or less leftist appeals in response to changing inequality.\footnote{Research on measuring party positions is very active (see the discussion by Elff 2013).}
To measure income inequality, I draw on the World Top Incomes Database painstakingly put together by economists using income tax records over the last decade (Atkinson and Piketty 2007; Atkinson and Piketty 2010; for an overview and thorough discussion of methodological issues, Atkinson, Piketty and Saez 2011). The database covers a much larger period than usual survey-based measures and it contains a measure that is close to the theoretical model. The model emphasizes a particular facet of inequality, namely, how much better the rich are doing compared to the middle classes. What matters for the electoral incentives of left parties is shaped to an important degree by the dynamics at the top of the income distribution. For this purpose the database is especially well suited. Survey-based measures, by comparison, usually have problems capturing the top of the distribution due to top-coding and relatively small sample sizes. The inverted Pareto coefficient characterizing top incomes is the main measure of income inequality. The Pareto law is usually considered to be a good approximation of the upper part (say top 10 %) of the observed income distribution. Reflecting the theoretical logic, an increase in the coefficient indicates higher top shares and generally higher income inequality. For instance, if the coefficient is 2 the average income of individuals with income above a threshold, say $100,000, is twice the threshold value (Atkinson, Piketty and Saez, 2011). In the data set, it varies from 1.3 to 2.9.\footnote{This implies a somewhat more selective and probably more common definition of the high-income group than the one based on equal group size used for convenience in the theory. The results are robust to using alternative operationalizations (Appendix S3).} The data are fairly homogenous within countries and researchers contributing to the database have made great efforts to harmonize them across countries as well. Including country fixed effects in the estimation absorbs possibly remaining country-specific variation in the measure. The variable has been rescaled as deviations from the overall mean to allow for an
easier interpretation of the interactive specification.

Thanks to a cumulative research effort, classifying electoral rules has become a largely uncontroversial affair. Following Iversen and Soskice (2006), for each election the electoral system for the legislature (lower house under bicameralism) is coded as PR or majoritarian (missing years were updated).\textsuperscript{14} As is well-documented in the literature, most of the variation in electoral rules is cross-sectional.

The literature discusses several other factors that may affect income inequality, electoral rules and party positions. Country fixed effects account for unobservable time-invariant country heterogeneity. For example, historical legacies, varieties of capitalist institutions, or cultural differences may drive trends in economic inequality and party positions; and the literature on the choice of electoral rules debates whether they are endogenous to redistributive politics.\textsuperscript{15} To account for unobserved time effects, the analysis includes a linear time trend; some specifications also include period effects (1950-1972, 1973-1989, 1990-2010) or country-specific trends. The analysis also includes time-varying control variables whose omission may bias the results. Income per capita captures difference in the capacity to redistribute and, according to some (Wagner’s Law), preferences for public spending. Voter turnout is often linked to the mobilization capacity of left parties (Pontusson and Rueda, 2010). Population size is include because size has been argued to impede the organization of working class interests (Alesina and Glaeser, 2004). A prevalent view is that economic globalization, which covaries with income inequality, increases demand for public spending; at the same time,\textsuperscript{14}See Appendix S3 for details.\textsuperscript{15}For instance, Boix (1999) emphasizes short-run calculations by old party elites and Cusack, Iversen and Soskice (2007) the legacy of labor market institutions.
it is held to constrain policy choices and thus party positions (Ward, Ezrow and Dorussen 2011). The measure used here is trade openness as a percentage of GDP. The percentage of population over 65 is a common proxy for the demand for social insurance. All time-varying right-hand side variables except electoral rules are measured as the average for the years from the last to the current election.

Results

The results from the basic specification are summarized in Table 1. While all model include country fixed effects, additional control variables are added sequentially. Throughout, the effects of the main variables of theoretical interest, income inequality and its interaction with PR, are fairly precisely estimated and have the predicted sign. The positive coefficient on the inequality variable indicates that an increase in top income inequality is associated with a right-ward move of the main left party in majoritarian systems. The magnitude of the negative coefficient on the interaction term indicates that the impact of inequality is practically reversed in PR systems. These estimates are consistent with the theoretical expectations and are not easily accounted for by existing theories. Adding control variables and time effects does not change the results very much. The results are politically relevant. Model (4), which includes time-varying controls as well as country-specific time trends, suggests that a one standard deviation increase in top income inequality causes left parties in majoritarian settings to move to the right by about one standard deviation. The sum of the coefficients on inequality and the interaction term suggests that the marginal effect of inequality is negative (about −0.51) under PR, consistent with the traditional partisan view, though not significant at conventional levels \((p = 0.27)\). What is clear is that left parties

---

16 Standard errors are clustered by country to allow for serial correlations. As shown in Appendix S3, using a lagged dependent variable yields similar results.
competing under PR do not move to the right in response to rising top inequality whereas left parties in majoritarian systems do.

The results are robust to various alternative specifications (see online Appendix S3). This includes using alternative measures of all three main theoretical concepts: top income inequality, electoral systems, and party positions. The sensitivity analysis also adds additional control variables sometimes used in related studies of party positions (union density, EU integration, incumbency, existence of a left competitor) and a lagged dependent variable. I also control for the centralization of candidate selection. Using data on the Gini coefficient of market inequality available from 1960 onwards, I rule out that the effect of top inequality is driven by changes in other parts of the income distribution.

One remaining concern is that top income inequality is endogenous to redistributive politics in a way that is not captured by the fixed effects, controls, and time trends. Given the observational nature of the data, it is impossible to completely rule out time-varying country-specific confounders. To address this possibility, an additional analysis exploits plausibly exogenous variation in top income inequality based on the evolution on top inequality in the United States. The idea is that the increase in top income inequality is driven to an important party by a dramatic increase in executive pay and this occurred first in the U.S. (Atkinson, Piketty and Saez, 2011). This rise was not driven by politics in the smaller economies outside the U.S. Through explicit compensation benchmarking and partially integrated market for top managers (DiPrete, Eirich and Pittinsky, 2010), the dramatic rise in executive compensation in the U.S. led to exogenous competitive wage pressures on executive pay other developed countries. The instrumental variable estimates confirm the results from the basic panel analysis that under majoritarian rules increased top inequality leads left parties to move to the right (S3, Table 8).

Altogether, the macro-level results are consistent with a novel implication of the theoretical model and run counter to conventional views of partisan conflict over redistribution.
Table 1: Panel Estimates of the Conditional Impact of Top Income Inequality on the Economic Position of the Main Left Party in 16 Democracies 1950-2012

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Income Inequality (TI)</td>
<td>0.98**</td>
<td>0.91**</td>
<td>0.79**</td>
<td>0.90**</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.31)</td>
<td>(0.37)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>Proportional Representation (PR)</td>
<td>-0.01</td>
<td>-0.15</td>
<td>-0.16</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>PR × TI</td>
<td>-1.13**</td>
<td>-1.21**</td>
<td>-1.18**</td>
<td>-1.41**</td>
</tr>
<tr>
<td></td>
<td>(0.52)</td>
<td>(0.34)</td>
<td>(0.36)</td>
<td>(0.43)</td>
</tr>
<tr>
<td>Income per capita (in 1000 USD)</td>
<td>-0.03</td>
<td>-0.04</td>
<td>-0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>Voter Turnout</td>
<td>-0.05</td>
<td>0.12</td>
<td>-0.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.49)</td>
<td>(0.44)</td>
<td>(0.96)</td>
<td></td>
</tr>
<tr>
<td>Trade Openness</td>
<td>-0.02**</td>
<td>-0.01**</td>
<td>-0.02**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>Population Over 65 (%)</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.13)</td>
<td></td>
</tr>
<tr>
<td>Log of Population</td>
<td>-1.63**</td>
<td>-1.54**</td>
<td>-3.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.27)</td>
<td>(2.52)</td>
<td></td>
</tr>
<tr>
<td>Time Trend (Year)</td>
<td>0.05**</td>
<td>0.04**</td>
<td>0.09**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.04)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>No</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country-specific time trend</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Period indicators</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Countries</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Observations</td>
<td>228</td>
<td>228</td>
<td>228</td>
<td>228</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.071</td>
<td>0.334</td>
<td>0.344</td>
<td>0.358</td>
</tr>
</tbody>
</table>

Standard errors in parentheses are clustered by country. The dependent variable is the left-right position of the main left party as estimated from manifestos by Elff (2013). All models include country fixed effects.

* $p < 0.10$, ** $p < 0.05$ (two-tailed tests)
They also qualify the previous finding that higher inequality generally induces left parties to move to the right (Barth, Finseraas and Moene, 2014). Accounts emphasizing the ability of rich elites to influence voting and party behavior through the media and political donations (Hacker and Pierson, 2010) also do not explain the institution-varying effect of income inequality.\textsuperscript{17}

**Analysis of Candidate Selection**

To provide some evidence on the candidate selection mechanism, I analyze decentralization in Britain as a natural experiment that compares the political profile candidates nominated by the same parties under alternative electoral institutions. The analysis exploits the fact that the 1997 elections for the lower house of the British parliament (House of Commons) in Scotland and the 1999 election of the new subnational Scottish parliament were held in the same districts but under different electoral systems. The established system based on single-member districts and plurality voting was maintained for the House of Commons, while mixed-member proportional representation (MMP) was adopted for the unicameral Scottish parliament.\textsuperscript{18}

The Scottish parliament has the power to make laws in a wide range of policy areas. Mirroring the executive-legislative institutions in Westminster, the Scottish government is responsible to its parliament. The MMP system used to elect the 129 members of the

\textsuperscript{17}Accordingly, higher top inequality increases the stakes of redistributive conflict and thus should lead elites across institutions to use their financial clout to blunt the left’s redistributive zeal.

\textsuperscript{18}Dewan and Spirling (2011) compare legislative behavior in these two parliaments.
Scottish parliament gives voters two votes, one for the familiar single-member district, where the candidate with a plurality of votes wins, and a second vote for a larger multi-member district where voters choose between closed party lists. Crucially, the second vote is used to allocate additional seats to parties that have received the least single-member districts seats relative to their total votes in the region, thereby increasing the proportionality in the translation of votes into seats. All but two of the 73 single-member districts for the Scottish parliament in 1999 were identical to the ones used for Westminster in 1997. The analysis focuses on the candidates nominated by the three parties three main center-left parties (Labour, Liberal Democrats, Scottish National Party) in these 71 districts (Appendix S4 provides further background information and data sources).

**Empirical Strategy**

In this institutional setting, the logic of endogenous commitment predicts that compared to its competitors Labour should exhibit the largest difference in its candidate selection strategy between the two electoral systems. Among the three main parties in Scotland at the time, Labour was the only party that had a chance of winning a majority of seats in the national election and thus faced the strongest incentives to moderate. This constraint was relaxed in the election for the Scottish parliament. Because of MMP, Labour was expected to fall short of a parliamentary majority and form a coalition government. Thus, there was less need to rely on moderate candidates to achieve commitment. The other parties might also move to the left by nominating less moderate candidates when competing for the Scottish parliament, given the presumed left-ward change in identity of the median voter at large across elections. However, the key observable implication is that Labour should overcompensate relative to its competitors. This clearly differs from the standard Downsian view, which assumes commitment and is concerned the overall party position and not individual candidates as such. It suggests, if anything, that parties move in parallel to
changes in the center of gravity, not that Labour is more responsive. To clarify, the basic estimation equation can be written as follows:

\[ C_{it} = \alpha_i + \beta_1 L A B_{it} + \beta_2 M M P_{it} + \beta_3 L A B_{it} \times M M P_{it} + \epsilon_{it} \]  

(3)

where \( C_{it} \) is a measure of the ideological orientation (higher values indicating more left-wing views) of candidate \( i \) and election \( t \) (Westminster 1997 or Scottish Parliament 1999), \( L A B_{it} \) is an indicator for whether the candidate belongs to the Labour party, \( M M P_{it} \) is an indicator for the 1999 election using MMP, \( \alpha_i \) is a district specific intercept, and \( \epsilon_{it} \) an idiosyncratic error. The Liberal Democrats and Scottish National Party are included as the control group. The district-specific intercept captures observable and unobservable district characteristics. Given the short time span between the elections, this likely captures the district median voter.\(^{19}\) The interaction term tests the implication of the commitment logic that Labour overcompensated relative to its competitors by undertaking a larger shift to leftist candidates (i.e., \( \beta_3 > 0 \)). In total, Labour can be expected to move to the left under MMP (\( \beta_2 + \beta_3 > 0 \)). If parties move in parallel, as suggested by the Downsian logic, then there should be such interaction effect (\( \beta_3 = 0 \)).

Finding a good proxy for candidates’ left-right orientation is tricky. Stated labor union membership is probably the best available one. Studies consistently show that union members are more supportive of redistribution and social insurance than non-members and unions were often considered a constraint on the ability of left parties to credibly move toward the center. This was widely held to be true for Labour in Britain during the 1980 and 1990s (Kitschelt 1994, pp. 249-252). ‘New’ Labour politicians portrayed unions as being part of ‘old’ Labour and unions were seen as a lobby for leftist candidates. An added advantage

\(^{19}\) Census data are only available 1991 and 2001.
is that union membership is not restricted to working class occupations, which are a small and shrinking part of the labor force.\textsuperscript{20}

Equation 3 is estimated with a linear probability model (LPM) as well as the fixed effects logit (FEL) estimator (e.g., see Wooldridge, 2002, pp. 490-3). The advantage of the LPM is that it can easily incorporate district fixed effects and provides substantively meaningful estimates. Robust standard errors are calculated to deal with heteroskedasticity. The appeal of the FEL model is that it explicitly models the discrete nature of the dependent variable and can also account for unobserved district heterogeneity by conditioning on them. Two drawbacks are that it drops districts for which there is no change in the outcome variable across elections and it does not allow the calculation of marginal effects or first differences.

