

Preferential Voting, Accountability and Promotions into Political Power: Evidence from Sweden*

Olle Folke⁺, Torsten Persson^φ, and Johanna Rickne^ψ

January 9, 2014

Abstract

Preferential voting has been introduced in a number of proportional election systems over the last 20 years, mainly as a means to increase the accountability of individual politicians. But most of these reforms have been criticized as blatant failures. In this paper, we discover a genuinely new fact, which calls into question this negative evaluation. We show that preferential voting in a general election can operate as a stand-in internal primary election for top party positions. To do this, we rely on a unique data set from four waves of Swedish local elections, which includes every nominated politician in each of 290 municipal assemblies. We use a natural-experiment (regression-discontinuity) approach to estimate the causal effect of winning the most preferential votes on becoming the local party leader, and find that narrow “list winners” are over 50 percent more likely to become party leaders than their runner-ups. Comparing across politicians, the effect of list winning is the strongest for competent politicians, who are also more likely to draw preferential votes than mediocre politicians. Comparing across municipalities, the response to narrow list winning is the strongest within unthreatened governing majorities, where voters also use the preferential vote the most frequently.

* The authors thank Jim Snyder, Pär Nyman, Daniel Markham Smith, and seminar participants at the Empirical Political Economics and Political Science Workshop (WATT), MPSA 2013, Uppsala University, Columbia University, Cal Tech, University of California at Berkeley, and Gothenburg University for helpful comments. Financial support from the ERC, the Swedish Research Council, and the Torsten and Ragnar Söderberg Foundation is gratefully acknowledged.

⁺ Columbia University, SIPA; and Research Institute for Industrial Economics; of2152@columbia.edu

^φ IIES, Stockholm University; LSE; and CIFAR; torsten.perssion@iies.su.se

^ψ Columbia University, SIPA; Research Institute for Industrial Economics; and Uppsala Center for Labor Studies; kjr2123@columbia.edu

1. Introduction

A key question in political economics is how institutions shape political behavior. In democracies, elections forge a critical link in the principal-agent chain between voters and politicians. When preferences of politicians and voters are misaligned, accountability to the voters and the threat of being ousted from office can induce politicians to act in the voters' interests (Barro, 1973; Ferejohn, 1986).

But the degree of accountability hinges on the electoral system in place. At the systemic level, a choice of proportional elections is often portrayed as a way to better representation than majoritarian elections, but at the cost of worse accountability (Powell, 1984; Lijphart, 1984; Taagepera and Shugart, 1989; Persson and Tabellini, 2000). Specifically, under Proportional Representation (PR) ballots refer to lists designed by parties – so-called closed lists – rather than specific candidates. As voters are unable to hold individual politicians to account, politicians lack direct incentives to please their party's voters, although they are motivated to please their party to get high enough a list position to secure an assembly seat (Carey and Shugart, 1995).

This fact has not gotten lost on electoral designers. Over the last 20 years, a number of countries with closed-list PR systems have introduced preferential voting – also known as flexible lists – to create a direct link between voters and candidates, as described e.g., for Austria by Mueller (2005) and for the Netherlands by Andeweg (2005).¹ Preferential voting allows voters to select one (or more) of the candidates on the list, and if any candidate amasses enough preferential votes to pass a certain threshold, s/he automatically obtains a seat. A strong argument in favor of these reforms has been that they will raise accountability, since politicians with "safe" positions high up on the ballot will face the threat of being replaced by politicians lower down on the ballot.

However, preferential-voting reforms have not received a passing grade from the scholarly community. Voters tend to cast preferential votes for the "wrong" candidates, those on top of the party list instead of those populating the lower portion. Since safely placed politicians have not really been threatened by replacement, preferential voting systems have been described as little more than "closed lists in disguise" (Farrell, 2001; Mueller, 2005; Andeweg, 2005; De Winter, 2005). For the same reason, ex post evaluations by electoral reformers have been critical. In Sweden, e.g., the preferential-vote threshold for automatic representation will be considerably lower in the 2014 election as a direct result of the criticism (SOU 2008:125).

In this paper, we present a genuinely new finding, which puts the negative verdict on preferential voting in doubt. We use data for all preferential votes cast in four waves of Swedish local elections to show that these votes have had a huge impact on promotions into political power. In effect, preferential voting has worked as a primary election: a politician who wins the most preferential votes of her local party sees her chance of being appointed this party's political leader go up by more than 50 percent. This means that a preferential voting reform can indeed be a forceful tool for increasing accountability, albeit not in the manner expected by reformers.

Why should a political party use the information contained in preferential votes? PR-elections, especially at the local level, take place in a low-information environment. Opinion polls, at best, concern parties rather than individual politicians, so preferential votes convey unique information about which individual politicians are popular among voters. An office-oriented party may find it

¹ Preferential voting is currently used in (at least): Austria, Belgium, the Czech Republic, Denmark, Estonia, Indonesia, Norway, the Netherlands, Slovakia, and Sweden, and is debated in many more PR systems.

rational to reward politicians who receive many preferential votes, since this create strong incentives for earning votes. Such rewards can also help us understand why many politicians "campaign" for preferential votes even though they are already placed high up on the ballot. Previous studies have indeed documented that politicians who generate many preferential votes are moved up to higher positions on the party list in the next election (Mueller, 2005; Crisp et al., 2013).

However, a positive correlation between a politician's preferential-vote support and her future ballot position does not guarantee that getting many votes caused the upward move. If more ambitious and motivated politicians move up the power ladder and also gain more preferential votes, we might mistakenly attribute the career move to the vote support, when the true reason is politician quality. Moreover, because characteristics that drive vote support as well as career advancement are likely unobserved by the researcher, control variables may be insufficient to remove omitted-variable bias in the estimated relation from preferential votes to political power. Thus, we need another method to evaluate whether preferential voting creates direct accountability of politicians to voters.

In this paper, we apply a Regression Discontinuity design, by examining the threshold for winning the most preferential votes within the party group.² If the top two preference-vote earners have similar vote tallies, they are likely similar in terms of both observed and unobserved characteristics. In this quasi-experimental setting, we can consider very close races approximately random draws of winners and losers. A battery of robustness checks confirms that this is indeed the case.

By comparing close winners and losers, we demonstrate that the causal impact of narrowly winning the most preferential votes is a 15 percentage point – more than 50 percent – hike in the probability of becoming the top-ranked candidate in the next election. We also show that this response to voters' preferences is not mere window dressing: a re-shuffling of the list but an intact distribution of political power. With the same empirical design, we examine the appointment directly after the election to the top appointments in the local political hierarchy: full time political positions³ and the chair of the municipal council board. Here, we find a huge impact – more than a doubling of the baseline probability – of winning the most preferential votes in the largest party.

Our result that preferential voting shapes appointments of political leaders is noteworthy in view of their disproportional political influence in PR systems. Local party leaders help formulate the party's political platform, bear the standard of the party, and serve as its main spokesperson. The importance of the position can also help us understand why parties respond to preferential votes. In Sweden, news articles often highlight the preferential vote "winner" and charge parties with weak accountability and ignorance of the popular will for not granting them the top appointment.⁴ This is similar to Katz and Bardi's (1980) observation on the symbolic importance of the top name in Italy's previous (now abandoned) electoral system, where the general expectation was that the "capolista" (list leader) should be the person with the most preferential vote support in the party group.

² Imbens and Lemieux (2008) give an overview of the RD methodology. Lee, et al. (2004) and Petterson-Lidbom (2008) are the first applications to elections. For a more extensive list of papers that have applied RD designs in electoral settings, see Caughey and Sekhon (2012).

³ Full time positions are usually only the chair and the vice chair of the municipal council board, and these positions are held by the top ranked politician from the largest majority and minority parties, respectively.

⁴ An editorial piece in the Swedish newspaper Kristinehamnsbladet pointed out which local party groups had failed to reward the preferential vote winners in the 2006 elections with the most prestigious political positions (Brommesson, 2006). The editor concluded that "the political parties have a curious inability to hand over power to the voters that they will be ruling for the next four years". In a subsequent op-ed, a disappointed voter responded to this story by bursting out that "[w]hat are our local parties doing? Is this what they consider to be ethical and moral behavior? All our preferential votes are being totally ignored [...]. How do we dare to continue voting for parties that have completely side-stepped our democratic principles?"

Although preferential voting may have raised accountability to voters, parties still hold the ultimate power in leader selection. Our analysis shows that winning the preferential vote is important for politicians that have already reached a high enough position in the party hierarchy, approximated by holding one of the three highest positions on the party ballot. In contrast, winning the most preferential votes does not propel a politician from the lower portion of the list to a leadership position. In this way, voter support serves as a complement rather than a substitute for party support.

The fact that party groups adhere to preferential votes raises two questions. Do voters replace bad leaders with good ones? Do preferential votes trigger leader changes even in political monopolies, where the governing party does not need to cater to voters in order to win the election? Our analysis gives a clear and categorical "yes" to both questions.

On the first question, we start by analyzing the characteristics of politicians who receive more preferential votes, and show strong correlations between different measures of competence and vote support. We then provide evidence that the treatment effect from winning the most preferential votes is stronger for competent politicians. Although these results suffer in precision due to small samples, voters appear at least as good at naming leaders as parties. Moreover, narrow "losers" in the preferential vote contest are no more likely to leave the party group than the winners. Competent losers that receive large shares of preferential votes are thus not lost to the party: they remain active in local politics, although they give up the leader position to the preferential-vote winner.

On the second question, we compare Swedish municipalities with regard to dominant parties. Many municipalities have seen a single political party hold power for most of its modern history. Just as industrial firms with large market shares, one may suspect that monopoly parties may lack incentives to cater to customers. However, we show that the responsiveness to the preferential vote is actually the strongest in dominant party groups. This result flies less in the face of the monopoly parallel if we consider voter participation, which is actually larger in dominant parties. Supposedly, this makes any signal of discontent stronger.