Results

The estimation results are displayed in Table 2. Model (1) does not include district fixed effects. Thus, the regression intercept corresponds to the relative frequency of union members among the Liberal Democrats and Scottish National Party in Westminster (14%). Unsurprisingly, Labour’s historical and organizational ties to unions, membership is considerably higher among Labour candidates for Westminster (by about 37%). The coefficient on the MMP indicator shows that union-membership hardly varies across electoral systems for the non-Labour parties. The interaction term indicates that Labour, relative to the other parties, has significantly increased (by about 19%) the share of union candidates under the more proportional electoral rules in place for the Scottish parliament. Those findings hold when district fixed effects are added (Model (2)). The coefficients from the FEL estimator convey the same pattern (Model (3)). Thus, the estimates are in line with theoretical expectations.\textsuperscript{20}

Candidate surveys contain relevant items but do not allow researchers to identify candidates to match them to their district.
Table 2: Effect of Electoral Rules on Candidate Selection in Scotland

<table>
<thead>
<tr>
<th></th>
<th>(1) LPM</th>
<th>(2) LPM</th>
<th>(3) FEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labour</td>
<td>0.37**</td>
<td>0.37**</td>
<td>2.15**</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.06)</td>
<td>(0.40)</td>
</tr>
<tr>
<td>Mixed Member PR (MMP)</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.45</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.38)</td>
</tr>
<tr>
<td>Labour × MMP</td>
<td>0.19**</td>
<td>0.18**</td>
<td>1.09**</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.55)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.14**</td>
<td>0.37**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.18)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>413</td>
<td>413</td>
<td>333</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.236</td>
<td>0.295</td>
<td></td>
</tr>
<tr>
<td>Pseudo $R^2$</td>
<td></td>
<td></td>
<td>0.374</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses. The dependent variable is an indicator for union membership. Models (1) and (2) are linear probability models and (2) includes district fixed effects. Model (3) is a fixed effects logit model.

* $p < 0.10$, * $p < 0.05$ (two-tailed tests)

Relative to its competitors, the change from majoritarian to proportional electoral rules has prompted Labour to significantly increased its share of candidates with a union background. Given the more conservative estimate in Model (2), the relative increase (measured by the coefficient on the interaction term, $\beta_2$) is about 18%. Also accounting for the negative (though insignificant) effect on the constituent MMP term, the total increase ($\beta_2 + \beta_3$) for Labour is 13% ($p = 0.08$). There is no evidence the parties in the control group have significantly changed their candidate selection strategy across electoral systems.$^{21}$

$^{21}$Appendix S4 shows that these results are robust to adding individual-level controls and excluding districts won by the SNP.
Concluding Remarks

This paper has provided a new theoretical perspective and brought new evidence to bear on how conflict over redistribution plays out in democracies. Together, the results make the case that endogenous commitment by political parties is a driving force in redistributive politics and that candidate selection is an important but neglected commitment device. The main theoretical innovation is to consider how parties’ efforts to solve the commitment problem influences electoral strategies, election outcomes, and policy. This has lead to several new testable implications, for example, concerning the institution-varying responsiveness of parties to income inequality. Using two different data sets, the empirical analysis has found evidence consistent with two key implications of the theoretical logic. Additional implications of the theory merit testing. One further empirical task, for example, is to try to unbundle the difference in government partisanship observed across electoral institutions into its positional and electoral component. Broadening the analysis of commitment beyond reputation also matters for the literature on the determinants of party policy positions and opens up questions about the underpinnings of credible positions.

References


ON-LINE APPENDIX

February 2, 2015

Abstract

This online appendix provides supporting information for the paper “Endogenous Credible Commitment and Party Competition Over Redistribution Under Alternative Electoral Institutions.” Appendix S1 contains proofs for the theoretical model (equilibrium, comparison with PR). Appendix S2 solves the theoretical extension informally discussed in the paper. Appendix S3 contains additional information for the cross-national empirical analysis of party positions. In particular, it provides more details on the data and reports the results of the sensitivity analysis briefly discussed in the paper. Appendix S4 provides background information (including sources for the data) and additional results for the empirical analysis of candidate selection in Great Britain.
S1. Proofs for Theoretical Model

Equilibrium

Proposition 1. In equilibrium, parties’ candidate selection strategies, the probability that LM wins the election and the expected size of the redistributive budget, $R \equiv E(\sum_j T_j)$ are as follows:

1. Suppose $\frac{T_H}{T_M} \leq \frac{2(1-\pi)^2}{\pi-1/2}$ or $c_{LM} \geq \left( \frac{3}{4} - \frac{\pi}{2} \right) T_H + (1-\pi)^2 T_M$
   - $\sigma^*_{LM} = \sigma^*_{MH} = 0$
   - $Pr(\text{LM wins}) = 0$
   - $R = \frac{1}{2} T_H$

2. Suppose $\frac{T_H}{T_M} > \frac{2(1-\pi)^2}{\pi-1/2}$, $c_{LM} \leq \left( \frac{3}{4} - \frac{\pi}{2} \right) T_H + (1-\pi)^2 T_M$, and $c_{MH} \geq (1-\pi) T_H$
   - $\sigma^*_{LM} = 1, \sigma^*_{MH} = 0$
   - $Pr(\text{LM wins}) = 1$
   - $R = T_H + (1-\pi)^2 T_M$

3. Suppose $\frac{T_H}{T_M} > \frac{2(1-\pi)^2}{\pi-1/2}$, $c_{LM} \leq \left( \frac{3}{4} - \frac{\pi}{2} \right) T_H + (1-\pi)^2 T_M$, and $c_{MH} \leq (1-\pi) T_H$
   - $\sigma^*_{LM} = 2 \left( \pi - 1/2 + \frac{c_{HM}}{T_H} \right), \sigma^*_{MH} = 1 - \frac{c_{LM}}{\left( \frac{3}{4} - \frac{\pi}{2} \right) T_H + (1-\pi)^2 T_M}$
   - $Pr(\text{LM wins}) = \sigma^*_{LM} (1 - \sigma^*_{MH})$
   - $R = \left( \frac{1}{2} (1 + \sigma^*_{LM})(1 - \sigma^*_{MH}) + \sigma^*_{MH} \pi \right) T_H + \sigma^*_{LM} (1 - \sigma^*_{MH})(1-\pi)^2 T_M$

PROOF

The proof proceeds by backward induction.
Policymaking. In the policymaking stage, consider legislative voting for a given proposal $b = (p^L_b, p^M_b, p^H_b)$. Legislators of type $J$ accept any proposal $b$ that specifies a (weakly) positive net transfer to their group $J$, $p^J_b \geq 0$. The optimal proposal by the agenda setter of type $J$ maximizes the net transfer to the group $J$ subject to legislators’ acceptance constraints and the policy constraints. First, consider an agenda setter of type $M$. If there is no majority of $H$ types, $|H| < \frac{N+1}{2}$, $M$’s optimal proposal is its most-preferred policy, $b^{M^*} = (\frac{T^H}{2}, \frac{T^H}{2}, -T^H)$, as it will be accepted by both $M$ and $L$ legislators. Else, $H$ legislators will veto any redistributive bill so any proposal is a best response. Second, suppose there is a $L$ agenda setter. If there is a majority of $L$ legislators, by the same logic, $|L| \geq \frac{N+1}{2}$, $b^{L^*} = (\bar{T}, -\bar{T}^M, -\bar{T}^H)$. If $|L| < \frac{N+1}{2}$ and $|L| + |M| \geq \frac{N+1}{2}$, $L^* = (T^H, 0, -T^H)$. Else, $|H| \geq \frac{N+1}{2}$, $H$ legislators will veto any redistributive bill so any proposal is a best response. Third, $H$ agenda setter. $H$ proposes the ideal policy of the high-income group, $b^{H^*} = (0, 0, 0)$. Taken together, the legislative outcome for a given type of agenda-setter and composition of the legislature are summarized by Lemma 1.

Lemma 1. For a given post-electoral distribution of preference types in the legislature, the outcomes of the legislative process is as follows:

<table>
<thead>
<tr>
<th>Seat allocation in legislature</th>
<th>Agenda setter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>L</td>
</tr>
<tr>
<td>$</td>
<td>L</td>
</tr>
<tr>
<td>$</td>
<td>H</td>
</tr>
</tbody>
</table>

General election. Lemma 2 summarizes the best-responding voting behavior of the pivotal voter $M$. To see why it holds, compare $M$’s expected utility of voting for $LM$ (given its choice of candidates $x_{LM}$) with that of voting for $MH$ (given $x_{MH}$). Note that by
the sequence of actions, the pivotal voter $M$ observes the (realized) choice of candidates $(x_{LM}, x_{MH})$ even if parties use non-degenerate mixed strategies.

**Lemma 2.** The median voter $M$ in the pivotal districts votes for the candidate of party MH unless $x_{LM} = 1, x_{MH} = 0$ and $\frac{\bar{T}_H}{\bar{T}_M} > \frac{2(1-\pi)^2}{\pi-1/2}$, in which case $M$ supports LM’s candidate.

By Lemma 1 and the random selection of the setter from the majority party in the legislature, we know the policy outcome for a given preference configuration in the legislature. At the time of the election, $M$ faces uncertainty about how her vote choice translates into policy because candidates’ policy preferences are private information. Recall that candidates’ reputation is an informative but noisy signal of their true type, where $\pi$ parameterizes the informational value of reputation. In the case where LM wins the election thanks to $M$’s support, let $Pr(a = M|x_{LM}, LM)$ denote the probability that the agenda setter will be a moderate type ($a = M$) given LM’s choice of candidates $x_{LM}$. In this case, the resulting policy corresponds to the ideal point of $M$ ($p = (\frac{\bar{T}_H}{2}, \frac{\bar{T}_M}{2}, -\bar{T}_H)$). Similarly, let $Pr\left(\#L \geq \frac{N+1}{2} \text{ and } a = L|x_{LM}, LM\right)$ denote the probability that the agenda setter is of type $L$ and there is a majority of $L$ types in the legislature given LM wins the election with candidates $x_{LM}$. In this case, by Lemma 1, policy correspond to $L$’s ideal policy ($p = (\bar{T}, -\bar{T}_M, -\bar{T}_H)$). Finally, $Pr\left(\#L < \frac{N+1}{2} \text{ and } a = L|x_{LM}, LM\right)$ denotes the probability that there is a $L$ agenda setter but no majority of $L$ legislators given that LM wins based on $x_{LM}$, which requires $|L| < \frac{N+1}{2}$ and $|L| + |M| \geq \frac{N+1}{2}$. Then, per Lemma 1, policy corresponds to $p = (\bar{T}_H, 0, -\bar{T}_H)$. Hence we can write $M$’s expected utility from voting for LM as

$$EU^M(vote \ LM|x_{LM}) = Pr(a = M|x_{LM}, LM) \left(y^M + \frac{\bar{T}_H}{2}\right) + Pr\left(\#L \geq \frac{N+1}{2} \text{ and } a = L|x_{LM}, LM\right) \left(y^M - \bar{T}_M\right) + Pr\left(\#L < \frac{N+1}{2} \text{ and } a = L|x_{LM}, LM\right) \left(y^M\right).$$
By the same logic, M’s expected utility from voting for MH can be written as

\[
EU^M(\text{vote MH}|x_{MH}) = Pr(a = H|x_{MH}, MH)(y^M) \\
+ Pr\left(\#H \geq \frac{N + 1}{2} \text{ and } a = M\big|x_{MH}, MH\right)(y^M) \\
+ Pr\left(\#H < \frac{N + 1}{2} \text{ and } a = M\big|x_{MH}, MH\right)(y^M + \frac{\bar{T}H}{2})
\]

Given that there are three districts and each party can at most win two, expected utilities simplify to

\[
EU^M(\text{vote LM}|x_{LM} = 1) = y^M + \frac{\bar{T}H}{2} - (1 - \pi)^2 T^M \\
EU^M(\text{vote LM}|x_{LM} = 0) = y^M + \frac{1}{2} \frac{\bar{T}H}{2} - \pi(1 - \pi) T^M \\
EU^M(\text{vote MH}|x_{MH} = 1) = y^M + \frac{\bar{T}H}{2} \\
EU^M(\text{vote MH}|x_{MH} = 0) = y^M + \frac{1}{2} \frac{\bar{T}H}{2}
\]

M chooses the party whose offer of candidates yields higher expected utility. It is straightforward to see that if \(x_{LM} \leq x_{MH}\), then M is always strictly better off voting for MH. If \(x_{LM} > x_{MH}\), then M continues to strictly prefer MH as long as \(\frac{\bar{T}H}{T^M} < \frac{2(1 - \pi^2)}{\pi - 1/2}\). M strictly prefers LM if \(\frac{\bar{T}H}{T^M} > \frac{2(1 - \pi^2)}{\pi - 1/2}\) and is indifferent if \(\frac{\bar{T}H}{T^M} = \frac{2(1 - \pi^2)}{\pi - 1/2}\). In the latter case I assume that M votes for MH as this is in the spirit of the commitment problem and reduces notation.

**Candidate selection and equilibrium outcomes.** At the candidate selection stage, parties anticipate the best response of the pivotal voter (Lemma 2) and legislators (Lemma 1). Each party chooses an electoral strategy, summarized by the probability of pursing a moderate candidate selection strategy, that is a best response given the other party’s strategy. First, consider LM’s decision given \(\sigma_{MH}\). (1) Suppose \(\frac{\bar{T}H}{T^M} \leq \frac{2(1 - \pi^2)}{\pi - 1/2}\). By Lemma 2, MH wins for sure and LM’s strategy does not affect policy. Given \(c_{LM} > 0\), \(\sigma^*_LM(\sigma_{MH}) = 0\) for
any $\sigma_{MH}$. (2) Suppose $\frac{\bar{T}H}{T} > \frac{2(1-\pi)^2}{\pi-1/2}$. By Lemma 2, LM only wins if it alone moderates ($x_{LM} = 1, x_{MH} = 0$). For a given $\sigma_{MH}$, then, LM’s expected utility of pursuing costly moderation for sure is as follows

$$EU^L(\sigma_{LM} = 1|\sigma_{MH}) = \sigma_{MH}Pr(#H < (N + 1)/2 \text{ and } a = M|x_{MH} = 1, MH) \frac{T^H}{2}$$

$$+ (1 - \sigma_{MH}) \left( Pr(#L < (N + 1)/2 \text{ and } a = L|x_{LM} = 1, LM) \bar{T}H \right)$$

$$+ Pr(#L \geq (N + 1)/2 \text{ and } a = L|x_{LM} = 1, LM) \bar{T}$$

$$+ Pr(a = M|x_{LM} = 1, LM) \left( \frac{\bar{T}H}{2} \right) - c_{LM}$$

Note that terms with zero redistribution were dropped. Given the probabilities computed above, this simplifies to $EU^L(\sigma_{LM} = 1|\sigma_{MH}) = (1 - \pi/2)T^H + (1 - \pi)^2T^M - c_{LM} - \sigma_{MH}((1 - \pi)\bar{T}H + (1 - \pi)^2\bar{T}M)$. LM loses the election for sure if it does not moderate ($x_{LM} = 1$).

Again taken $\sigma_{MH}$ as given, LM’s expected utility of not moderating is

$$EU^L(\sigma_{LM} = 0|\sigma_{MH}) = \sigma_{MH} \left( Pr(#H < (N + 1)/2 \text{ and } a = M|x_{MH} = 1, MH) \frac{\bar{T}H}{2} \right)$$

$$(1 - \sigma_{MH}) \left( Pr(#H < (N + 1)/2 \text{ and } a = M|x_{MH} = 0, MH) \frac{\bar{T}H}{2} \right)$$

$$= \sigma_{MH} \pi \frac{\bar{T}H}{2} + (1 - \sigma_{MH}) \frac{\bar{T}H}{2}$$

Comparing expected utilities between the two pure strategies, note that when $c_{LM} > (\frac{\pi}{4} - \frac{\pi}{2})\bar{T}H + (1 - \pi)^2\bar{T}M$, then LM is strictly better off not moderating even if $\sigma_{MH} = 0$. Otherwise, LM is just indifferent between costly moderation and no moderation if $EU^L(\sigma_{LM} = 1|\sigma_{MH}) = EU^L(\sigma_{LM} = 0|\sigma_{MH})$. Solving for $\sigma_{MH}$ yields the cutoff below (above) which LM chooses $\sigma_{LM} = 1$ ($\sigma_{LM} = 0$). It is denoted by $\bar{\sigma}_{MH} \equiv 1 - \frac{c_{LM}}{(1 - \bar{T})T^H + (1 - \pi)^2T^M}$. In the case of indifference, any $\sigma_{LM} \in [0,1]$ is a best-response. If $\sigma_{MH} < \bar{\sigma}_{MH}$, LM is better off moderating and if $\sigma_{MH} > \bar{\sigma}_{MH}$ LM is better off not moderating. Lemma 3 summarizes LM’s
best-response.

Lemma 3. If \( \frac{T^H}{T^M} \leq \frac{2(1-\pi)^2}{\pi-1/2} \) or \( c_{LM} > (\frac{3}{4}-\frac{\pi}{2})T^H + (1-\pi)^2T^M \), then \( \sigma^*_L M(\sigma_{MH}) = 0 \) for any \( \sigma_{MH} \). Else, the best-responding candidate selection strategy of \( L M \) is

\[
\sigma^*_L M(\sigma_{MH}) = \begin{cases} 
1 & \text{if } \sigma_{MH} < \bar{\sigma}_{MH} \\
[0, 1] & \text{if } \sigma_{MH} = \bar{\sigma}_{MH} \\
0 & \text{if } \sigma_{MH} > \bar{\sigma}_{MH}
\end{cases}
\]

where \( \bar{\sigma}_{MH} = 1 - \frac{c_{LM}}{(\frac{3}{4}-\frac{\pi}{2})T^H + (1-\pi)^2T^M} \).

Next, consider \( MH \)'s choice of candidates given \( MH \)'s decision. (1) Suppose \( \frac{T^H}{T^M} \leq \frac{2(1-\pi)^2}{\pi-1/2} \). In this case, as shown by Lemma 2, \( MH \) wins for sure even when not pursuing costly moderation. It thus minimizes losses from income redistribution by not selecting further moderates. Hence, \( \sigma^*_M H(\sigma_{LM}) = 0 \) for any \( \sigma_{MH} \). (2) Suppose \( \frac{T^H}{T^M} > \frac{2(1-\pi)^2}{\pi-1/2} \). By Lemma 2, \( MH \) only loses if it does not moderate while \( LM \) does \( (x_{LM} = 1, x_{MH} = 0) \). If \( MH \) pursues costly moderation, it wins for sure and its expected utility is

\[
EU^H(\sigma_{MH} = 1|\sigma_{LM}) = y^H - (Pr(\#H < (N + 1)/2 \text{ and } a = M|x_{MH} = 1, MH)\bar{T}^H - c_{MH})
\]

If \( MH \) refrains from further moderating, it will surely suffer from being taxed at capacity if \( LM \) moderates. Otherwise it wins the election for sure with a 1/2 chance of seeing its most preferred policy implemented. Expected utility is

\[
EU^H(\sigma_{MH} = 0|\sigma_{LM}) = y^H - (1 - \sigma_{LM})Pr(\#H < (N + 1)/2 \text{ and } a = M|x_{MH} = 0, MH)\bar{T}^H - \sigma_{LM}\bar{T}^H
\]

\[= y^H - \frac{1}{2}(1 + \sigma_{LM})\bar{T}^H \]
As \( \text{LM} \), \( \text{MH} \) will not moderate when the cost of nominating moderate types are too high. Comparing utilities makes apparent that if \( c_{MH} > (1 - \pi) \bar{T}^H \), then \( \text{MH} \) is strictly better off not moderating even if \( \sigma_{LM} = 1 \). Else, for sufficiently low \( c_{MH} \), \( \text{MH} \) is indifferent between costly moderation and no moderation if \( \sigma_{LM} = 2 \left( \pi - 1/2 + \frac{c_{MH}}{\bar{T}^M} \right) \equiv \bar{\sigma}_{LM} \). If \( \sigma_{LM} > \bar{\sigma}_{LM} \), \( \text{MH} \) is better off moderating and if \( \sigma_{LM} < \bar{\sigma}_{LM} \) \( \text{MH} \) is better off not moderating. Lemma 4 summarizes \( \text{MH} \)’s best-response.

**Lemma 4.** If \( \frac{T^H}{T^M} \leq \frac{2(1-\pi)^2}{\pi-1/2} \) or \( c_{MH} > (1 - \pi) \bar{T}^H \), then \( \sigma_{MH}^*(\sigma_{LM}) = 0 \) for any \( \sigma_{LM} \). Else, the best-responding candidate selection strategy of \( \text{MH} \) is

\[
\sigma_{MH}^*(\sigma_{LM}) = \begin{cases} 
1 & \text{if } \sigma_{LM} > \bar{\sigma}_{LM} \\
[0, 1] & \text{if } \sigma_{LM} = \bar{\sigma}_{LM} \\
0 & \text{if } \sigma_{LM} < \bar{\sigma}_{LM}
\end{cases}
\]

where \( \sigma_{LM} = 2 \left( \pi - 1/2 + \frac{c_{MH}}{\bar{T}^H} \right) \).

The equilibrium candidates follow from Lemma 3 and 4. (1) Suppose \( \frac{T^H}{T^M} \leq \frac{2(1-\pi)^2}{\pi-1/2} \). Then \( \sigma_{LM}^* = \sigma_{MH}^* = 0 \). (2) Suppose \( \frac{T^H}{T^M} > \frac{2(1-\pi)^2}{\pi-1/2} \). (a) If \( c_{LM} \geq \left( \frac{3}{4} - \frac{\pi}{2} \right) \bar{T}^H + (1 - \pi)^2 \bar{T}^M \), then \( \sigma_{LM}^* = \sigma_{MH}^* = 0 \). (b) If \( c_{LM} \leq \left( \frac{3}{4} - \frac{\pi}{2} \right) \bar{T}^H + (1 - \pi)^2 \bar{T}^M \) and \( c_{MH} \geq (1 - \pi) \bar{T}^H \), then \( \sigma_{LM}^* = 1, \sigma_{MH}^* = 0 \). (c) Finally, if \( c_{LM} \leq \left( \frac{3}{4} - \frac{\pi}{2} \right) \bar{T}^H + (1 - \pi)^2 \bar{T}^M \) and \( c_{MH} \leq (1 - \pi) \bar{T}^H \) then the equilibrium is in mixed strategies \( \sigma_{LM}^* = \bar{\sigma}_{LM}, \sigma_{MH}^* = \bar{\sigma}_{MH} \). There are no other equilibria. Suppose otherwise and let \( \sigma_{MH} < \bar{\sigma}_{MH} \). Hence, as summarized by Lemma 3, \( \sigma_{LM}^*(\sigma_{MH}) = 1 \). Given the decision to moderate, Lemma 4 shows that \( \text{MH} \) wants to deviate to \( \sigma_{MH}^*(1) = 1 \). But now \( \text{LM} \) is better off not moderating for sure, \( \sigma_{LM}^*(1) = 0 \). This obviously cannot be an equilibrium either as \( \text{MH} \) is better off with a deviation to no moderation, \( \sigma_{MH}^*(0) = 0 \). This is where we started. Given \( \sigma_{MH} = 0 \), \( \text{LM} \)’s best response is to invest in moderate candidates.
Given the best-responding pair of candidate selection strategies \((\sigma_{LM}^*, \sigma_{HM}^*)\), optimal voting behavior (Lemma 2) and policymaking (Lemma 1) determine the expected electoral and policy outcome. Hence, in cases (1) and (2a): \(Pr(\text{LM wins}) = 0\) and the expected policy vector is \(E(p) = \frac{1}{2}\hat{p}^M + \frac{1}{2}\hat{p}^H\), which entails an expected tax effort of \(R = \frac{1}{2}\bar{T}^H\). Case (2b) implies \(Pr(\text{LM wins}) = 1\) and \(E(p) = \pi\hat{p}^M + (1 - \pi)^2\hat{p}^L + \pi(1 - \pi)(\bar{T}^H, 0, -\bar{T}^H)\). The latter entails \(R = \bar{T}^H + (1 - \pi)^2\bar{T}^M\). Case (2c) implies \(Pr(\text{LM wins}) = \frac{\sigma_{LM}^*}{\sigma_{MH}^*}(1 - \sigma_{MH}^*)\) and expected policy \(E(p) = \sigma_{LM}^*(1 - \sigma_{MH}^*)\left(\pi\hat{p}^M + (1 - \pi)^2\hat{p}^L + \pi(1 - \pi)(\bar{T}^H, 0, -\bar{T}^H)\right) + \sigma_{MH}^*(\pi\hat{p}^M + (1 - \pi)\hat{p}^H) + (1 - \sigma_{LM}^*)(1 - \sigma_{MH}^*)\left(\frac{1}{2}\hat{p}^M + \frac{1}{2}\hat{p}^H\right)\). Hence, \(R = \left(\frac{1}{2}(1 + \sigma_{LM})(1 - \sigma_{MH}) + \sigma_{MH}\pi\right)\bar{T}^H + \sigma_{LM}(1 - \sigma_{MH})(1 - \pi)^2\bar{T}^M\). □

**Comparison with PR**

**Proposition 2.** The expected size of the redistributive budget is smaller in the majoritarian electoral system compared to the PR system analyzed by Iversen and Soskice (2006).