That preferential voting can serve as an accountability mechanism under strong party dominance suggests a parallel with primary elections in majoritarian systems. In the US, such elections were introduced to break out of a non-competitive status quo. Power was moved from parties to voters, especially in places like the South where the Democrats had a near monopoly on power from the 1880s to the 1950s. Primary elections became a democratic outlet for selecting the candidate who was assured to win the subsequent general election. Thus, primary elections may be adopted by heterogeneous parties wishing to avoid fractionalization (Hortala-Vallve and Mueller, 2012), and by dominant parties facing little external competition (Ansolabehere et. al., 2007; Serra, 2006; Kemahlioglu, et. al. 2009).

In the following, Section 2 gives a background by describing Sweden's reform towards preferential voting, as well as its system of municipal politics. Section 3 describes our data and measurement. Then, Section 4 presents descriptive evidence on how Swedish voters use their preferential votes. Section 5 introduces the Regression Discontinuity design used to obtain our main findings, which are presented in Section 6. Sections 7 and 8 discuss the two extensions, corresponding to the competence of promoted candidates, and the divisions by political context (majority vs. minority party, and high vs. low competition municipality). Section 9 concludes. Some auxiliary results and robustness checks are relegated to an Appendix.

2. Background

The Preferential Voting Reform

The debate on preferential voting started in Sweden almost a hundred years ago. In 1909, when the majoritarian election system was replaced by multimember districts and closed lists, preferential voting was discussed as a way to curtail the increased influence of political parties. Over time, various reform proposals were put forth and rejected. Mandatory preferential voting was perceived as risking an "Americanization" of the Swedish politics. Incentives to run individual campaigns, where private funds could be used to gain elected positions, was (and is still) seen as undemocratic.

In the early 1990s, a government white paper suggested a compromise. As before, voters would face ordered lists with the party's top politician at the top. Seats would continue to be counted from the top: for a party winning three seats, the top three persons would get one each. The news was that each voter got the option of casting one voluntary preferential vote. Little boxes were to be placed next to each politician's name on the electoral ballot, to allow the voter to tick one of them (see Figure 1). If a politician would amass enough votes to clear a threshold level, he or she would be guaranteed a seat. The remaining seats obtained by the party would then be awarded to the candidates with the highest list rank who did not clear the threshold.

The Swedish political parties largely found this approach both simple and balanced. The principle of party rather than individual representation was preserved, as voluntary preferential

Figure 1 – Example of Electoral Ballot

VAL TILL KOMMUNFULLMÄKTIGE

Moderata Samlingspartiet

Du får bara markera en av dessa **anmälda** kandidater.

☐	1	Hans Jonsson, Lantbrukare, Ringarum
☒	2	Anna Nilsson, Leg. Sjuksköterska, Fil.mag., Gryt
☐	3	Karin Magnusson, Fritidspedagog, Valdemarsvik
☐	4	Monica Stillnert, Fru, Ringarum
☐	5	Hans Andersson, Key Account Manager, Valdemarsvik
☐	6	Per Hollertz, Lantbrukare, Redovisningskonsult, Ringarum
☐	7	Anita Esbjörnsson, Revisor, Valdemarsvik
☐	8	Charlotta Hollertz, Agronom, Mäklarassistent, Ringarum
☐	9	Jan Ekroth, Företagare, Östra Ed
☐	10	Göran Österdahl, Projektledare, Ringarum
☐	11	Lennart Andersson, Yrkesofficer, Valdemarsvik
☐	12	Tord Andersson, Egen företagare, Valdemarsvik
☐	13	Lars Ekblad, Konsult, Valdemarsvik
☐	14	Torbjörn Stackling, Företagsekonom, Gryt
☐	15	Rolf Swärd, F.d. officer, Gryt
☐	16	Joel M. Hodt, Organisationskonsult, Valdemarsvik
☐	17	Per Gunnarsson, Lantbrukare, Östra Ed

Valdemarsviks Kommun
0001—01416

Notes: Electoral ballot for the Conservative Party in Valdemarsvik municipality, with a preferential vote cast for Anna Nilsson, an authorized nurse (Leg. Sjuksköterska), holding a university degree (Fil. Mag.), and residing in the village of Gryt.

voting gave voters the option of supporting the party's candidate order by abstaining from casting their preferential vote. With a sufficiently high threshold, parties also felt protected from dramatic changes to their candidate order. Thresholds were set to 8% of the party's total vote for the national parliament and 5 percent of the party's total vote (with at least 50 votes) for local elections. The new system was introduced in 1998 for all three levels of government: the one-chamber national parliament, the 20 county assemblies, and the 290 municipal councils.

Sweden's Municipal Politics and Institutions

Compared to many other countries, municipalities in Sweden have substantial political power. They have the right of self-government, guaranteed by the Swedish Instrument of Government (1991 Local Government Act 2.1). This includes the power to tax. In 2010, the average local income tax was 21.6%. Through their responsibility for large areas of social spending – schools, child care and elderly care – municipalities are the main employer of about a quarter of Sweden's total labor force. Because of this political importance, a leadership position in a large municipality is generally considered as a position of greater political power than a seat in the national parliament.

The municipal council (the local parliament) is led by the council board (the local government). The chairperson of the board (the closest equivalent to a mayor) is indirectly selected from the governing party, or the largest party in the governing coalition.⁵ Specific policy areas are dealt with in subcommittees, of which the average council has seven, with chairpersons appointed from the governing majority. The board, assembly, and committee chairpersons are the most influential politicians, but the chairperson of the municipal council is often the only person employed as a full-time politician.⁶

Sweden's party system is highly stable with seven political parties represented in nearly all municipal assemblies. Small local parties hold about 4-5% of the total council seats. The only electoral threshold is defined by district magnitude, which ranges from about 1.5 to 5 percentage points of the vote.

Nomination procedures for local elections take place in local party organizations. Within the municipality, members are organized into clubs based on either geographical areas or party sub-organizations. Large parties typically have clubs for the youth league, women, university students, and (in left-wing parties) trade-unionists. The local party has a leader who leads the group in formulating policy proposals functions and acts as its political spokesperson. Standard procedure is to place this leader on top of the ballot and appoint him or her to the top political position available to the party. Specifically, the person at the top of the largest party in the governing majority is generally appointed chairperson of the municipal council board.

Party lists are composed in two steps. In the two left-wing parties (Social Democrats and Left Party), clubs nominate members to a committee that aggregates the nominations into a proposed list. In the other parties, the nomination committee organizes test elections, in which all party members can vote. On the face of it, such an internal primary may seem open to turnover within the

⁵ Municipalities differ widely in both land area (from 9 to 19,447 square kilometers) and population (from 2,558 to 780,817 inhabitants). Population differences are reflected in different-sized municipal councils, from at least 31 seats to a maximum of 101, with an average of 40. Seats are allocated across parties in proportion to their vote shares in electoral districts with a minimum magnitude of 15 seats.

⁶ Others receive piece-rate compensation for meetings and offices, and less than 10% of politicians receive more than 40% of a standard full-time salary. A survey carried out in 1991 shows that an average council member spends 8.3 hours per week on her duties, while a chairperson spends 17.8 hours.

list, but in reality both nomination procedures strongly favor the status quo. Dominant factions within the party control the aggregation of nominations, both through the clubs and the committee which makes up the final list. In the internal primaries, votes are mobilized by strong actors and networks, which organize around candidates. These networks are important, since only a minor fraction of the members use their right to vote.

Also important is that local elections take place in a low-information environment. There are typically not even opinion polls for parties, let alone for individual politicians. Still, about one quarter of the electorate split their vote between the national and municipal level. The extent of such vote splitting has been shown to vary with local political conditions (Karlsson, 2012). Interviews conducted for this paper suggest that local party groups do not automatically expect the same election result as in the national election. It appears to be commonly understood that vote splitting in the national and local election can arise rapidly to punish or reward local parties. Numerous local party groups either fail miserably or by far outperform the national party.⁷

3. Data and Measurement

Candidate Characteristics

The data for our analysis refer to four waves of elections in Sweden's 290 municipalities, from 1998, when preferential voting was introduced, until the most recent election in 2010. For each of these four elections, we have collected all electoral ballots for all political parties. These ballots are registered at the Swedish election agency with the personal-identification code of each politician and then stored at Statistics Sweden. Using these identification codes, we can link each politician to extensive socioeconomic information from other registers. In the combined dataset, we can follow politicians over time, knowing how many preferential votes they receive in each election, their list placements on the electoral ballots, and whether their socioeconomic status such as education, occupation and income changes. For the two most recent election periods, we also have additional data on the political appointments held by each politician. This includes positions as municipal-council executives, in the council board, and on committees handling specific policy areas.

Preference Votes

The most important political variable in our dataset is the number of preference votes obtained by each politician in each election. Drawing on this information, we construct a binary treatment variable that takes a value of one for the person who gets the largest number of preferential votes of all candidates on the party list in a particular election. Around the point when a top candidate just barely beats his or her most important challenger to the highest preferential vote count, assignment

⁷ A typical example is the municipality of Sigtuna, north-west of Stockholm. Here, the Social Democrats have gained ten percentage points of electoral support in the last three elections, while the national party lost about the same fraction of votes. In the 2010 election, 14 percentage of the local electorate split their votes to support the local Social Democrats instead of their first choice for the parliamentary election. This success is largely accredited to energetic local party leader Anders Johansson. In 2010, he received 40% of the party's preferential vote, 6-7 times more than the closest challenger.

Another example is the Left party in Fagersta municipality, in the middle of Sweden. The party's local leader, Stig Henriksson, has been described as a charismatic opponent to the party's national policy. In the 2006 election, the local Left party received 58% of the vote, while the national party only received 9%. Even though the confidence crisis of the local Social-Democratic party in the 1990s may have helped pave the way for this spectacular success, it is commonly considered a Stig-Henriksson effect (Etzler, 2008). In 2010, Henriksson received a staggering 89% of the party's total preferential vote, fifty times more than his main challenger.

to treatment can be considered random. Thus, when the preferential-vote count is very close, we do not expect any *systematic* differences in the political and socioeconomic characteristics of politicians close to the threshold. This lack of differences is the identifying assumption of our RD design: it is further explained and empirically verified in subsequent sections of the paper.

Although they are outcomes in a PR-election system, our “close elections” are defined in a similar manner as in the archetypical RDD study of a majoritarian election system with two candidates. We compute the relative support of the winner of the preferential-vote tally relative to his or her “top contender”, i.e., the politician on the list with the second largest number of preferential votes. We then divide this vote difference by the total number of votes for the two candidates, which gives the margin of victory.⁸ This variable has a natural threshold of zero, the point at which the winner and top contender have equal preferential vote shares.

Party Leadership

Our main dependent variable is a binary indicator of holding the position as municipal party leader at the time of the next municipal election. We define this as holding the top position on the party ballot. This top-ranked candidate is almost always awarded the top appointment in the municipal political hierarchy available to each party, which we can verify using data for appointments following the 2006 and 2010 elections. In nine cases out of ten, the chairperson of the municipal council board is the top-ranked person in the largest political party of the governing majority. In eight cases out of ten, the vice chairperson is the top-ranked person from another party, usually the largest opposition party. In seven cases out of ten, when a smaller party has a seat on the municipal council board, this is occupied by their top-ranked politician. In Section 6 below, we also provide direct evidence that the effect we estimate on becoming the top-ranked candidate on the ballot in the next election extends to receiving the party’s top political appointment directly after the current election.

4. Descriptive Evidence: Preferential Voting Behavior

Swedish elections are held every fourth year with a very high turnout, typically in the 80-90 percent range of eligible voters. About one third of these voters also cast a preferential vote. Surveys show that preferential voters are generally more knowledgeable about politics and identify more strongly with a certain political party (SOU 2007:68). Their socioeconomic characteristics are fairly representative for the population at large with a slight over-representation of foreign-born, highly educated, and younger people (Holmberg and Oscarsson, 2013).

Surveys conducted shortly after the preferential voting reform could dismiss the fear that voters did not understand the mechanisms of the new system (SOU 1999:92).⁹ Interestingly, surveys have also dismissed that voters abstain from preferential voting in order to signal support for the party’s candidate order. Instead, the most common reason for abstention, self-reported by half the

⁸ This is the equivalent of the margin of victory in a two-candidate election in a majoritarian system.

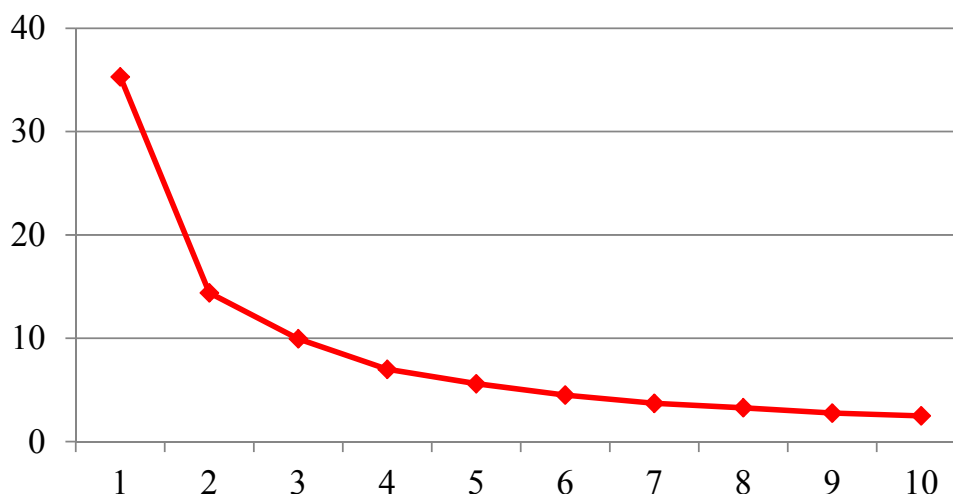
⁹ The Swedish government spent a lot of resources on informing voters. The Swedish system is also substantially less complicated than rules in some other countries with flexible lists. As noted by Mueller (2005) in the case of Austria, a lack of information about the system in popular media was likely “to cause many voters [to] tick the first name on the list just to be sure that their vote will be valid”.

respondents in various surveys, is a perceived lack of knowledge about the candidates,¹⁰ and the second-most common reason is a dislike for preferential voting (SOU 2007:68).

Abstention due to a lack of information on candidates is perhaps not surprising since individual political campaigns for preferential vote are modest. Voters need to figure out more than the basic information about these relatively unknown politicians than what is provided on the ballot (see Figure 1), and also weigh them against the characteristics of the party's own suggested candidates which populate the higher portion of the list. Selecting a politician from the lower portion of the list hence requires extensive research.

Largely ignored in the discussion of the Swedish preferential voting system is the fact that voters have incentives to protect candidates that they like. If a voter favors one of the top names and does not have information on the voting by others, the favorite candidate can be replaced unless he or she amasses more preferential votes than the challengers. This creates a strong incentive to vote for one's preferred candidate rather than to vote "strategically" for persons at lower ballot ranks. As in other countries with flexible list systems, the list rank of a candidate is the single most important predictor of preferential votes in Sweden (SOU 1999:92; SOU 2007:68).

Figure 2 – Preferential Votes Obtained by List Rank



Notes: The figure shows the preferential votes cast for different candidates at different positions on the list as shares of the party's total number of votes in the municipal election.

Figure 2 shows the average fraction of preferential votes within a party group received by candidates on list positions one through ten. On average, a candidate placed at the very top of the list receives more than a third of all preferential votes (or 12 percent of the total vote).¹¹ The average fraction received by the second-in-rank is less than half this number. The curve flattens out gradually.

¹⁰ Nevertheless, half of all voters, and a fifth of those who said a lack of knowledge caused them to abstain, could still name at least one politician for whom they could have voted.

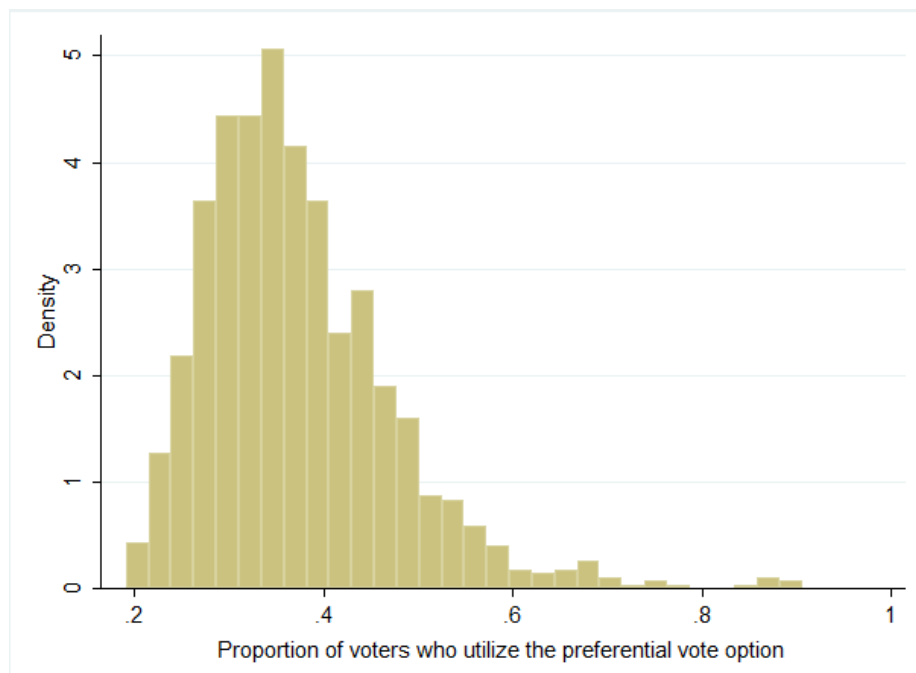
¹¹ Surely, this phenomenon can at least in part be attributed to the tendency of voters to choose the top-ranked persons by "default". This has been shown to be a basic aspect of voter psychology, and exists also in countries where ballot ranks are decided by alphabetical order or at random (e.g. Montabes and Ortega, 2002).

Clearly, the figure shows no tendency for voters to cast preferential votes for low-ranked candidates, whose list ranks do not guarantee a council seat.

Since voters cast their preferential votes for politicians with top positions on the party list, most of those who clear the preferential voting threshold belong to the top group. Out of all nominated, one fifth clear the 5-percent threshold, and of this fifth, only about 5 percent are placed low enough not to be guaranteed a seat by their position on the list.

Let us briefly examine the variation in preferential voting across municipalities and election years. Figure 3 shows the municipality-election year distribution between 1998 and 2010. The four highest participation rates all come from the small northern municipality of Övertorneå. In this municipality, the Social Democrats in fact announced publicly before the 2010 election that whoever among their top three ranked candidates that received the most preferential votes would also become the council-board chair.

Figure 3 – Proportion of Preferential Voting across Municipalities and Time



Notes: Share of voters, by municipality and election year, who used the preference vote option in the elections from 1998 to 2010.

More broadly, the high-participation municipalities are all rural municipalities with below-average population. The lowest participation rate, of 19 percent, refers to the 2006 election in Nacka, a suburban municipality just east of Stockholm. The vast majority of other low-participation municipalities are either larger cities or suburbs to Sweden's three largest cities.

Unconditionally, voters for small parties on the center-right are more likely to participate than voters for the largest parties: the Social Democrats (on the left) and the Conservative party (on the right). Controlling for party, however, more voters cast a preferential vote the larger the party group in the municipal council. This correlation is interesting, because the absolute number of votes needed to clear the threshold *increases* with both the size of the party and the size of the municipality. As

participation rises rather than falls with the relative size of the party, a lower likelihood that their vote is pivotal does not appear to deter voters from casting a preferential vote.

5. Statistical Methodology

To demonstrate our main finding that preferential voting in the general election works much like an indirect primary for leadership of the party, we consider different specifications for estimating the effect of preferential-vote winning on leadership selection.

OLS

First, we show the effects from a simple OLS regression. This specification quantifies the correlation between receiving the most preferential votes and the likelihood that a politician becomes the leader in the next election period, but does not allow us to estimate the causal effect of receiving one of the two treatments. We let index i denote the individual, index p the party, and index t the election year. The dependent variable, $Y_{i,p,t+1}$ is set to one if politician i is appointed leader of local party p in the next election ($t+1$) (i.e., if s/he is the top ranked person on the party's ballot) and zero otherwise. The treatment indicator $P_{i,p,t}$ is a dummy for individual i being the preferential vote winner in party p and election t . The basic specification assumes the following simple relation between $Y_{i,p,t+1}$ and $P_{i,p,t}$,

$$Y_{i,p,t+1} = \beta_0 + \beta_1 P_{i,p,t} + \varepsilon_{i,p,t+1} , \quad (1)$$

where β_1 measures the correlation between the treatment and becoming leader in the next period.