**PROOF**

The result directly follows from comparing the equilibrium policy described in Proposition 1 with the equilibrium policy of the PR model derived by Iversen and Soskice (2006, Proposition 2). The equilibrium policy under PR (with \(M\) as formateur) is \(p_{PR}^* = (\bar{T}, \bar{T} - \bar{T}^M, -\bar{T}^H)\) (recall \(\bar{T} = \bar{T}^M + \bar{T}^H\)). It is straightforward to verify that the size of the redistributive budget under PR, \(R_{PR}^* = \bar{T}^H + \frac{\bar{T}^M}{2}\) is strictly larger than that in the majoritarian model (recall that \(1 > \pi > 1/2\)): \(\bar{T}^H + \frac{\bar{T}^M}{2} > \max \{R_{MAJ}^*\} = \bar{T}^H + (1 - \pi)^2\bar{T}^M\). □
S2. Theoretical Extension: A Dynamic Game

This appendix formally develops and solves the dynamic extension of the basic model informally discussed in the main text. The point of the extension is to highlight how inter-temporal considerations - the interplay between repeated elections and economic effects of redistribution - can influence the commitment problem that lies at the core of the theoretical argument. In particular, current redistribution influences the future growth and distribution of market income and thereby also shapes future conflict over redistribution. To explore how such dynamic considerations shape endogenous commitment, I analyze a two-period version of the one-period model analyzed so far, in the spirit of models like that of Besley and Coate (1998). The results from the static model still go through in this repeated elections framework. But accounting for dynamics generates additional insights.

Set-up

Suppose there are two electoral periods, $t \in \{1, 2\}$. Each period repeats the basic model, consisting of the sequence of candidate selection, general election and policymaking spelled out before. Players are forward-looking and care about the sum of utility over the two periods, where second-period utility is discounted by $\delta \in (0, 1]$. After the first period ends, there is a shock to the income distribution. Then there is a second and final round of candidate selection, general election and policymaking. The income shock depends on previous redistributive policy. To focus on the political dynamics, economic effects are modeled in a reduced-form way. Denote the exogenous market income of a voter belonging to group $J \in \{L, M, H\}$ in period $t$ as $y_t^J$. Crucially, second-period incomes before depend on the size of the redistributive budget in the first period ($R_1$). $R_1$ measures the taxes levied one or more groups taxed to finance redistribution to other groups. The underlying economic assumption is that higher income redistribution through taxes and transfers in $t = 1$ leads
to lower income growth, at least for the middle-income and high-income group, and lower inequality between the middle and the top incomes before taxes and transfers in $t = 2$. To clarify the intuition, it is sufficient to focus on two cases. First, the total size of the tax burden in the first period is relatively low ($R_1 \leq K$), where $K > 0$ captures how the economy responds to redistribution, and so future income growth and income inequality before taxes and transfers are high. Denote second-period incomes in this case by $\tilde{y}_2^J$ and taxable capacity by $\tilde{T}_2^J$. In particular, the second-period income distribution satisfies the condition $\frac{\tilde{T}_2^H}{\tilde{T}_2^M} > \frac{2(1-\pi)^2}{\pi - 1/2}$. Second, the total size of the tax burden in the first period is relatively high ($R_1 > K$). In this case, future income growth and inequality before taxes and transfers are relatively low. Denote second-period incomes in this case by $\tilde{y}_2^L$, taxable capacity by $\tilde{T}_2$ and let $\frac{\tilde{T}_2^H}{\tilde{T}_2^M} \leq \frac{2(1-\pi)^2}{\pi - 1/2}$. By definition, $\tilde{y}_2^M > \tilde{y}_2^L$, $\tilde{y}_2^H > \tilde{y}_2^L$, $\tilde{T}_2^M > \tilde{T}_2^L$ and $\tilde{T}_2^H > \tilde{T}_2^L$. For simplicity and in line with Iversen and Soskice (2006), let $\tilde{y}_2^L = \tilde{y}_2^L = 0$. To exclude cases with trivially low economic cost of redistribution, I assume $K < \tilde{T}_1$.

The economic effects of current redistribution on future income are common knowledge to the players. As in the static model, the equilibrium concept is subgame perfect Nash equilibrium and weakly dominated strategies in legislative voting are excluded. The game is solved by backward induction. The analysis focuses on the most interesting scenario of sufficiently low commitment cost.

Results

First-period policy of $L$-majority

A first result concerns the policy chosen by a majority of policymakers representing low-income citizens in the first period. In the dynamic framework, the commitment problem does not go away even if policymakers in the first period put a large weight on second-period utility and anticipate how current redistributive policy affects future income distribution and
voting behavior. To see this, consider the scenario where the center-left party LM wins the
election in the first period and there is a majority of L-types in the legislature that determines
policy for $t = 1$. While policymakers representing low-income voters are forward-looking,
they will generally choose more redistribution to the low-income group than preferred by the
pivotal middle-income voter $M$ (Proposition 3). Intuitively, first-period policy in the dynamic
framework can be more moderate than in the static model if policymakers put a high weight
on the second period. There are two related reasons for this. First, L-policymakers face a
trade-off between current and future redistribution. This is a standard economic constraint
concerning how to maximize revenue for redistribution over time. Second, L-policymakers
understand that high redistribution in the first period can undermine political support among
the middle class in the second period. This is a political constraint that goes to the heart
of the commitment problem. In response, the L-majority will choose to moderate current
redistribution if it cares sufficiently about the future - though not enough to eliminate the
commitment problem. Policymakers in the first period care about what happens in the
second period, but they do at least equally care about first period policy and preserving
their future electoral opportunities does not generally require implementing the ideal point
of the middle-income, median voter $M$. In a framework of finitely repeated elections, $M$
cannot credibly commit to punish all deviations from its most preferred policy.

**Proposition 3.** Suppose center-left party LM wins the election in the first period and there
is a majority of L-types in the legislature. Even if policymakers put a large weight on the
second-period, they will choose strictly more redistribution to the low-income group than
preferred by the middle-income voter $M$.

**PROOF**

Policymakers in the first period anticipate the best-responding behavior of voters and
parties in the second period. As the game ends after the second period, the equilibrium in
$t = 2$ corresponds to that of the static game characterized by Proposition 1. Concerning the
optimal policy choice legislative majority of \( L \)-types in \( t = 1 \), there are two relevant cases. First, suppose the \( L \)-majority chooses a policy of high redistribution in \( t = 1 \), setting total taxes on \( M \) and \( H \) above \( K \) to finance transfers exclusively to \( L \). Following the economic assumption, second-period incomes of the two upper groups are compressed such that \( \tilde{T}^H \frac{H}{2} \leq \frac{2(1-\pi)^2}{\pi - 1/2} \). By the proof of Proposition 1, this means that in the second period voter \( M \) has a dominant strategy to support party \( MH \). As a result, neither party opts for full commitment, \( MH \) wins the election for sure and the expected policy vector in \( t = 2 \) is \( \frac{1}{2} \hat{p}^M + \frac{1}{2} \hat{p}^H \). Hence, expected second-period utility for \( J \)-type policymakers is \( \hat{V}_J^J(\hat{p}^M) + \hat{V}_J^J(\hat{p}^H) \equiv \hat{V}_2^J \) For \( L \)-types, this simplifies to \( \tilde{T}^H \frac{M}{4} \). Second, the legislative majority of \( L \)-types chooses a policy of moderate or low redistribution in \( t = 1 \), setting total tax on \( M \) and \( H \) at or below \( K \) to finance transfers to \( L \). This results in higher incomes growth and inequality between the top and the middle, \( \tilde{T}^M \frac{L}{2} > \frac{2(1-\pi)^2}{\pi - 1/2} \). This means that the condition for a mixed strategy equilibrium in the second period holds: \( LM \) and \( MH \) choose full moderation with probability \( \sigma^*_{LM} \) and \( \sigma^*_{MH} \), respectively. In this case, we know from the proof of Proposition 1 that expected policy in the second period is

\[
E(p_2) = \sigma^*_{LM}(1 - \sigma^*_{MH}) \left( \pi \hat{p}^M + (1 - \pi)^2 \hat{p}^L + \pi(1 - \pi)(\tilde{T}^H, 0, -\tilde{T}^H) \right) \\
+ \sigma^*_{MH} \left( \pi \hat{p}^M + (1 - \pi)\hat{p}^H \right) + (1 - \sigma^*_{LM})(1 - \sigma^*_{MH}) \left( \frac{1}{2} \hat{p}^M + \frac{1}{2} \hat{p}^H \right)
\]

Thus, \( J \)'s expected utility in the second period is

\[
\sigma^*_{LM}(1 - \sigma^*_{MH}) \left( \pi V^J(\hat{p}^M) + (1 - \pi)^2 V^J(\hat{p}^L) + \pi(1 - \pi)V^J((\tilde{T}^H, 0, -\tilde{T}^H)) \right) \\
+ \sigma^*_{MH} \left( \pi V^J(\hat{p}^M) + (1 - \pi)V^J(\hat{p}^H) \right) + (1 - \sigma^*_{LM})(1 - \sigma^*_{MH}) \left( \frac{1}{2} V^J(\hat{p}^M) + \frac{1}{2} V^J(\hat{p}^H) \right) \equiv \tilde{V}_2^J
\]

Note that \( \tilde{V}_2^L > \tilde{V}_2^J \). Suppose \( L \) puts a sufficiently high weight on the second period so that it prefers not to tax \( M \) and \( H \) at capacity to maximize first-period redistribution to \( L \).
(δ > \frac{\bar{T}_2 - K}{V_2^L - V_2^L}). Then the L-majority maximizes its utility by setting the total tax burden on M and H to K. If K > \bar{T}^H, this implies a net transfer of K to L. The distribution of taxes among M and H is restricted by the non-regressivity assumption, but this still leaves a set of feasible policies that make L indifferent. At maximum, H can be taxed at capacity and M be taxed to finance the remainder: \( p_1 = (K, \bar{T}^H - K, -\bar{T}^H) \). L-policymakers can also choose a more equal distribution of taxes between M and H. Given non-regressivity, however, the net transfer to H is at least \(-K/2\). In any case, M is a net loser from redistribution and strictly prefers all policies that reduce its redistributive burden. If \( T^H_1 \geq K \), then all the tax burden may be concentrated on H, \( p_1 = (K, 0, -K) \), or be shared more equally among M and H subject to non-regressivity. While the former is the most favorable outcome for M among the policies that yield the same benefit to the L-majority, M would clearly nonetheless prefer to receive a positive net transfer. □

**Equilibrium**

A second result is that dynamic consideration can actually increase the commitment problem faced by the center-left party (LM) - because forward-looking middle-income voters (M) now have an additional motivation to fear a left government: not just the current loss imposed by redistribution to the poor, but also the reduction in future income as a consequence of high redistribution. As a consequence, they can be more reluctant to support the LM party. In the theoretical framework, the severity of the commitment problem is indicated by the relative income threshold below which middle-class voters have a dominant strategy to vote for the center-right party (MH). Proposition 4 describes an equilibrium in which the endogenous threshold is more demanding in the dynamic than in the static model. Accordingly, median
voter $M$ always votes for party $MH$ as long as

$$\frac{\bar{T}_H}{\bar{T}_M} \leq \frac{2(1-\pi)^2}{\pi-1/2} \left(1 + \frac{\delta (\bar{V}^M_M - \bar{V}^H_M)}{\bar{T}_M^M} \right)$$

where $\bar{V}^M_2$ (formally defined above) refers to $M$'s expected utility in second period given that a moderate or low redistributive policy is chosen in the first period, so that $\frac{\bar{T}_H}{\bar{T}_M} > \frac{2(1-\pi)^2}{\pi-1/2}$, and $\bar{V}^M_2$ refers to $M$'s expected utility in second period in the opposite case of high first-period redistribution, leading to $\frac{\bar{T}_H}{\bar{T}_M} \leq \frac{2(1-\pi)^2}{\pi-1/2}$. Note that the threshold obtained in the dynamic model includes the threshold of the static model as a special case where $\delta = 0$. The right-hand-side of the threshold is strictly larger than in the static model if $\bar{V}^M_2 - \bar{V}^H_2 > 0$. In substantive terms, this condition is plausible and requires that $M$ prefers the equilibrium with higher market income and moderation by both parties to the equilibrium with lower market income and no moderation. Inspecting the threshold makes apparent that commitment becomes more of a problem for the left party as voters attach more weight to the future (i.e., $\delta$ increases) - up to the point of $\delta = \frac{\bar{T}_1 - K}{V^L_2 - V^L_2}$.

Also, note that the equilibrium recovers the key empirical implications obtained in the static model. For instance, higher inequality between $M$ and $H$ increases the probability of a credible right-ward shift by party $LM$.