RDD

To estimate a causal effect, we use a Regression Discontinuity Design (RDD), in three specifications: (i) one with a flexible control polynomial, (ii) one with a linear control function, and the sample restricted to observations close to the threshold, (iii) one where the sample is restricted to observations close to the threshold, but with no controls included. The RDD regressions add a so-called forcing variable, $m_{i,p,t}$, to Equation (1), a variable that measures the distance of politician i from the threshold for being the list winner in party p , $P_{i,p,t}$. Approaches (i) to (iii) strike different balances between the size of the window around the threshold and the account for the relationship between the outcome and the forcing variable.

The first approach uses a wide sample and a control function, a third-order or fourth-order polynomial function of the forcing variable. We still drop some observations in the tails of the forcing variable to limit the risk of over-fitting the control polynomials to capture outliers. For each type of polynomial we set the data window in accordance with the density of observations in a common-sense manner (see Figure A1, of the Appendix for the distribution of observations). We set the maximum distance from the threshold as a win margin by 50 percentage points of the combined preferential votes obtained by the winner and the runner up.

The basic idea behind the flexible control polynomials is that the treatment variable, $P_{i,p,t}$, is entirely determined by the forcing variable, $m_{i,p,t}$. Because of this, we can control for potential endogeneity of the treatment variable, and also deal with other problems, such as omitted-variable bias, by controlling flexibly for the forcing variable. To do that, we estimate separate control functions on each side of the threshold. The specification becomes:

$$Y_{i,p,t+1} = \beta_0 + \beta_1 P_{i,p,t} + f^L(m_{i,p,t}) + f^U(P_{i,p,t} m_{i,p,t}) + \varepsilon_{i,p,t+1} , \quad (2)$$

where f^L and f^U denote the control functions below and above the threshold.

The second RDD approach uses a more narrow estimation window and a local linear-control function, which allows us to account for the possibility of a relationship between the forcing variable and the outcome. We show the results for two estimation windows, the optimal bandwidth suggested by the Imbens and Kalyanaraman (2012) test, and a 10 percentage point window around the threshold. Graphically, we also show the estimate for every estimation window between 0 and 30 percentage points by gradually increasing the window by 0.5 percentage point increments and plotting each treatment effect and its 95% confidence interval.

The third approach is to use a close margin around the threshold and exclude the control functions for the forcing variable. In tables, we show the results for a 10 percentage point window and a 5 percentage point window. We also show these results graphically by plotting the estimate for all window sizes between 0 and 30 percentage points in a figure.

The results for the three RDD approaches are shown both with and without control variables. These include list rank interacted with four dummies for the party group's size in the municipal council, a dummy for being a woman, seven dummies for educational categories, a dummy for a candidate elected to the parliament in $t+1$, and four dummies for age categories. We also include interaction terms between a dummy for an individual being top-ranked on the list and three dummy variables for election outcomes: (i) party p becomes a part of the governing majority between t and $t+1$, (ii) party p is removed from the governing majority between t and $t+1$, or (iii) party p loses voter support between t and $t+1$. If the relationship between the treatment and the selection of party leader in the next election still hold up when these controls are interacted, this is further evidence that our estimated relationship is causal.

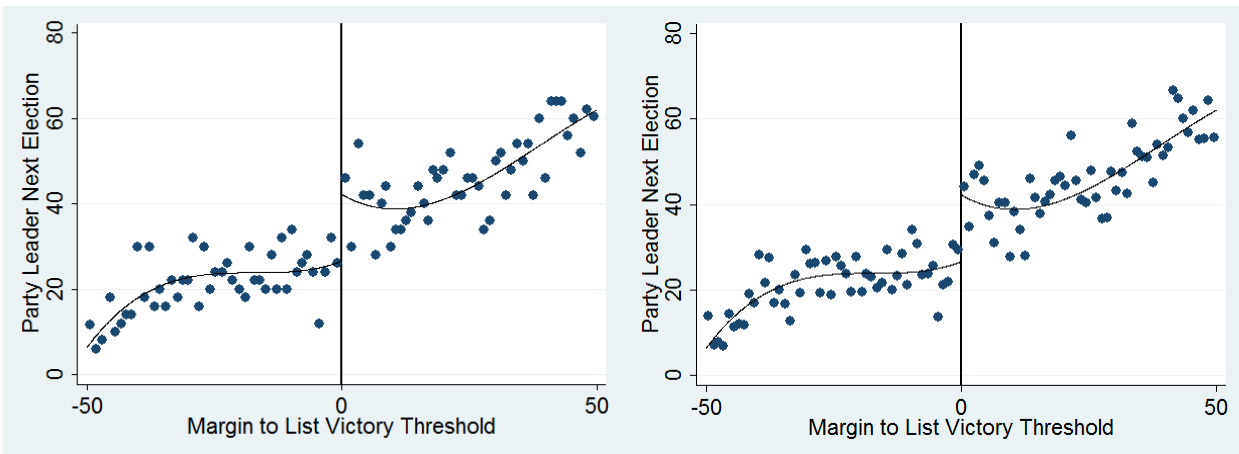
Our starting sample includes 6,862 individual-election observations from competitive preferential vote elections, defined as the absolute win margin for the politician with the most preferential votes being less than 50 percentage points. From this sample we omit 1,266 individuals in party groups with less than three representatives, since the notion of a political leader is not as relevant in these small groups. We also omit any observation where more than two candidates are close to thresholds, as the probability of winning is less than 50% in these cases. Specifically, we exclude all observations where the *second* runner up is within a 10 percentage point margin from winning the most votes, which applies to 118 observations. Finally, we exclude 992 party groups where at least one of the two top contenders in the preferential vote was not guaranteed to be elected based on their list rank, i.e., their list rank was lower than the number of candidates elected. This leaves us with a final sample of 3,640 individual-election observations.

6. Baseline Empirical Results: Party Behavior

Graphical Analysis

Following the "industry standard" for RDD research, we begin with a graphical analysis. Figure 4 shows the binned averages of the percentages of candidates that are selected as party leaders in the next election period, $Y_{i,p,t+1}$ as a function of the win/loss margin to being the party's list winner in preferential votes, $m_{i,p,t}$. As noted above, we set the range of $m_{i,p,t}$ to [-50% to 50%]. The binned averages are computed in two different ways. In the left graph, each dot corresponds to 50 observations, while it corresponds to a 1 percentage-point interval for the forcing variable $m_{i,p,t}$ in the right graph.

Figure 4 – Graphical Analysis of the List Winner Effect



Notes: The graphs show binned averages of politicians chances of becoming the party leader in $t+1$ as a function of the win/loss margin to being the list winner in the party's preferential votes. In the graph to the left, each bin contains 50 observations. In the graph to the right, each bin contains all observations within a 1 percentage-point interval for the win/loss margin. The lines in the graphs are third-order polynomials fitted to the observations.

When we cross the list-winner threshold in the party, both graphs show a clear shift in the probability of being selected as party leader. This probability is 25 percentage points among narrow losers, but 40 percentage points among narrow winners. In other words, the probability of becoming the party leader increases by 15 percentage points, or by more than half (60 percent in relative terms), if the candidate narrowly wins the most preferential votes in the party group.

Regression Analysis

Table 1 shows the regression estimates corresponding to Figure 4. For comparison, column 1 shows the results from estimating Equation (1) by OLS. Columns 2-7 then moves to estimating Equation (2), using the three RDDs outlined above. Columns 2 and 3 contain results from the control-polynomial specifications, column 4 and 5 the local-linear control-function specifications, specification, and columns 6 and 7 the close-election specifications. The upper pane of the table does not include control variables, while the lower pane does. Each cell displays the estimated coefficient on the dummy variable for the forcing variable being above the threshold, i.e., θ_1 in Equation (1) or (2).

The OLS results show that, overall, the politician who receives the largest number of preferential votes in his or her party group is about 19 percentage points more likely to be selected as the party leader compared to those politicians who do not. However, including our extensive set of control variables cuts this estimate by more than half.

The estimates from the RDD specifications show that the correlation captured by the OLS estimates is indeed driven by a causal effect. The estimates are similar to what we would expect from the graphical analysis. The two polynomial specifications both have large estimates and are statistically significant at the 5 and 10 percent significance levels respectively. The I-K test suggests an optimal estimation window of 21 percentage points and a point estimate of 12 percentage

Table 1 – Estimates of the Relationship between Being the List Winner in Preferential Votes in Election t , and Becoming Party Leader in $t+1$

		Flexible Polynomials		Local Linear		Close Margin	
		3 rd Pol	4 th Pol	I-K test	10% window	10 % window	5% Window
<i>Without Controls</i>	OLS						
Treatment Effect	24.79*** (1.64)	15.59** (7.06)	16.40* (9.04)	12.07** ^a (5.38)	21.18*** (8.00)	13.51*** (3.71)	20.20*** (5.48)
Observations	4,486	4,486	4,486	1,870	888	888	406
<i>With Controls</i>							
Treatment Effect	12.09*** (1.78)	14.10** (6.87)	17.32* (8.90)	11.31** ^b (5.11)	19.39** (7.82)	11.73*** (3.63)	17.87*** (5.41)
Observations	4,437	4,437	4,437	1,892	878	878	406