**Proposition 4.** Suppose $K > \bar{T}_1^H$, $\delta < \frac{\bar{T}_1 - K}{V^L_2 - V^L_2}$, $\bar{V}^M_2 > \bar{V}^H_2$ and $\bar{V}^H_2 > \bar{V}^H_2$. In equilibrium, parties’ candidate selection strategies, the probability that $LM$ wins the election and the expected size of the redistributive budget in the first period ($t = 1$) are as follows:

1. Suppose $\frac{\bar{T}_H}{\bar{T}_M} \leq \frac{2(1-\pi)^2}{\pi-1/2} \left(1 + \frac{\delta (\bar{V}^M_2 - \bar{V}^H_2)}{\bar{T}_M^M} \right)$
   
   - $\sigma^*_L = \sigma^*_M = 0$
   
   - $Pr(\text{LM wins}) = 0$
   
   - $R_1 = \frac{1}{2} \bar{T}_1^H$
2. Suppose \( \frac{T^H}{T^M} > \frac{2(1-\pi)^2}{\pi-1/2} \left( 1 + \frac{\delta(V^M_2 - V^M_1)}{T^M_1} \right) \)

- \( \sigma^*_L = \frac{(\pi-1/2)T^H_L + c_{LM}}{T^H_L + (1-\pi)^2 \delta(V^M_2 - V^M_1)}, \sigma^*_M = 1 - \frac{c_{LM}}{(\delta-\pi)T^H_L + (1-\pi)^2(\delta(V^M_2 - V^M_1))} \)

- \( \Pr(\text{LM wins}) = \sigma^*_L (1 - \sigma^*_M) \)

- \( R_1 = \left( \frac{1}{2}(1 + \sigma^*_L)(1 - \sigma^*_M) + \sigma^*_L \pi \right) T^H_1 + \sigma^*_L (1 - \sigma^*_M)(1 - \pi)^2 T^M_1 \)

**PROOF**

Following the logic of backward induction and as in the proof of Proposition 3, start by considering policymaking in \( t = 1 \). Policymakers at this stage rationally anticipate the best-responding behavior of voters and parties in the second period. As the game ends after the second period, the equilibrium in \( t = 2 \) corresponds to that of the static game characterized by Proposition 1.

**Policymaking in** \( t = 1 \). Legislators of type \( J \) accept any proposal \( b = (p^L_0, p^M_0, p^L_0) \) with total tax burden \( R_1 \leq K \) if \( p^J_0 + y^J + \delta \tilde{V}^J_2 \geq y^J + \delta \tilde{V}^J_2 \) and reject it otherwise. The acceptance condition simplifies to \( p^J_0 \geq 0 \). A proposal with a tax burden of \( R_1 > K \) is accepted if \( p^J_0 + y^J + \delta \tilde{V}^J_2 \geq y^J + \delta \tilde{V}^J_2 \) and rejected otherwise. The acceptance condition simplifies to \( p^J_0 \geq \delta(\tilde{V}^J_2 - \tilde{V}^J_2) \). Given \( \tilde{V}^J_2 > \tilde{V}^J_2 \), any type of policymaker needs to be compensated to support a proposal with \( R_1 > K \) in \( t = 1 \). Given non-regressivity, this is never feasible for \( H \). Hence, policymakers representing the high-income group only accept proposals with \( R_1 \leq K \) and weakly positive net transfers.

The optimal proposal by the agenda setter of type \( J \) maximizes the utility of group \( J \) over both periods subject to legislators’ acceptance constraints and the policy constraints. First, consider an agenda setter of type \( M \). If there is no majority of \( H \) types, \( |H| < \frac{N+1}{2} \), \( M \)'s optimal proposal is \( b^M* = (\frac{\bar{T}^H}{2}, \frac{T^H}{2}, -\bar{T}^H_1) \), which entails \( R_1 < K \). The proposal will be accepted by both \( M \) and \( L \) legislators. This is makes \( M \) better off than any proposal with \( R_1 > K \) that by non-regressivity yields an expected utility for \( M \) of at most \( y^M_1 + \frac{T^H}{2} + \delta \tilde{V}^J_2 \).
Given $\tilde{V}_2^M > \tilde{V}_2^M$, this is strictly smaller than the utility $M$ received from proposal $b^M = (\bar{T}_H^M, \bar{T}_M^M, -\bar{T}_H^M)$: $y_1^M + \bar{T}_H^M + \delta V^J$. If they constitute a majority, $H$ legislators will veto any redistributive bill so any feasible proposal is a best response.

Second, assume there is a $L$ agenda setter. If there is a majority of $L$ legislators, $|L| \geq \frac{N+1}{2}$, the optimal proposal is $b^* = (\bar{T}_H^L, -\bar{T}_M^L, -\bar{T}_H^L)$, which entails $R_1 > K$. This is preferred to proposals that redistributes $K$ or less to $L$ because $\delta < \frac{\bar{T}_H^M - K}{V^L_2 - V^M_2}$. If $|L| < \frac{N+1}{2}$ and $|L| + |M| \geq \frac{N+1}{2}$, then the best proposal is $b^* = (\bar{T}_H^H, 0, -\bar{T}_H^H)$ with $R_1 < K$. This just satisfies the acceptance condition of $M$. It makes $L$-policymakers better off than any alternative proposal with $R_1 > K$. Recall from above that such a proposal has to compensate $M$ by at least $\delta(\tilde{V}_2^M - \tilde{V}_2^M)$ to be accepted. This leaves $L$ with a net transfer of no more than $\bar{T}_H^H - \delta(\tilde{V}_2^M - \tilde{V}_2^M)$ in the first period, lower than under proposal above, as well as lower second-period utility, because $\tilde{V}_2^L > \tilde{V}_2^L$. Else, if $|H| \geq \frac{N+1}{2}$, $H$ legislators will veto any redistributive bill so any proposal is a best response.

Finally, suppose $H$ is agenda setter. It follows that $H$ proposes the ideal policy of the high-income group, $b^H = (0, 0, 0)$, and it will be accepted. Taken together, the legislative outcome for a given type of agenda-setter and composition of the legislature are summarized by Lemma 5.

**Lemma 5.** For a given post-electoral distribution of preference types in the legislature, the outcomes of the legislative process is as follows:

<table>
<thead>
<tr>
<th>Seat allocation in legislature</th>
<th>Agenda setter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>L</td>
</tr>
<tr>
<td>$(\bar{T}_1, -\bar{T}_1^M, -\bar{T}_H^M)$</td>
<td>$(\bar{T}_H^H, \bar{T}_H^H, \bar{T}_H^H)$</td>
</tr>
<tr>
<td>$</td>
<td>L</td>
</tr>
<tr>
<td>$</td>
<td>H</td>
</tr>
</tbody>
</table>
General election in $t = 1$. The pivotal voter $M$ has rational expectations about the policy outcome of the post-electoral game in $t = 1$ and its consequences for pre-fiscal income distribution and police in $t = 2$ and votes accordingly. Lemma 6 summarizes the best-responding voting behavior of the pivotal voter $M$ for a given electoral set of candidates nominated by each party. Recall that by the sequence of actions, the pivotal voter $M$ observes the (realized) choice of candidates $(x_{LM}, x_{MH})$ even if parties use non-degenerate mixed strategies ($\sigma_{LM}$ and $\sigma_{MH}$, respectively).

Lemma 6. The median voter $M$ in the pivotal districts votes for the candidate of party $MH$ unless $x_{LM} = 1, x_{MH} = 0$ and $\frac{T^H_1}{T^M_1} > \frac{2(1-\pi)^2}{\pi - 1/2} \left(1 + \frac{\delta (V^M_2 - \bar{V}^M)}{\bar{T}^M_1 - 2}\right)$, in which case $M$ supports LM’s candidate.

The structure of the argument is the same as in the static model. The difference is that the dynamic extension requires that $M$ also takes expected outcomes in $t = 2$ into account. For a given $x_{LM} \in \{0, 1\}$ (where 1 means full, costly moderation and 0 no moderation), $M$’s expected utility from voting for LM is

\[
EU^M_{\text{vote LM}}(x_{LM}) = P_r(a = M|x_{LM}, LM) \left(y^M_1 + \frac{T^H_1}{2} + \delta \bar{V}^M_2\right) + P_r \left(\frac{N + 1}{2} \geq \#L \text{ and } a = L \left|x_{LM} = 1, LM\right\} \left(y^M_1 - T^M_1 + \delta \bar{V}^M_2\right) + P_r \left(\frac{N + 1}{2} < \#L \text{ and } a = L \left|x_{LM} = 1, LM\right\} \left(y^M_1 + \delta \bar{V}^M_2\right)
\]
For a given $x_{MH}$, $M$’s expected utility from voting for MH is

$$EU^M(\text{vote } MH|x_{MH}) = Pr\left(# H < \frac{N + 1}{2} \text{ and } a = M \bigg| x_{MH}, MH\right) \left( y_1^M + \frac{T^H}{2} + \delta V_2^M \right)$$

$$+ Pr\left(# H \geq \frac{N + 1}{2} \text{ or } a = H \bigg| x_{MH}, MH\right) \left( y_1^M + \delta V_2^M \right)$$

After substitution of the probabilities (same as Proposition 1),

$$EU^M(\text{vote } LM|x_{LM} = 1) = \pi(y_1^M + \frac{T^H}{2} + \delta V_2^M) + (1 - \pi)^2(y_1^M - T_1^M + \delta V_2^M)$$

$$+ \pi(1 - \pi)(y_1^M + \delta V_2^M)$$

$$EU^M(\text{vote } LM|x_{LM} = 0) = \frac{1}{2}(y_1^M + \frac{T^H}{2} + \delta V_2^M) + \pi(1 - \pi)(y_1^M - T_1^M + \delta V_2^M)$$

$$+ \frac{1}{2}(1 - \pi)(y_1^M + \delta V_2^M)$$

$$EU^M(\text{vote } MH|x_{MH} = 1) = \pi(y_1^M + \frac{T^H}{2} + \delta V_2^M) + (1 - \pi)(y_1^M + \delta V_2^M)$$

$$EU^M(\text{vote } MH|x_{MH} = 0) = \frac{1}{2}(y_1^M + \frac{T^H}{2} + \delta V_2^M) + \frac{1}{2}(y_1^M + \delta V_2^M)$$

$M$ chooses the party whose offer of candidates yields higher expected utility. It is straightforward to verify that if $x_{LM} \leq x_{MH}$, then $M$ is always strictly better off voting for MH. If $x_{LM} > x_{MH}$, then $M$ continues to strictly prefer MH if $\frac{T^H}{T_1^M} < \frac{(1 - \pi)^2}{\pi - 1/2} \left( 1 + \frac{\delta (V_2^M - V_2^L)}{T_1^M} \right)$; $M$ strictly prefers LM if $\frac{T^H}{T_1^M} > \frac{(1 - \pi)^2}{\pi - 1/2} \left( 1 + \frac{\delta (V_2^M - V_2^L)}{T_1^M} \right)$ and is indifferent otherwise, in which case I assume that $M$ votes for MH. This completes the proof of Lemma 6.

**Candidate selection and equilibrium outcomes in $t = 1$.** The majority faction in each party chooses a candidate selection strategy, summarized by the probability of pursuing full moderation, that is a best response given the other party’s strategy, the best response of the pivotal voter (Lemma 6) and legislators (Lemma 5). First, consider LM’s decision given
\(\sigma_{MH}\). (1) Suppose \(\frac{T_H^H}{T_M^H} \leq \frac{2(1-\pi)^2}{\pi-1/2} \left( 1 + \frac{\delta(V_M^M-V_L^M)}{T_M^M} \right)\). By Lemma 6, \(MH\) wins for sure and \(LM\)'s strategy does not affect policy. Given \(c_{LM} > 0\), \(\sigma^*_L M(\sigma_{MH}) = 0\) for any \(\sigma_{MH}\). (2) Suppose \(\frac{T_H^H}{T_M^H} > \frac{2(1-\pi)^2}{\pi-1/2} \left( 1 + \frac{\delta(V_M^M-V_L^M)}{T_M^M} \right)\). By Lemma 6, \(LM\) only wins if it alone moderates \((x_{LM} = 1, x_{MH} = 0)\). Hence, \(LM\)'s expected utility of pursing costly moderation for sure given \(\sigma_{MH}\) is

\[
EU^L(\sigma_{LM} = 1|\sigma_{MH}) = \sigma_{MH} \left( \frac{T_H^H}{2} + \delta V^L_2 \right) + (1 - \sigma_{MH}) \left( (1-\pi)(T_H^H + \delta V^L_2) \right) + (1 - \pi)^2(T_1 + \delta V^L_2) + \pi \left( \frac{T_H^H}{2} + \delta V^L_2 \right)
\]

\(LM\) loses the election for sure if it does not moderate \((x_{LM} = 1)\). In this case \(LM\)'s expected utility is

\[
EU^L(\sigma_{LM} = 0|\sigma_{MH}) = \sigma_{MH} \left( \frac{T_H^H}{2} + \delta V^L_2 \right) + (1 - \sigma_{MH}) \left( \frac{T_H^H}{2} + \delta V^L_2 \right)
\]

By assumption we have ruled out cases \(c_{LM}\) where large enough so that \(L\) has a dominant strategy to never moderate. Equating the utility from both pure strategies and solving for \(\sigma_{MH}\) yields the cutoff below (above) which \(LM\) chooses \(\sigma_{LM} = 1\) \((\sigma_{LM} = 0)\). It is denoted by \(\bar{\sigma}_{MH} \equiv 1 - \frac{c_{LM}}{(\frac{1}{4}-\frac{1}{2})T_H^H + (1-\pi)^2(T_M^M - \delta(V_L^L-V_L^M))}\). Note that \(T_M^M - \delta(V_L^L-V_L^M) > 0\) as \(\delta < \frac{T_1-K}{V_L^L-V_L^M}\) and \(K > T_H^H\). In the case of indifference, any \(\sigma_{LM} \in [0,1]\) is a best-response. If \(\sigma_{MH} < \bar{\sigma}_{MH}\), \(LM\) is better off moderating and if \(\sigma_{MH} > \bar{\sigma}_{MH}\) \(LM\) is better off not moderating. Lemma 7 summarizes \(LM\)'s best-response.