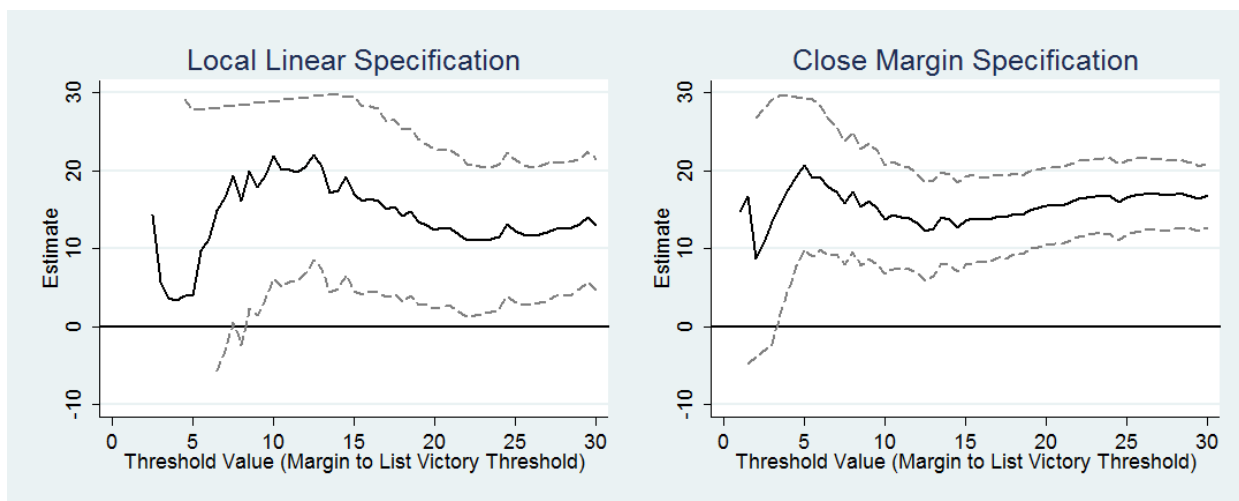
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors clustered at the party-group level are reported in parentheses. Dummy variables are scaled as 0 or 100. Control variables include three dummy variables for election outcomes: the party becomes a part of the governing majority between t and $t+1$, the party is removed from the governing majority between t and $t+1$, or the party loses voter support between t and $t+1$. They also include the following dummy variables, included separately for the leader and the challenger: list-ranked fixed effect (for the challenger), four dummies for the size of the party group in the municipal council, a dummy for being a woman, seven dummies for educational categories, a dummy for a candidate being elected to the parliament in $t+1$, four dummies for age categories where one specifically for persons reaching retirement age before election $t+1$. Bandwidth: ^a20.3% ^b21.0%.

points, which is also statistically significant at the five percent level. The local linear specification with the 10% estimation window gives a larger point estimate, which likely reflects that the control function is fitted to outliers closer to the threshold.

When we include the large set of controls in the lower pane of the table, the estimates barely change for our RDD approaches compared to the naive OLS. For the specifications that use a wide estimation window, we also see a reduction in the standard errors. That the results are unaffected by a large set of variables, which are strongly related to a candidate's share of preferential votes as well as her selection as local party leader, strongly support the claim that our RDD estimates are indeed causal.

Figure 5 shows estimates for the local linear and close-elections specifications over numerous sizes of the estimation window. The solid line is the estimated treatment effect corresponding to each window size, and the dotted lines represent a 95% confidence interval. The confidence intervals are censored and only shown in the interval $[-10, 30]$. As the figure shows, it is hard to obtain precise estimates below a certain range for the window around the threshold and thus the size of the estimation samples. However, at reasonable sample sizes, the estimates are remarkably stable and significant at the 95% level for all bandwidths. The results clearly show that the list winner in the party group becomes substantially more likely to be selected as party leader in the next election.

Figure 5 – Estimates of the Treatment Effect by Estimation Bandwidth



Notes: The solid lines in the figure show the estimated treatment effects from being the list winner, as we vary the width of the estimation window around the threshold, with a local linear specification and a close-margin specification, respectively. The dotted lines are 95% confidence intervals.

Robustness Checks

To validate our identifying assumption that treatment status is as good as random around the threshold, we conduct a large set of robustness checks. We now summarize these tests, leaving a more detailed explanation with tables and figures for the Appendix. Our first sensitivity check is the McCrary test (McCrary, 2008) for balance in the density of observations around the threshold to verify that the first-ranked candidate on the list is not more likely to become the list winner close to the threshold. Restricting the sample to only the party leaders in period t , the McCrary test shows no difference in the number of times that these current leaders win or lose close contests.

A second sensitivity analysis checks that variables that should not be systematically affected by the treatment are indeed balanced on the two sides of the threshold. To do that we perform graphical analyses as well as placebo regressions on four relevant pre-determined covariates: being the current leader, the individual's current list rank, gender, and years of education. This analysis strongly suggests that covariates are balanced around the threshold. There is no significant difference in the chosen characteristics for individuals that narrowly win or lose close contests.

In a third set of tests, we vary the placement of the treatment to create so-called "placebo thresholds" at false values of the forcing variable. In this analysis, we only find a significant treatment effect at the true threshold. Taken together, all the robustness checks provide very strong support for our identifying assumption.

Validation of the Party Leader Measure

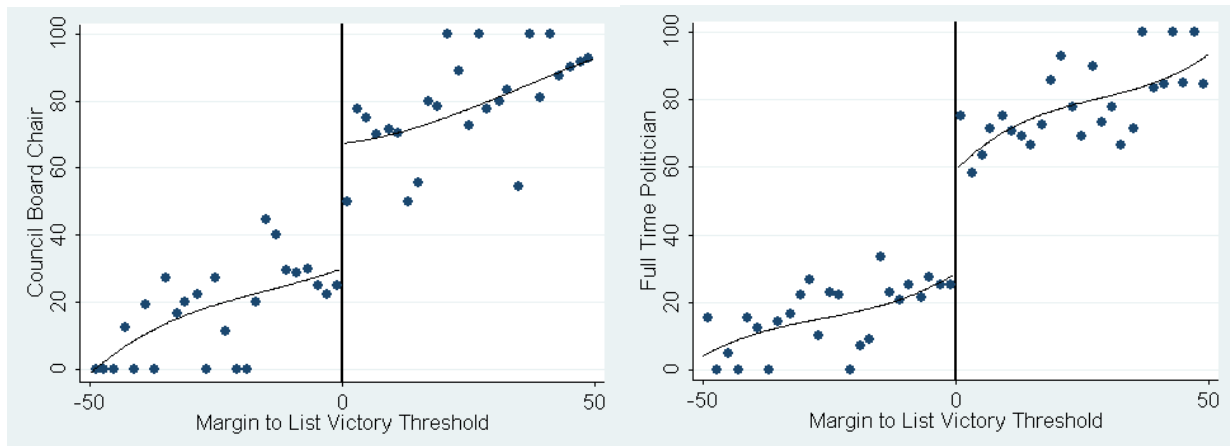
To verify that preferential voting causes a reallocation of political power rather than a mere "reshuffling" of the list, we use our data on specific political appointments. For the largest party of the political majority in each municipal assembly, we can check which politician was appointed chair of the municipal council board, an appointment made immediately after the general election. We also use data for full-time political employment, a strong indicator of influence as

the full-time politicians usually include only the chair and the vice chair of the council board. The sample therefore includes many of the largest minority parties, which appoint the vice chair. To interpret the estimates for full-time employment in a similar way as those for the chair, we restrict the sample to those cases where the party gets to select only one full-time politician.

For this estimation, we only have only for two elections. Combined with the restriction that we only include parties which get to nominate the council board chair, or one full time politician, restricts the sample to 496, or 640, observations¹². This makes our RDD estimates less precise and more sensitive to the precise specification.

To account for the smaller sample size, we increase the bin size to 2 percentage points of the forcing variable in the graphical analysis shown in Figure 6. The figure shows a striking shift in the probability of becoming the chair of the council board as well as becoming a full time politician as we cross the threshold for winning the most votes. For chairperson, around 30 percent of the close losers receive this appointment after the election, compared to around 70 percent of the close winners. For full-time politicians, around 30 percent of the close losers receive this appointment versus 60 percent of the close winners.

Figure 6 – Graphical Analysis of the List Winner Effect on Municipal Council Board Chairpersonships and Full-Time Politicians.



Notes: The graphs show binned averages of the chance to become the council board chair, or a full time politician, after election t as a function of the win/loss margin to being the list winner in the party's preferential votes in t . Each bin contains all observations within a 2 percentage-point interval for the win/loss margin. The lines in the graphs are third-order polynomials fitted to the observations.

The regression analysis in Table 2 corroborates the graphical evidence, by showing significant treatment effects in most RDD specifications for both outcome variables. Due to the small samples, we cannot include the full set of controls, so instead we include a single dummy variable for being the current list leader, the strongest predictor of becoming the council chair as well as the preferential votes received by a candidate. Still, the exact estimates should be interpreted with caution as the estimates are sensitive to both the specification and the control for being the current leader. Despite the uncertainty in pinpointing the exact magnitude of the effect, the results validate our main finding that political parties respond to the distribution of

¹² 388 observations are in both samples, which implies that about 80% of the observations in the council-board chair sample are also in the full-time politician sample, while 60% of the observations in the full-time politician sample are also in the council-board chair sample.

preferential votes in their appointments to the most powerful positions in the political hierarchy, which are made directly following the election.

Table 2 – Estimates of the Relationship between Being the List Winner in Preferential Votes in Election t and Becoming Appointed Chair of the Municipal Council at t

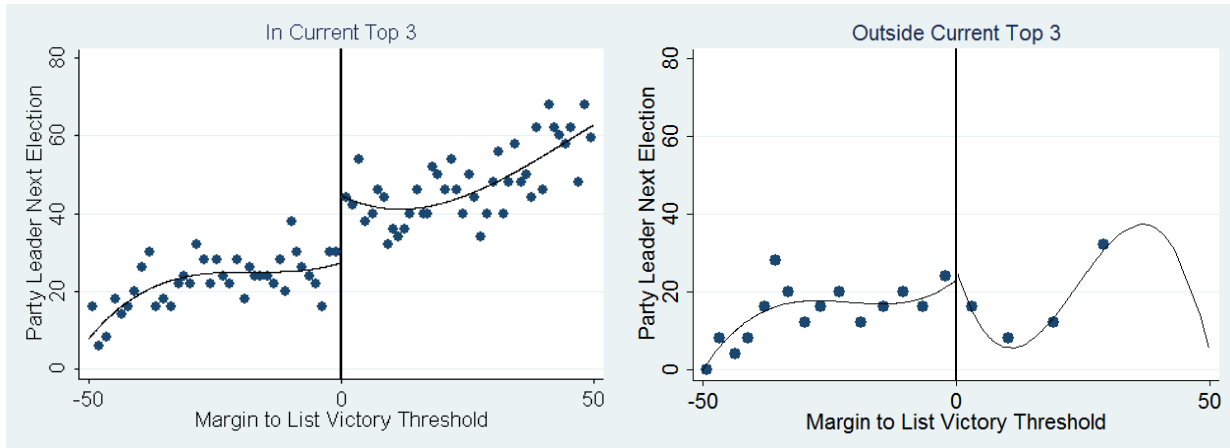
		Flexible Polynomials		Local Linear		Close Margin	
Pane A: Chair of the Municipal Council							
<i>Without Control</i>	OLS	3 rd Pol	4 th Pol	I-K test	10% window	10 % window	5% Window
Treatment Effect	62.90*** (4.78)	37.29 (25.01)	63.70* (33.25)	26.87 ^a (17.61)	51.39 (31.13)	44.74*** (14.56)	50.00** (20.51)
Observations	496	496	496	228	76	76	36
<i>With Control</i>							
Treatment Effect	23.46*** (6.31)	23.44* (13.46)	39.78** (18.25)	10.33 ^a (9.28)	33.30* (17.11)	4.64 (8.14)	18.21 (12.23)
Observations	496	496	496	280	76	76	36
Pane B: Full Time Politician							
<i>Without Control</i>							
Treatment Effect	63.75*** (4.11)	30.52 (23.32)	32.75 (32.39)	43.72** ^c (20.40)	34.00 (29.70)	42.86*** (12.44)	50.00** (18.74)
Observations	640	640	640	206	98	98	40
<i>With Control</i>							
Treatment Effect	23.71*** (5.76)	21.67 (14.46)	18.98 (17.97)	18.27* ^d (10.81)	10.78 (15.49)	19.03* (10.92)	27.20 (16.84)
Observations	640	640	640	258	98	98	40

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors clustered at the party-group level are reported in parentheses. Dummy variables are scaled as 0 or 100. The control variable is a dummy for being the party leader, i.e. top ranked on the party list, in the current election. Bandwidth: ^a24.8%; ^b21.4%. ^c17.3%; ^d20.6%.

Treatment Effect Conditional on Party Support

Next, we differentiate our estimates for party leadership to check for limitations in voter influence. If a candidate with a relatively low list rank wins the most preferential votes, this could arguably be considered a stronger signal of popularity since these candidates are relatively unknown to voters. On the other hand, candidates with lower list ranks have not made it to the top of the party hierarchy and are thus more likely to differ from the party's own leader preference. We examine these contending claims by dividing the sample in two groups, politicians that hold one of the top three positions on the party list, and those who do not.

Figure 7 – Graphical Analysis of the List Winner Effect on Party Leadership in the Next Election



Notes: The graphs show binned averages of politicians chances of becoming party leader in $t+1$ as a function of the win/loss margin to being the list winner in the party's preferential votes. In the graph to the left, each bin contains 50 observations, and in the graph on the right each bin contains 25 observations. The lines are third-order polynomials fitted to the observations.

Figure 7 shows the graphical results for our RDD in these two samples. The baseline finding in Figure 5 is replicated in the left graph, based on the sample of top-three politicians. Here, winning the most preferential votes dramatically raises the likelihood of becoming party leader. The right graph shows that candidates outside of the top three do not experience the same career boost. Although the sample is smaller, there is no indication that close winners have a larger average likelihood of becoming party leader than close losers.

Table 3 – Estimates of the Relationship between Being the List Winner in election t , and Becoming Party Leader in $t+1$, by Party List Rank

	OLS	Flexible Polynomials		Local Linear		Close Margin	
		3 rd Pol	4 th Pol	I-K test	10% window	10 % window	5% Window
Party List Rank 1-3							
Treatment Effect	25.04*** (1.72)	17.24** (7.45)	18.98** (9.56)	13.90** ^a (5.81)	24.78*** (8.47)	15.30*** (3.94)	23.15*** (5.76)
Observations	4,018	4,018	4,018	1,622	790	790	359
Party List Rank 4 and up							
Treatment Effect	1.79 (4.14)	3.28 (15.48)	0.51 (19.84)	-5.85 ^a (10.65)	3.16 (16.85)	-7.81 (7.17)	-2.22 (12.51)
Observations	468	468	468	224	98	98	47

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors clustered at the party group level reported in parentheses. The table shows OLS and RDD estimates of the heterogeneous relationship between the treatment of being the list winner in preferential votes in election period t and the current list rank. Bandwidth: ^a19.4%; ^b24.3%.

Table 3 contains the results from applying the RDD to the subsamples based on list rank. The regressions tell a clear story and validate the inference from the graphs. The treatment effect of receiving the most preferential vote exists only for those politicians that are already part of the party elite. We do not have a large enough sample to break up the analysis by individual list rank and get

precise estimates. However, even though the estimates are not always statistically significant the treatment effect is not significantly different for the first, second and third ranked candidates.

All in all, the analysis strongly suggests that a candidate must reach a high enough position in the party hierarchy to benefit from winning the most votes in the party group. This suggests that voters are only able to influence the leadership of the party group within a pre-determined pool of candidates. It is as if candidates first have to pass a vetting process before they are eligible to participate in the primary. In this sense, voter support is a complement rather than a substitute for party support.

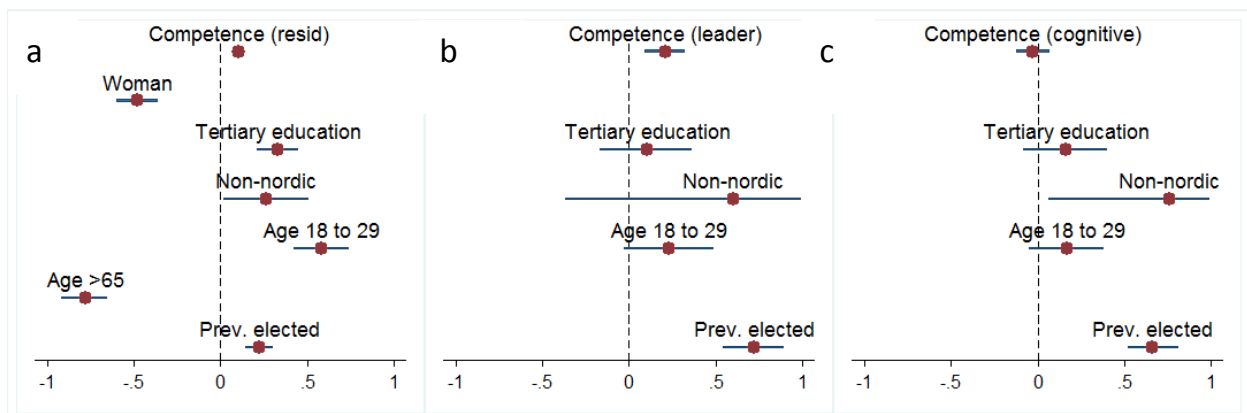
7. Extension 1: Candidate Competence

In this section, we examine if the voters' influence over the selection of party leaders translates into the selection of more competent leaders. This would require two conditions to hold. Voters must vote for highly competent candidates, and parties must promote these high-competence candidates.

To check if voters demand more competent politicians we run a simple regression for the relationship between on the one hand candidate competence, as well as key socioeconomic characteristics of the politician, and on the other hand his or her preferential votes. The socioeconomic characteristics are gender, country of birth (outside Nordic countries or not), age, re-election (at least once), and education (tertiary education or not). The last two of these have been used to measure politician quality in previous work, but we also include three other measures.

Our main income-residual measure is the one in Besley et al. (2013), who exploit data on Swedish local politicians' earnings on the labor market. It assumes that competence can be gauged from average income differences (over 20 years) between people with the same education, occupation, age, and gender. The other two competence measures are from Sweden's prior mandatory military draft: one cognitive score from a written IQ-type test and one leadership score from a formalized interview with a psychologist meant to sort recruits across training programs. However, the draft data are only available for men born between 1951 and 1980.

Figure 8 – Share of Preferential Votes and Politician Characteristics



Notes: Each graph shows the coefficients and their 95% confidence bands in an OLS regression of the share of preferential votes for an individual candidate on her list rank and a number of personal characteristics, as described in the text.

The results from this analysis are presented in Figure 8, which plots the point estimates and a 95% confidence interval for each candidate characteristic. To reduce the risk of conflating socioeconomic characteristics with trends in preferential voting and/or list rank, all specifications include non-parametric trends in preferential voting over time (year-fixed effects); non-parametric trends within each specific party group (year-fixed effects interacted with party-group-fixed effects)¹³; and preferential vote differences across list ranks (list-rank fixed effects interacted with four categorical dummy variables for the size of the party group). Since they include list-rank fixed effects, our estimates are likely to be downward biased as more competent politicians more likely have higher ranks. The control for list rank may thus partly absorb the correlation between competence and preferential votes.

The income-residual and leadership measures are positively correlated with preferential votes, even though we control for list rank. Higher wage-residual competence by one standard deviation is associated with a 0.1 percentage point higher share of preferential votes. The estimate based on the leadership score has the same magnitude. Even though the estimates are relatively small, they are still striking as they refer to a characteristic that voters cannot learn about from ballot information. One potential reason is that charismatic politicians with good leadership skills are more likely to run personal campaigns or be known through the local media. Some Swedish local politicians do run campaigns for preferential votes, but these are often modest with spending typically falling below 5,000 SEK (~ USD 800) in the early 2000s.

For the competence measure based on cognitive skills, we do not find a statistically significant association. This may be due to the limited sample of male politicians for whom we know the military draft results. In the full sample (figure 8a), the coefficient on tertiary education is significant, but it is not in the smaller sample, suggesting that the analysis may pick up a different sample rather than voters disregard for cognitive skills. It is also notable that preferential votes are positively and strongly associated with previous election, a common measure of politician quality.

Given that voters indeed appear to reward more competent candidates with more preferential votes, we move on to the behavior of parties. Here, we split the baseline sample by the median of our primary competence measure, the income residuals. As previously, we start with a graphical analysis shown in Figure 9. When comparing the two groups we see that the treatment effect of list winning is much more pronounced for those with competence above the median.