**Lemma 7.** If \(\frac{T_H^H}{T_M^H} \leq \frac{2(1-\pi)^2}{\pi-1/2} \left( 1 + \frac{\delta(V_M^M-V_L^M)}{T_M^M} \right)\), then \(\sigma^*_L M(\sigma_{MH}) = 0\) for any \(\sigma_{MH}\). Else, the
best-responding candidate selection strategy of $\mathbf{LM}$ is

$$
\sigma^*_\text{LM}(\sigma_{\text{MH}}) = \begin{cases} 
1 & \text{if } \sigma_{\text{MH}} < \bar{\sigma}_{\text{MH}} \\
[0, 1] & \text{if } \sigma_{\text{MH}} = \bar{\sigma}_{\text{MH}} \\
0 & \text{if } \sigma_{\text{MH}} > \bar{\sigma}_{\text{MH}}
\end{cases}
$$

where $\bar{\sigma}_{\text{MH}} = 1 - \frac{c_{\text{LM}}}{(\frac{1}{4} - \frac{\pi}{2}) T_1^H + (1 - \pi)^2 (T_1^M - \delta (V_2^L - V_2^R))}$.

Following same logic, the best-response for the majority faction of party $\mathbf{MH}$ is characterized by Lemma 8.

**Lemma 8.** If $\frac{T_1^H}{T_1^M} \leq \frac{2(1-\pi)^2}{\frac{\pi}{1-\frac{1}{2}} \left(1 + \frac{\delta (V_2^M - V_2^L)}{T_1^M}\right)}$, then $\sigma^*_\text{MH}(\sigma_{\text{LM}}) = 0$ for any $\sigma_{\text{LM}}$. Else, the best-responding candidate selection strategy of $\mathbf{MH}$ is

$$
\sigma^*_\text{MH}(\sigma_{\text{LM}}) = \begin{cases} 
1 & \text{if } \sigma_{\text{LM}} > \bar{\sigma}_{\text{LM}} \\
[0, 1] & \text{if } \sigma_{\text{LM}} = \bar{\sigma}_{\text{LM}} \\
0 & \text{if } \sigma_{\text{LM}} < \bar{\sigma}_{\text{LM}}
\end{cases}
$$

where $\bar{\sigma}_{\text{LM}} = \frac{(\frac{\pi}{1-\frac{1}{2}} T_1^H + c_{\text{MH}})}{T_2^H + \delta (V_2^M - V_2^L)}$.

The equilibrium candidate selection strategies and policy are apparent from Lemma 5 - 8. $\square$
S3. Appendix for Analysis of Party Positions

Data

This section provides more details on the data used in the cross-national analysis of party positions. Table 1 reports the summary statistics and sources for all variables used in the analysis, including additional variables introduced in the sensitivity analysis discussed in the next section.

Main Left Party What matters to identify the main left party is not the absolute left-right position of each party, as the theoretical expectation is that they differ across electoral rules, but the relative position in the party system. It is generally straightforward to identify the main left party in the countries under study. Following Pontusson and Rueda (2010), it refers to the labor parties of Australia, Britain, Ireland, the Netherlands, New Zealand and Norway, the social democratic parties of Denmark, Finland, Germany, Sweden and Switzerland, the socialist parties in France, France, Japan, Portugal and Spain, and liberal party in Canada. In the case of Italy, where the communist party (and its successor) often was stronger as strong as the socialist party, the left-right position is the weighted average of both, with vote share as weights; not including the communists, who were excluded from government during the cold war, produces the same results.

PR Following the coding of Iversen and Soskice (2006), the following elections are classified as PR = 1: Denmark, Finland, France (until 1956 and 1986), Germany, Italy, Netherlands, New Zealand (since 1996), Norway, Portugal, Spain, Sweden, Switzerland. All the other elections are coded as majoritarian (PR = 0).

Top Income Inequality The data described in the paper are from the World Top Incomes Database (Atkinson and Piketty 2007; Atkinson and Piketty 2010) and were retrieved from http://topincomes.parisschoolofeconomics.eu/ (last accessed on February 3, 2014). The variables were rescaled as deviations from the mean. The inverted Pareto
coefficient characterizing top incomes is the main measure of income inequality. The Pareto distribution is a power law distribution that is usually considered to be a good approximation of the top part (e.g., top 10%) of the observed income distribution. The Pareto law for top incomes is given by the distribution function \( F(y) = \frac{k}{y^\alpha} \) where \( \alpha \) is called the Pareto parameter and \( k \) is another parameter (Atkinson, Piketty and Saez, 2011, pp. 13-15). According to the documentation of the World Top Incomes Database, the inverted Pareto-Lorenz coefficient, \( \beta \) was calculated by Atkinson and colleagues from the Pareto-Lorenz coefficient, \( \beta \), using the formula \( \beta = \alpha/(\alpha - 1) \), where \( \alpha \) was computed using the top share estimates. They write that as a rule the coefficient was estimated using the following formula: \( \alpha = 1/[1 - \log(S1%/S0.1%)/\log(10)] \), where \( S1\% \) refers to the income share of the top 1 percent and \( S0.1 \) is the income share of the top 0.1 percent. When the top 0.1% and top 1% shares were not available, the closest substitutes were used.
Table 1: Descriptive Statistics and Sources

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
<th>N</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic Left-Right Position</td>
<td>0.21</td>
<td>0.62</td>
<td>-1.21</td>
<td>2.45</td>
<td>228</td>
<td>Eloff (2013)</td>
</tr>
<tr>
<td>Left-Right Position (Logit Scale)</td>
<td>-0.86</td>
<td>0.76</td>
<td>-3.78</td>
<td>1.10</td>
<td>228</td>
<td>Lowe et al. (2011)</td>
</tr>
<tr>
<td>Top Income Inequality (inverted Pareto coeff.)</td>
<td>1.77</td>
<td>0.30</td>
<td>1.28</td>
<td>2.90</td>
<td>228</td>
<td>World Top Incomes Database (see text)</td>
</tr>
<tr>
<td>Top Income Share (1%)</td>
<td>7.94</td>
<td>2.02</td>
<td>4.00</td>
<td>14.56</td>
<td>227</td>
<td>Same as above</td>
</tr>
<tr>
<td>PR</td>
<td>0.54</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
<td>228</td>
<td>Iversen and Soskice (2006)</td>
</tr>
<tr>
<td>Log of District Magnitude</td>
<td>1.60</td>
<td>1.48</td>
<td>0</td>
<td>5.01</td>
<td>228</td>
<td>Bormann and Golder (2013)</td>
</tr>
<tr>
<td>Per Capita Income (in 1000 USD)</td>
<td>18.7</td>
<td>7.8</td>
<td>3.7</td>
<td>38.2</td>
<td>228</td>
<td>Feenstra, Inklaar and Timmer (2013)</td>
</tr>
<tr>
<td>Voter Turnout</td>
<td>0.76</td>
<td>0.11</td>
<td>0.35</td>
<td>0.95</td>
<td>228</td>
<td>International IDEA (2012)</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>54.5</td>
<td>24.9</td>
<td>17.2</td>
<td>172.5</td>
<td>228</td>
<td>WDI (2013); Franzese (2002) before 1960</td>
</tr>
<tr>
<td>Population Over 65 (%)</td>
<td>12.3</td>
<td>2.9</td>
<td>5.8</td>
<td>20.0</td>
<td>228</td>
<td>WDI (2013); Franzese (2002) before 1960</td>
</tr>
<tr>
<td>Log of Population</td>
<td>16.6</td>
<td>1.2</td>
<td>14.5</td>
<td>18.7</td>
<td>228</td>
<td>Feenstra, Inklaar and Timmer (2013)</td>
</tr>
<tr>
<td>Union Density</td>
<td>0.42</td>
<td>0.19</td>
<td>0.08</td>
<td>0.87</td>
<td>223</td>
<td>Visser (2013); Golden (2009) before 1960</td>
</tr>
<tr>
<td>EU Integration</td>
<td>0.14</td>
<td>0.19</td>
<td>0</td>
<td>0.55</td>
<td>228</td>
<td>Hooghe and Marks (2001)</td>
</tr>
<tr>
<td>Incumbent Cabinet (%)</td>
<td>0.51</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
<td>228</td>
<td>Doring and Manow (2012)</td>
</tr>
<tr>
<td>Left Competitor (%)</td>
<td>0.69</td>
<td>0.46</td>
<td>0</td>
<td>1</td>
<td>228</td>
<td>Eloff (2013)</td>
</tr>
<tr>
<td>Centralization (of candidate selection)</td>
<td>0.35</td>
<td>0.29</td>
<td>0</td>
<td>1</td>
<td>220</td>
<td>Ranney (1981); Bille (2001); Lundell (2004)</td>
</tr>
<tr>
<td>Gini coefficient (pre-fisc)</td>
<td>42.3</td>
<td>5.59</td>
<td>28.54</td>
<td>53.75</td>
<td>192</td>
<td>Solt (2014)</td>
</tr>
</tbody>
</table>
Sensitivity Analysis

A first set of robustness checks reported in Table 2 uses alternative measures for top inequality and electoral institutions, respectively. The Top Income Share (TIS) is also derived from World Top Incomes Database. It captures the income share of the top 1%. This a measure used by Pontusson and Rueda (2010) as well as Scheve and Stasavage (2009). The signs and significance of the coefficients for the inequality ratio and its interaction with PR are the same as in the basic specification (models (1) - (3)). The alternative measure of electoral rules consists of the natural log of the average district magnitude derived from the data set of Bormann and Golder (2013). While perhaps not as closely related to the theoretical argument, it captures the fact that there may be considerable variation in district magnitude among systems coded as PR or majoritarian and a larger magnitude is associated with more proportionality and, ultimately, coalition government. Similar measures have been used widely in the literature on the fiscal effects of electoral systems (Persson and Tabellini, 2003).\footnote{Concretely, the measure corresponds the average magnitude in the first tier in the data set from Bormann and Golder (2013) if the electoral system has no second tier. In the mixed proportional systems of Germany, New Zealand (since 1996 and Italy between 1994 and 2001), the average district magnitude is measured based on the second-tier districts. In the case of Japan since 1996 and Denmark, the calculation takes into account both tiers.}

Models (4) - (6) indicate that in systems with a district magnitude of one (i.e., majoritarian systems), a rise in top inequality leads to a rightward move of the left. The ideological effect on top inequality is conditional on district magnitude. It declines as district magnitude increases and changes sign for a sufficiently large magnitude (of about 11, according to model (4)).
Lowe et al. (2011) provide an alternative measure of the left-right position of each party based on the raw data collected by the Comparative Manifesto Project (CMP) (Budge et al., 2001). Similar to the measure calculated by Elff (2013), it deals with important limitations of the standard left-right scale provided by the CMP. In particular, they use a logit transformation to transform the underlying counts of left and right sentences in the manifestos and argue that the logit scale has better theoretical properties and find that it correlates stronger with independent expert surveys than the original CMP measure. While the measurement approach of Lowe et al. (2011) is different than that of Elff (2013), the estimation results are essentially the same. This is shown by Table 3. The exception is Model (3) that includes country-specific time trends. Given the time varying controls and country-specific year trends, there is little within-country variation left to be exploited. Altogether, the results are robust to alternative measurements of the three key theoretical concepts relevant for the analysis: party positions, top inequality, and electoral rules.

**Additional controls**

As shown by Table 4, the results are also robust to controlling for additional variables used in the literature on party positions (union density, EU integration, incumbent cabinet, left competitor). In the power resource approach, union density is linked to the strength of left parties as well as to redistributive policies. Union density measures union membership as a proportion of wage and salary earners in employment. From 1960 onwards, it is from Jelle Visser’s ICTWSS Database (Visser, 2013). For the years between 1950 and 1960, it is from Miriam Golden’s dataset on Union Centralization (Golden, 2009). Several authors have also suggested that European integration has restricted the police menu available to left parties. Following Nanou and Dorussen (2013), EU integration is a multiplicative term between membership in the EU and the competencies of the EU in economic policy, where competencies in economic policy is the sum of integration levels (ranging from 1 to 5) for all 15 economic
policy areas minus 15 (to set no integration to 0) and divided by 60 (so variable ranges between 0, no integration, to 1, full integration). The measure of EU integration is from Hooghe and Marks (2001). Some studies also suggest that incumbency entails moderation (Schumacher, de Vries and Vis, 2013). The variable is an indicator for cabinet membership taken from Döring and Manow (2012). Furthermore, following Pontusson and Rueda (2010), the analysis also controls for the existence of a left competitor. This is an indicator variable set to 1 if there is a party competing in the election on the left of the mainstream left party. It is calculated based on the positional data from Elff (2013). To be considered serious, a left competitor needs to receive at least 3 percent of the vote or win at least one seat in parliament.

Another question is whether candidate selection rules influence the relationship between inequality and party positions. The theoretical model does not assume or imply that the national party leadership fully controls the candidate selection process. However, it assumes that parties rather than voters select the candidates that run for national parliamentary office under the party label. This is true for all parties included in the analysis, but there is some variation in how the parties under study select their candidates. To capture this, I draw on data concerning the centralization of the candidate selection process from Ranney (1981), Bille (2001), and Lundell (2004). The sources cover all countries included in the analysis except Portugal. The Centralization variable constructed from these sources varies from 0 (subnational party organs alone control candidate selection) to 1 (national party organs make the final decision). The mid-point of the variable indicates that national organs either have veto power over lists proposed by the subnational organs or they draw up the lists subject to approval from the subnational organs. Subnational organs control candidate selection in, for example, Germany and Sweden. Note that this is not equivalent with full decentralization via local primaries as in the United States. Rather, this means that regional party organizations draw up lists and select them with the involvement of
card-carrying members. National organs have more power but do not decide alone in, for example, the Labour Party in Britain or the Socialist Party in France. Spain is a case of high centralization. Strategic party politicians can select moderate types through a variety of candidate selection procedures. This suggests that variations in the selection procedure observed in the data should not fundamentally alter the results. Furthermore, Iversen and Soskice (2006) argue that centralization is endogenous to parties’ efforts to achieve commitment. Thus, including it may lead to biased estimates of the effect of inequality. A priori, it is not clear if this outweighs concerns about omitted variable bias. Table 5 shows the results when centralization of candidate selection and its interaction with PR are included in the panel regression. Reassuringly, the two key coefficients on inequality and the interaction between PR and inequality, exhibit the same pattern as before and are precisely estimated. Interestingly, estimates indicate that the impact of candidate selection rules on party positions varies across electoral systems.

The theoretical argument emphasizes the importance of top income inequality and the measure used in the analysis reflects this concern. To further rule out that the observed relationship is not driven by changes in other parts of the income distribution, it is desirable to include a proxy for changes in the income distribution as a whole. The Gini coefficient concerning household income before taxes and transfers is the most commonly used measure for this in the political economy literature (e.g., Persson and Tabellini, 2003). It is typically calculated from household surveys. As income items in surveys are typically top-coded and random samples do not capture the very top of the income distribution very well, the Gini coefficient is not as sensitive to changes in inequality based on growing top incomes as the measures calculated by Atkinson, Piketty and Saez (2011) and their colleagues. This makes the Gini index a useful control capturing other aspects of inequality, including changes at the bottom of the income distribution. Unfortunately, the required Gini index is not available for the whole period under study. Solt (2014) has constructed the most comprehensive available
series of pre-fisc or market Gini coefficients from 1960s onwards that combines information from high quality surveys such as the Luxembourg Income Study with additional surveys and uses multiple imputation to fill in the gaps. As an additional robustness test I have included Solt’s Gini measure and its interaction with PR into the regression equation. Due to completely missing Gini coefficients for the first decade covered by the data set, this robustness analysis drops 16% of the elections. The correlation between the Gini coefficient and the inverted Pareto coefficient characterizing top incomes is 0.36, consistent with the known fact that the two measures capture somewhat different features of the income distribution.

Table 6 shows the results from the fixed effects regression. Models 1-3 repeat the basic specifications using top income inequality, without the Gini variables, in the smaller sample for which the Gini coefficient is available. This demonstrates that the main findings hold up in the subsample. Models 4-6 drops the top inequality variables and instead include the Gini coefficient and its interaction with PR. The results from this analysis provide an indication that the findings based on top inequality are not driven by other aspects of inequality. While the signs of the relevant variables suggests that the Gini index appears to be associated with more right-ward positions in majoritarian electoral systems but not PR, the standard errors are large and so the effects are never individually or jointly significant. Also note that the adjusted $R^2$ is considerably lower in the these models compared to the comparable specifications based on top inequality. Finally, models 7-9 include both top inequality and the Gini coefficient and their interaction with PR. This specification is quite demanding. Nonetheless, the estimated effects of top inequality are only slightly smaller these models than in restricted models 1-3 and while not all coefficient reach standard levels of significant, most do and otherwise they are close. In fact, Hausman tests fail to reject the null hypothesis that the coefficients vary systematically across the restricted specification only including top inequality and the full specification.² Hence, models 1-3 should be more efficient and thus

²The p-value for the Hausman test statistic comparing model 1 and model 7 in Table 6 is 0.2; p = 0.3 for
deliver more reliable measures of estimation uncertainty. In contrast, the estimated impact of the Gini measures of inequality is reduced by more than half in the full specification and the effects are never statistically significant. In model 9, the sign of the interaction term actually flips. Taken together, the results in Table 6 back up the interpretation that top inequality is an important and institution-varying determinant of the electoral position of mainstream left parties.

Table 7 reports estimates including a lagged dependent variable. It is well-known that combining unit fixed effects and a lagged dependent variable leads to biased estimates and that the Nickell bias goes to zero as the number of time periods becomes large. Here the number of periods is moderately large (14.2 on average), so the bias is not a first order problem (Beck and Katz, 2011). The coefficient on the lagged dependent variable is relatively small (only about 1/3 including controls) and throughout the two coefficients of substantive interests behave as before. Top income inequality exhibits a positive relationship with the left-right position of the main left party in majoritarian systems. In proportional systems, the relationship is reversed.

**Instrumental variable estimation**

An additional robustness check exploits plausibly exogenous variation in top income inequality. The basic intuition for the identification strategy is based on the fact that the increase in top income inequality is driven to an important party by a dramatic increase in executive pay and this occurred first in the United States (e.g., for a discussion see Atkinson, Piketty and Saez, 2011). This rise was exogenous to redistribute politics in the smaller economies outside the United States. Through explicit compensation benchmarking and partially integrated market for top managers, the dramatic rise in executive compensation in the US led to exogenous competitive wage pressures on executive pay other developed countries. Economists

delivered express measures of estimation uncertainty. In contrast, the estimated impact of the Gini measures of inequality is reduced by more than half in the full specification and the effects are never statistically significant. In model 9, the sign of the interaction term actually flips. Taken together, the results in Table 6 back up the interpretation that top inequality is an important and institution-varying determinant of the electoral position of mainstream left parties.

Table 7 reports estimates including a lagged dependent variable. It is well-known that combining unit fixed effects and a lagged dependent variable leads to biased estimates and that the Nickell bias goes to zero as the number of time periods becomes large. Here the number of periods is moderately large (14.2 on average), so the bias is not a first order problem (Beck and Katz, 2011). The coefficient on the lagged dependent variable is relatively small (only about 1/3 including controls) and throughout the two coefficients of substantive interests behave as before. Top income inequality exhibits a positive relationship with the left-right position of the main left party in majoritarian systems. In proportional systems, the relationship is reversed.

**Instrumental variable estimation**

An additional robustness check exploits plausibly exogenous variation in top income inequality. The basic intuition for the identification strategy is based on the fact that the increase in top income inequality is driven to an important party by a dramatic increase in executive pay and this occurred first in the United States (e.g., for a discussion see Atkinson, Piketty and Saez, 2011). This rise was exogenous to redistribute politics in the smaller economies outside the United States. Through explicit compensation benchmarking and partially integrated market for top managers, the dramatic rise in executive compensation in the US led to exogenous competitive wage pressures on executive pay other developed countries. Economists
and sociologists have studied the increasing use of explicit compensation benchmarking by companies to determine executive pay and several studies link this practice to the increase in top inequality (e.g., DiPrete, Eirich and Pittinsky, 2010; Faulkender and Yang, 2010). There is little reason to believe that, given the controls, top inequality in the US has a direct influence of left party positions in other countries. Note that this identification strategy is only possible for the subset majoritarian electoral systems as, so far, the field has not generated credible instruments for electoral rules (Acemoglu, 2005). The instrumental variable estimates reported in Table 8 confirm the main results from the basic panel analysis. In majoritarian electoral systems, increased top inequality caused left parties to move to the right. Note that there is a strong and statistically significant first stage relationship between top inequality in the US and that in the other countries. This holds after including all the control variable and time trends. The second stage estimates are in line with the OLS estimates. If anything, the estimated effect of top inequality is somewhat larger.
Table 2: Panel Results Using Alternative Measures of Top Income Inequality and Electoral Institutions

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Income Share (TIS)</td>
<td>0.10**</td>
<td>0.09**</td>
<td>0.07**</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Proportional Representation (PR)</td>
<td>-0.03</td>
<td>-0.06</td>
<td>-0.09</td>
<td>(0.05)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>PR × TIS</td>
<td>-0.16**</td>
<td>-0.16**</td>
<td>-0.15*</td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Top Income Inequality (TI)</td>
<td>0.82**</td>
<td>0.69*</td>
<td>0.61</td>
<td>(0.30)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>District Magnitude (DM)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>DM × TI</td>
<td>-0.34**</td>
<td>-0.32**</td>
<td>-0.29</td>
<td>(0.10)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Country-specific time trend</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Period indicators</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Countries</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Observations</td>
<td>227</td>
<td>227</td>
<td>227</td>
<td>228</td>
<td>228</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.336</td>
<td>0.351</td>
<td>0.350</td>
<td>0.322</td>
<td>0.329</td>
</tr>
</tbody>
</table>

Standard errors in parentheses are clustered by country. The dependent variable is the left-right position of the main left party as estimated from manifestos by Elff (2013). All models include country fixed effects and the following time-varying controls: income per capita, voter turnout, trade openness, population over 65, log of population, time trend (year).

* p < 0.10, ** p < 0.05 (two-tailed tests)
Table 3: Panel Results Using an Alternative Dependent Variable

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Income Inequality (TI)</td>
<td>0.90***</td>
<td>0.75**</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.29)</td>
<td>(0.40)</td>
</tr>
<tr>
<td>Proportional Representation (PR)</td>
<td>-0.10</td>
<td>-0.12</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(0.32)</td>
<td>(0.55)</td>
</tr>
<tr>
<td>PR × TI</td>
<td>-1.13**</td>
<td>-1.11**</td>
<td>-0.43</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.42)</td>
<td>(0.62)</td>
</tr>
<tr>
<td>Country-specific time trend</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Period indicators</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Countries</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Observations</td>
<td>228</td>
<td>228</td>
<td>228</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.135</td>
<td>0.151</td>
<td>0.199</td>
</tr>
</tbody>
</table>

Standard errors in parentheses are clustered by country. The dependent variable is the left-right position of the main left party as estimated from manifestos by Lowe et al. (2011). All models include country fixed effects and the following time-varying controls: income per capita, voter turnout, trade openness, population over 65, log of population, time trend (year).

* $p < 0.10$, ** $p < 0.05$ (two-tailed tests)
Table 4: Panel Estimates Including Additional Control Variables

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Income Inequality (TI)</td>
<td>0.93**</td>
<td>0.84**</td>
<td>0.92**</td>
<td>0.88**</td>
<td>0.83**</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.31)</td>
<td>(0.32)</td>
<td>(0.31)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>Proportional Representation (PR)</td>
<td>-0.10</td>
<td>-0.21</td>
<td>-0.16</td>
<td>-0.13</td>
<td>-0.21</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.13)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>PR × TI</td>
<td>-1.25**</td>
<td>-1.22**</td>
<td>-1.20**</td>
<td>-1.16**</td>
<td>-1.22**</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
<td>(0.29)</td>
<td>(0.34)</td>
<td>(0.35)</td>
<td>(0.33)</td>
</tr>
<tr>
<td>Union Density</td>
<td>0.30</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.91)</td>
<td>(0.86)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EU Integration</td>
<td>-1.04*</td>
<td></td>
<td>-1.16**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.53)</td>
<td></td>
<td>(0.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incumbent Cabinet</td>
<td>0.05</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left Competitor</td>
<td>0.06</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Countries</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Observations</td>
<td>223</td>
<td>228</td>
<td>228</td>
<td>228</td>
<td>223</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.321</td>
<td>0.358</td>
<td>0.333</td>
<td>0.333</td>
<td>0.348</td>
</tr>
</tbody>
</table>

Standard errors in parentheses are clustered by country. The dependent variable is the left-right position of the main left party as estimated from manifestos by Elff (2013). All models include country fixed effects and the following time-varying controls: income per capita, voter turnout, trade openness, population over 65, log of population, time trend (year).

* $p < 0.10$, * $p < 0.05$ (two-tailed tests)
Table 5: Panel Results Controlling for Centralization of Candidate Selection

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Income Inequality (TI)</td>
<td>0.95**</td>
<td>0.85**</td>
<td>0.90**</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.35)</td>
<td>(0.33)</td>
</tr>
<tr>
<td>PR × TI</td>
<td>-1.09**</td>
<td>-1.07**</td>
<td>-1.47**</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.35)</td>
<td>(0.39)</td>
</tr>
<tr>
<td>Centralization</td>
<td>0.28</td>
<td>0.39</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.29)</td>
<td>(0.33)</td>
</tr>
<tr>
<td>PR × Centralization</td>
<td>-1.49**</td>
<td>-1.51**</td>
<td>-1.48**</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.31)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>Proportional Representation (PR)</td>
<td>0.75**</td>
<td>0.76**</td>
<td>0.76**</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.21)</td>
<td>(0.32)</td>
</tr>
<tr>
<td>Country-specific time trend</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Period indicators</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Countries</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Observations</td>
<td>220</td>
<td>220</td>
<td>220</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.354</td>
<td>0.363</td>
<td>0.372</td>
</tr>
</tbody>
</table>

Standard errors in parentheses are clustered by country. The dependent variable is the left-right position of the main left party as estimated from manifestos by Elff (2013). All models include country fixed effects and the following time-varying controls: income per capita, voter turnout, trade openness, population over 65, log of population, time trend (year).