The regression analysis in Table 4 sheds further light on how the treatment effect might differ across the two samples. The estimated effect(s) is significantly different from zero in only one specification for politicians with below-median competence, while the coefficients are larger and statistically significant in all but two specifications for those with above-median competence. This suggests not only that the parties accept the competent candidates as leaders when they win the most votes, but that these candidates are even *more* likely to be promoted. Competence may thus serve as a complement to voter demand rather than as a substitute.

¹³ When we use the draft data to measure competence, we often only have a measure one member, or none, of each party group. To avoid throwing these observations away we do not include the party-group-year fixed effect directly in the regressions. Instead we take the residual from a regression for the full sample in which we include the fixed effect. The estimates are not sensitive to using this procedure or not, but the standard errors of the estimates become much smaller.

Figure 9 – Graphical Analysis of Heterogeneous Treatment Effects by Individual Competence



Notes: The graphs show the effect of passing the threshold of list winning in the local party group on being selected as the party leader in the next election period. Each bin contains 100 observations, and the lines represent fitted third order polynomials. The margin to list victory is computed as the vote difference in percentage points between the winner and the first challenger. The sample is divided by median competence.

Table 4 – Estimates of Heterogeneous Treatment Effects by Individual Competence

		Flexible Polynomials		Local Linear		Close Margin	
<i>Below Median</i>	OLS	3 rd Pol	4 th Pol	I-K test	10% window	10 % window	5% Window
Treatment Effect	23.16***	8.21	4.49	5.70 ^a	17.47	7.95	14.50*
	(2.21)	(9.49)	(12.11)	(7.20)	(10.64)	(4.95)	(7.44)
Observations	2,097	2,097	2,097	864	417	417	178
<i>Above median</i>							
Treatment Effect	24.63***	19.79**	22.13*	16.59** ^b	20.96*	18.60***	22.81***
	(2.24)	(9.43)	(12.01)	(7.10)	(10.70)	(5.16)	(7.24)
Observations	2,097	2,097	2,097	931	410	410	202

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered at the party group level reported in parentheses. The table shows OLS and RDD estimates of the heterogeneous relationship between the treatment of being the list winner in preferential votes in election period t and party p being selected as party leader in election $t+1$, by individual competence measures being above and below median competence as measured by average income residuals. Bandwidth: ^a20.64%; ^b21.8%.

8. Extension 2: Political Competition

Which party groups drive our finding that parties, on average, respond forcefully to preferential votes when distributing power? It is particularly important to check how party response relates to political competition. As recently noted by Hirano and Snyder (2013), in the US context primary elections are mostly needed when competition between political parties is poor. Poor competition is not uncommon in Sweden, where many municipalities have been governed by the same party since

1976, the first election since the current municipal structure was established. In these municipalities, preferential votes might introduce an element of within-party competition even though between-party competition is next to absent.

When more than two parties compete in an election, measuring competition is not self-evident. However, Swedish politics is centered around two stable left- and right-wing blocks, and has therefore been classified as a bipartisan political system (Alesina et al., 1997). The vote-shares difference between these blocks thus becomes an appropriate measure for the main dimension of political competition (see also empirical studies by, e.g., Svaleryd and Vlachos, 2009). To capture the persistent component of competition, we use a three-election moving average of the (absolute) difference in vote shares. We divide the party groups into four different subsamples based on majority status (majority or opposition) and political competition (above or below the median).

Table 5 – Average Share of Voters Casting a Preferential Vote, by Competition and Majority-Minority Status of the Municipality

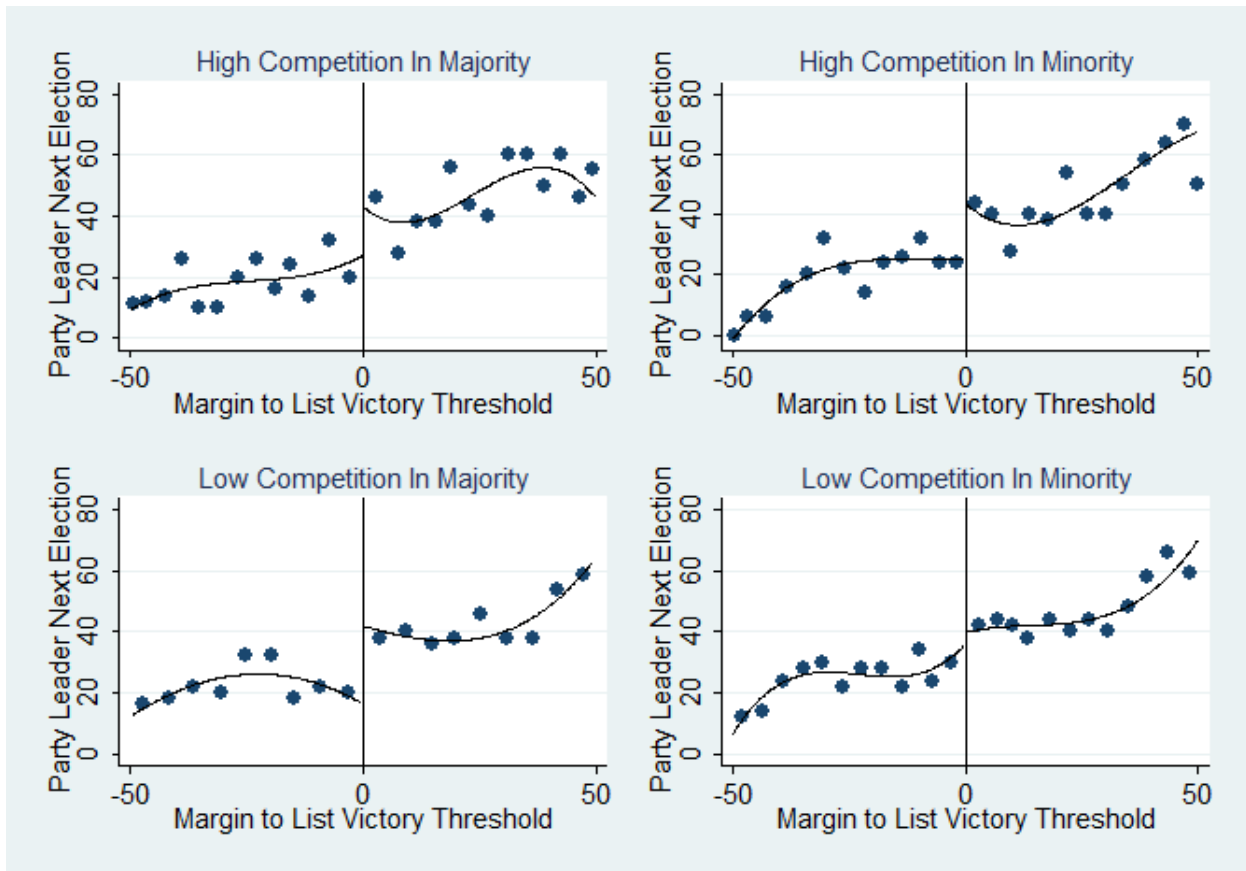
	Majority Parties	Opposition Parties
Above-median Competition	28.6%	26.5%
Below-median Competition	31.1%	25.7%

Before turning to the effects on leader selection, we briefly examine voter behavior along these dimensions. The cross-tabulation in Table 5 shows that the propensity to cast preference votes is greater in the municipalities with weaker political competition. About 31% of the voters cast a preferential vote for parties belonging to the governing majority vs. 26% for those in the minority. In municipalities with stronger political competition, the gap between majority and opposition parties is almost gone.

The RDD analysis of preferential votes and party leadership again starts with graphs. Figure 10 clearly shows that the treatment effect is at its largest in governing parties of low competition municipalities. In this subcategory, we see a large shift in the probability of becoming the local party leader as we cross the list-winning threshold for preferential votes. In municipalities with high competition, an effect – although a weaker one – appears to be present in both majority and opposition parties. Finally, opposition parties in low competition municipalities show no sign of a treatment effect. These parties have a low probability of exerting political influence, so we would expect their choice of leaders to be of little importance for voters.

The regression results, presented in Table 6, reinforce the graphical analysis. We see a large effect of winning the most preferential votes in majority parties in low competition municipalities. The estimated treatment effect varies between 20 to 30 percentage points across specifications, and is significant at the 10 percent level in half of the specifications. In the high competition subsamples, for majority and minority parties, the estimated effect is not as consistent, but always has a positive sign. Opposition parties in low-competition municipalities show no sign of a treatment effect.

Figure 10 – Graphical Analysis of Heterogeneous Treatment Effects by Political Context



Notes: The figures show the binned averages of politicians being the party leader in $t+1$ as a function of the win/loss margin in the election in t . Each bin contains 50 observations and the lines are fitted third-order polynomials. The full sample is divided into four categories depending on majority status of the party and the level of political competition in the municipality.

These empirical results support our conceptual discussion in the introduction. Specifically, they show that preferential voting can serve a similar role as did primary elections in the U.S. – at the time of their introduction – namely to introduce individual electoral accountability within dominant parties in localities where political competition is weak.

Table 6 – Estimates of Heterogeneous Treatment Effects by Political Context

		Flexible Polynomials		Local Linear		Close Margin	
	OLS	3 rd Pol	4 th Pol	I-K test	10% window	10 % window	5% Window
High Competition, Majority Party							
Treatment Effect	28.74*** (3.04)	15.76 (13.01)	29.25* (15.92)	12.98 ^a (8.32)	26.11* (14.63)	12.50* (7.46)	23.40** (10.69)
Observations	1,218	1,218	1,218	728	208	208	94
High Competition, Minority Party							
Treatment Effect	26.90*** (3.13)	18.48 (12.26)	20.97 (15.43)	13.91 ^b (9.78)	26.13* (13.93)	14.17** (7.06)	25.00** (9.49)
Observations	1,212	1,212	1,212	504	254	254	136
Low Competition, Majority Party							
Treatment Effect	19.95*** (3.69)	25.69 (16.35)	31.09 (21.06)	25.05* ^c (13.93)	26.34 (19.01)	20.24** (8.11)	21.62 (13.02)
Observations	872	872	872	276	168	168	74
Low Competition, Minority Party							
Treatment Effect	22.13*** (3.31)	3.66 (15.67)	-19.55 (20.50)	5.18 ^d (13.37)	5.07 (17.69)	9.30 (7.18)	9.80 (11.69)
Observations	1,184	1,184	1,184	408	258	258	102

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered at the party-group level are reported in parentheses. The table shows OLS and RDD estimates of the heterogeneous relationship between being the list winner in preferential votes in election period t and party p becoming the party leader in election $t+1$, by majority status of the party and political competition. Dummy variables are scaled as 0 or 100 to give results in percentage terms. Bandwidth: ^a30.2%; ^b19.5%; ^c16.3%; ^d15.8%.