* $p < 0.10$, * $p < 0.05$ (two-tailed tests)
Table 6: Panel Results For Gini-Measure of Inequality

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI</td>
<td>0.69**</td>
<td>0.62</td>
<td>0.79*</td>
<td></td>
<td></td>
<td>0.59</td>
<td>0.51</td>
<td>0.77*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(0.42)</td>
<td>(0.37)</td>
<td></td>
<td></td>
<td>(0.47)</td>
<td>(0.58)</td>
<td>(0.38)</td>
<td></td>
</tr>
<tr>
<td>PR × TI</td>
<td>-1.27**</td>
<td>-1.25**</td>
<td>-2.07**</td>
<td></td>
<td>-1.18*</td>
<td>-1.12*</td>
<td>-2.08**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.44)</td>
<td>(0.65)</td>
<td></td>
<td>(0.60)</td>
<td>(0.63)</td>
<td>(0.62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td></td>
<td></td>
<td></td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>PR × Gini</td>
<td></td>
<td></td>
<td></td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.01</td>
<td>-0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>PR</td>
<td>-0.21*</td>
<td>-0.20</td>
<td>-0.14</td>
<td>-0.17</td>
<td>-0.14</td>
<td>-0.02</td>
<td>-0.22</td>
<td>-0.19</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.11)</td>
<td>(0.24)</td>
<td>(0.10)</td>
<td>(0.11)</td>
<td>(0.14)</td>
<td>(0.13)</td>
<td>(0.12)</td>
<td>(0.23)</td>
</tr>
</tbody>
</table>

Country-specific
time trend  No  No  Yes  No  No  Yes  No  No  Yes

Period indicators  No  Yes  No  No  Yes  No  No  Yes  No

Countries  16  16  16  16  16  16  16  16  16

Observations  192  192  192  192  192  192  192  192  192

Adjusted $R^2$  0.309  0.314  0.383  0.281  0.287  0.329  0.302  0.308  0.376

Standard errors in parentheses are clustered by country. The dependent variable is the left-right position of the main left party as estimated from manifestos by Elff (2013). All models include country fixed effects and the following time-varying controls: income per capita, voter turnout, trade openness, population over 65, log of population, time trend (year).

$^*$ $p < 0.10$, $^* $ $p < 0.05$ (two-tailed tests)
Table 7: Panel Estimates Including a Lagged Dependent Variable

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous Left-Right Position</td>
<td>0.54**</td>
<td>0.35**</td>
<td>0.33**</td>
<td>0.29*</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.13)</td>
<td>(0.12)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Top Income Inequality (TI)</td>
<td>0.53**</td>
<td>0.70**</td>
<td>0.67**</td>
<td>0.57*</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.19)</td>
<td>(0.24)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>Proportional Representation (PR)</td>
<td>-0.02</td>
<td>-0.13*</td>
<td>-0.13</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.07)</td>
<td>(0.09)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>PR × TI</td>
<td>-0.96**</td>
<td>-1.13**</td>
<td>-1.13**</td>
<td>-1.25**</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.21)</td>
<td>(0.23)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>Country-specific time trend</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Period indicators</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Countries</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Observations</td>
<td>224</td>
<td>224</td>
<td>224</td>
<td>224</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.334</td>
<td>0.413</td>
<td>0.414</td>
<td>0.424</td>
</tr>
</tbody>
</table>

Standard errors in parentheses are clustered by country. The dependent variable is the left-right position of the main left party as estimated from manifestos by Elff (2013). All models include country fixed effects and all models (2)-(4) include the following time-varying controls: income per capita, voter turnout, trade openness, population over 65, log of population, time trend (year).

* $p < 0.10$, ** $p < 0.05$ (two-tailed tests)
Table 8: Instrumental Variables Estimates of the Impact of Inequality on Left Party Positions in Majoritarian Parliamentary Democracies

<table>
<thead>
<tr>
<th></th>
<th>First Stage Estimates</th>
<th>Second Stage Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Top Income Inequality</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td></td>
</tr>
<tr>
<td>US Top Inequality</td>
<td>0.62**</td>
<td>0.50**</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Income per capita</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Voter Turnout</td>
<td>-0.12</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Population Over 65 (%)</td>
<td>-0.05</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Log of Population</td>
<td>0.17</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>Time Trend (Year)</td>
<td>-0.01</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Country-specific trend</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Period indicators</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Countries</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Observations</td>
<td>106</td>
<td>106</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.704</td>
<td>0.722</td>
</tr>
</tbody>
</table>

Standard errors in parentheses are clustered by country. All models include country fixed effects.

* $p < 0.10$, * $p < 0.05$ (two-tailed tests)
S4. Background on Analysis of Candidate Selection

Scottish Parliament

The Scotland Act of 1998 established a unicameral parliament with the power to make laws for Scotland in a wide range of policy areas, including the power to vary the income tax, and with a government responsible to parliament, mirroring the executive-legislative institutions in Westminster. The Scottish parliament has the power to legislate in all areas that are not explicitly (in the Scotland Act 1998) retained by (reserved for) the British Parliament (which has the right to nullify acts of the Scottish parliament and to change the balance of power). Devolved matters include, among others, health, education, local government, transportation, home affairs, environment, agriculture, sports and arts.\(^3\)

The basic institutional design of the Scottish parliament, including the electoral system, was drafted by a constitutional convention in which Labour did not have a majority. Labour leaders accepted proportional representation to obtain the support of the Liberal Party. Some Labour politicians also considered it a longer term bargain to prevent the Scottish National Party from winning a majority of seats with a minority of votes (Taylor, 2002, pp. 53-56, 64).

As explained in the paper, the electoral system used to elect the 129 members of the Scottish parliament is a mixed member proportional representation system known as the additional member system. There are 73 constituency and 56 regional seats. Until the redrawing of the district boundaries for the House of Commons in 2005, all but two constituencies for the Scottish parliament were identical to the constituencies for the House of Commons. The exception are the Orkney Islands and the Shetland Islands, which form separ-\(^3\)For more information, see McFadden and Lazarowicz (2010).
rate constituencies for the Scottish parliament and one for Westminster. Note that the right to vote is almost identical in both cases with the exception that for the Scottish parliament it extends to all those who are eligible to vote in local elections, including non-British citizens of EU member states.

**Candidate Selection Procedure**

In the period under study, the formal candidate selection procedure in the Labour party involved the national party executive vetting and interviewing prospective candidates before putting an approved list up to a vote, in which each party member has one vote, to determine the candidate for each constituency. The party executive also retains the power to veto selected candidates (for more details, see Biezen and Hopkin, 2006; Hopkin, 2001). This procedure was used to select candidates for the House of Commons and candidates for the Scottish parliament. The literature shows that during the leadership of Tony Blair the party executive used the vetting procedure to influence the nomination process, excluding a large number of candidates “on purely political grounds” (Biezen and Hopkin, 2006, p. 27), consistent with arguments about strategic moderation.

Studies of the candidate nomination process in the Labour party for the first election of the Scottish parliament in 1999 suggest that the national party leadership was concerned with maintaining a unified party image of ‘new’ Labour in Scotland, even at the cost of alienating some voters in the more left-leaning (relative to national median) Scottish electorate, and probably also in order to avoid giving more radical critics in their own ranks a foothold in the new parliament (Biezen and Hopkin, 2006; Bradbury et al., 2000; Shaw, 2001). There was the perception of “heavy central party manipulation to promote candidates loyal to the New Labour project and to screen out dissidents” (Bradbury et al., 2000, p. 58). Biezen and Hopkin (2006) argue that party officials used the formal vetting process for candidates to exclude a large number of candidates because they were seen as not sufficiently in line
with the centrist stance of new Labour. In one instance the party executive used its ex-post
veto to reject a prospective candidate and sitting member of the House of Commons, Dennis
Canavan, for his left-wing views after he had been endorsed though the membership vote
(Shaw, 2001). This reflects the fact that political competition in different levels of a multi-
level political systems is never fully independent, even when elections and policymaking
in one level are institutionally separate from elections and policymaking in another, as is
the case in Scotland. In particular, the literature on federalism and political decentralization
emphasizes that the objectives and strategies of party leaders are linked across levels (Biezen
and Hopkin, 2006). In the context of this study, however, the countervailing incentives of the
Labour leadership to maintain a unified brand should make it less likely to find differences in
candidates across electoral systems. Thus, the Scottish comparison constitutes a hard test
for the hypothesis.

Also note that in the 1999 election Labour implemented a policy of gender parity in
candidate selection. The party nominated an equal number of male and female candidates
in winnable seats. This may have changed the pool of available candidates in a way that
systematically affects the nomination of union members. In particular, historically union
membership in Great Britain has always been higher among men than women. By the end
of the 1990s, however, the differences had diminished and union membership among both
men and women was about 30 percent (Sneade, 2001).

4The party also introduced a gender quota in the 1997 Westminster election, in the form
of all-women short-lists in 50 percent of all open winnable constituencies. The large number
of incumbents in Westminster meant that the impact of the policy was smaller than in the
new Scottish parliament. Moreover, its full implementation was abandoned after a successful
legal challenge.
Data Sources

The coding of candidates’s union membership is based on multiple sources. For the House of Commons, a standard source is *The Times Guide to the House of Commons* (Times Newspapers, 1997). It contains biographical information on all candidates, including those who did not win a seat, and has been used by previous studies of candidates in British politics (e.g., Eggers and Hainmueller, 2009). There is a special category for memberships in professional or other organizations. For instance, the entry for the Labour candidate in Glasgow Rutherglen, Tommy McAvoy, is as follows:


For the Scottish parliament, there are several sources. One is the electronic archive of candidate materials collected by the University of Strathclyde. Strathclyde’s ASPECT archive is available at [http://gdl.cdlr.strath.ac.uk/aspect/](http://gdl.cdlr.strath.ac.uk/aspect/). For example, the flyer for Glasgow Sprinburn candidate Paul Martin, describes the candidate as follows:

Paul Martin is 32 years old is married and has stayed in the Springburn Constituency all his life.

He has been Local Councillor for over 5 years and has held a number of senior positions in the Council. He is presently the Vice-Chair of the Economic & Industrial Development Committee. He has been an effective campaigner in areas such as Public Transport, Housing, and Education. Paul joined the Labour Party at the age of sixteen and is an active member of the AEEU Union.

Another source is the website of the Scottish parliament, which publishes biographical statements of its members, which include a category for trade union and professional mem-
bership (http://www.scottish.parliament.uk/msps/). For previous members the biographies are shorter, but web archives have often stored the original site. For example, see http://web.archive.org/web/20130707120840/http://www.scottish.parliament.uk/msp/membersPages/margaret_curran/. In addition, The Liberal Democrats also have published biographies of their candidates online (see http://web.archive.org/web/19990508123713/http://www.scotlibdems.org.uk/bioindex.htm). When information was missing, candidates’ personal websites were consulted.

Additional Results

Table 9 reports additional estimation results. All models include several individual-level controls: age, gender an an indicator whether the candidate has a blue collar background. In addition, models (3) and (4) reports estimates when districts won by the Scottish National Party (SNP) are excluded. This reflects the concern that Labour may be especially sensitive in the election for the Scottish parliament to competition by the SNP. These alternative specifications support the initial results reported in the manuscript.
Table 9: Effect of Electoral Rules on Candidate Selection in Scotland: Separating Control Group

<table>
<thead>
<tr>
<th></th>
<th>(1) LPM</th>
<th>(2) FEL</th>
<th>(3) LPM</th>
<th>(4) FEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labour</td>
<td>0.35**</td>
<td>2.23**</td>
<td>0.34**</td>
<td>2.02**</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.44)</td>
<td>(0.06)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>Mixed Member PR (MMP)</td>
<td>-0.05</td>
<td>-0.50</td>
<td>-0.13**</td>
<td>-1.76**</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.39)</td>
<td>(0.04)</td>
<td>(0.76)</td>
</tr>
<tr>
<td>Labour × MMP</td>
<td>0.18**</td>
<td>0.97</td>
<td>0.26**</td>
<td>2.31**</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.60)</td>
<td>(0.09)</td>
<td>(0.88)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.35*</td>
<td></td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td></td>
<td>(0.17)</td>
<td></td>
</tr>
<tr>
<td>District effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>SNP-districts excluded</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>405</td>
<td>320</td>
<td>346</td>
<td>274</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.309</td>
<td></td>
<td>0.316</td>
<td></td>
</tr>
<tr>
<td>Pseudo $R^2$</td>
<td></td>
<td>0.397</td>
<td></td>
<td>0.462</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses. The dependent variable is an indicator for union membership. Col. (1) and (3) report coefficients from a linear probability model, col. (2) and (4) from a fixed effects logit model. Controls include age, gender and blue collar occupation.

* $p < 0.10$, ** $p < 0.05$ (two-tailed tests)
References


URL: [http://www.parlgov.org/](http://www.parlgov.org/)


Feenstra, Robert C., Robert Inklaar and Marcel P Timmer. 2013. “The Next Generation of the Penn World Table [PWT Version 8.0].”.

URL: [www.ggdc.net/pwt](http://www.ggdc.net/pwt)


URL: [http://hdl.handle.net/1902.1/10193](http://hdl.handle.net/1902.1/10193)


URL: http://www.uva-aias.net/208