9. Conclusions

We evaluate the behavior of voters and parties after the introduction of preferential voting in Sweden's PR election system. Our main new finding is a causal effect of preferential-vote support for the promotion of politicians to leading positions in the party. Although this effect was not anticipated when carrying out the 1998 electoral reform, it calls into question the negative assessments of preferential voting in the public and scholarly debate. Like elsewhere, Swedish voters scarcely use preferential votes to upset the ballot ranking by appointing candidates with low list ranks. But their preferential votes importantly influence the distribution of political power among high-ranked candidates. The indirect leadership selection in PR-systems has been criticized for poor popular accountability. Our findings suggest that preferential voting can help address this criticism. It would thus be very interesting to investigate whether preferential voting has similar effects on leadership selection also in other countries than Sweden.

Preferential voting allows voters to signal candidate popularity to politicians and parties. As we have shown, parties indeed internalize these signals in their nomination procedures. The common complaint that voters waste their preferential votes on high-ranked politicians may need to be re-considered, as the downstream effects of these supposedly wasted votes appear to be considerable. At a general level, this finding is in line with a recent branch of voting theory (Piketty, 2000; Razin,

2003; Meirowitz and Shotts, 2009), although the precise theoretical mechanism remains poorly understood.

A preferential voting system appears to place high requirements on gathering and evaluating political information. To exploit the system, a rational voter would first need to inform herself of the number of seats a party is expected to win and then evaluate the candidates above and below that threshold. If she prefers a politician below the threshold to all those above it, a preferential vote should be cast. Indeed, the most common explanation for not casting a preferential vote is a lack of information about the candidates¹⁴. However, a voter may also have motives to support a “safe” candidate, if she lacks information about other voters' intentions. Unless she votes for her true favorite, this favorite may be replaced. If party lists are ordered to reflect a political hierarchy, with popular and known political leaders on the top positions, this may be where preferential votes accumulate. This brief discussion indicates that a detailed analysis of the interplay between voters and parties in the context of preferential voting is an exciting topic for future theoretical research.

Our empirical analysis also addresses the link between preferential voting and the selection of competent politicians. Descriptive evidence shows that competent politicians receive more preferential votes, even controlling for list rank and socioeconomic characteristics. In addition to this, we find indicative evidence that competent politicians are more awarded than incompetent politicians for their preferential votes. Together, these two findings suggest that the introduction of preferential votes may help improve the selection of leading politicians.

We also show that party responses to preferential voting results are greater for politically dominant parties under weak political competition. This result mimics the *voluntary* adoption of internal primaries among U.S. local political party organizations, where strong parties adopted primaries to avoid defection and fractionalization. That we observe the same dynamic after the introduction of preferential voting implies that such reform may boost party leadership accountability as voters get an opportunity to “throw the rascals out” within the dominant party. This possibility is made more plausible as we observe that voters indeed make more use of their voluntary preferential votes when they cast their ballot for a dominant party.

Overall, our paper uncovers important and previously overlooked dimensions of preferential-voting reform. The results suggest that PR-systems can introduce strong elements of individual accountability, without losing the ability of parties to keep vote-buying politicians in check (Chang and Golden, 2007). In the Swedish system, voters have gained influence, while parties have remained the ultimate gatekeepers for final ballot nominations. However, this seemingly fortunate mix of voter and party control should be further evaluated in terms of politician behavior in office and economic policy outcomes. Do party leaders whose appointment is based on preferential votes behave differently in office? If they turn out to, say, communicate more directly with voters, or shift the policy focus from the party line to their own preferences, these auxiliary consequences should also be considered in normative judgments of preferential-voting reform. These are important topics for further empirical research.

¹⁴ At a general level, this is in line with the theory of rational abstentions in Feddersen and Pesendorfer (1996).

References

- Alesina, Alberto, Nouriel Roubini, and Gerard Cohen 1997. *Political Cycles and the Macroeconomy*, Cambridge, MA: MIT Press.
- Andeweg, Rudy 2005. The Netherlands: The Sanctity of Proportionality, in Gallagher, Michael and Paul Mitchell (eds.), *The Politics of Electoral Systems*, Oxford: Oxford University Press, 491–511.
- Ansolabehere, Stephen, Shigeo Hirano, and James Snyder 2007. What Did the Direct Primaries Do to Party Loyalty in Congress, Party and Policy Making?, in Brady, David and Matthew McCubbins (eds.) *Further New Perspectives on the History of Congress*, Stanford: Stanford University Press.
- Barro, Robert 1973. The Control of Politicians: An Economic Model, *Public Choice* 14, 19-42.
- Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne 2013. Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden, Mimeo.
- Brommesson, Henrik (2006-11-10) Consequences of the preferential voting result (*Att dra konsekvenser av personvalet*), Editorial in Kristinehamnsbladet.
<<http://www.kristianstadsbladet.se/opinion/article1094536/Att-dra-konsekvenser-av-personvalet.html>>
- Carey, John M. and Matthew Shugart 1995. Incentives to Cultivate a Personal Vote: A Rank Ordering of Electoral Formulas, *Electoral Studies* 14, 417-439.
- Caughey, Devin and Jasjeet Sekhon 2012. Elections and the Regression Discontinuity Design: Lessons from Close U.S. House Races, 1942-2008, *Political Analysis* 19, 385-408.
- Chang, Eric, and Miriam Golden 2007. Electoral Systems, District Magnitude and Corruption, *Journal of Political Science* 37, 115-137.
- Crisp, Brian, F., Santiago Olivella, Michael Malecki, and Mindy Sher 2013. Vote-Earning Strategies in Flexible List Systems: Seats at the Price of Unity, *Electoral Studies*, forthcoming
- De Winter, Lieven 2005. Belgium: Empowering Voters or Party Elites?, in Gallagher, Michael and Paul Mitchell (eds.), *The Politics of Electoral Systems*, Oxford: Oxford University Press.
- Etzler, Aron 2008. Article series on the Left Party in Fagersta Municipality, *Flamman*.
<<http://www.flamman.se/fagerstaspecial-del-1-fagerstakoden>>;
<<http://www.flamman.se/fagerstaspecial-del-2-det-kan-vara-skont-att-vara-gossen-ruda>>
- Farrell, David 2001. *Electoral Systems: A Comparative Introduction*, Palgrave.
- Feddersen, Timothy, and Wolfgang Pesendorfer 1996. The Swing Voter's Curse, *American Economic Review* 86, 408-424.
- Ferejohn, John 1986. Incumbent Performance and Electoral Control, *Public Choice* 50, 5-25.
- Hirano, Shigeo and James M. Snyder 2013. "Primary Elections and the Quality of Elected Officials" Mimeo.
- Holmberg, Sören and Henrik Oscarsson 2013, *Nya svenska väljare*, Stockholm: Norstedts Juridik.
- Hortala-Vallve, Rafael and Hannes Mueller 2012. Primaries the Unifying Force, Mimeo.
- Imbens, Guido and Karthik Kalyanaraman 2012. Optimal Bandwidth Choice for the Regression Discontinuity Estimator, *Review of Economic Studies* 79, 933-959.

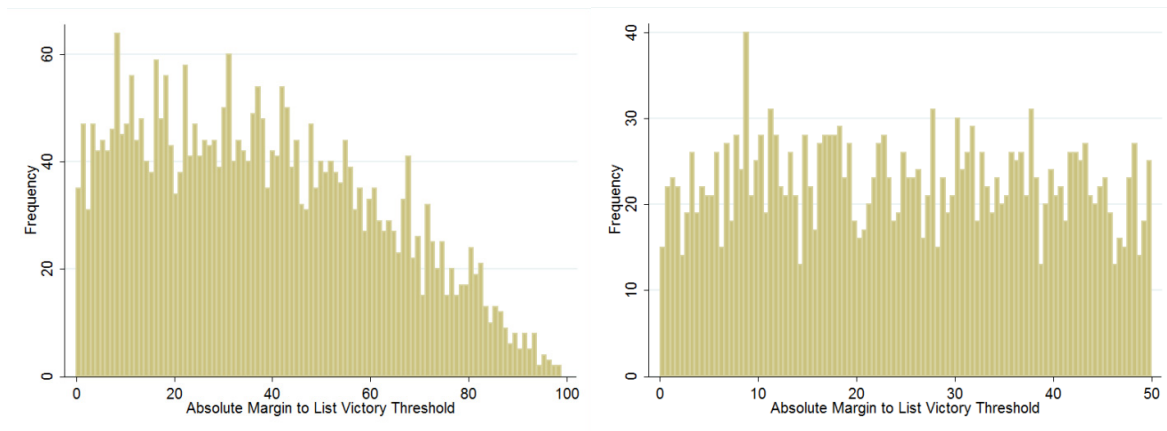
- Imbens, Guido and Thomas Lemieux 2008. Regression Discontinuity Designs: A Guide to Practice, *Journal of Econometrics* 142,615–35.
- Karlsson, David 2012. Regionala valrörelser och medborgarnas attityder i regionpolitiska frågor, in Berg, Linda and Oscarsson, Henrik (eds.) *Omstritt omval*, Göteborg: SOM-institutet.
- Katz, Richard S. and Luciano Bardi 1980. Preference Voting and Turnover in Italian Parliamentary Elections, *American Journal of Political Science* 24, 97-114.
- Kemahlioglu, Ozge, Rebecca Weitz-Shapiro, and Shigeo Hirano 2009. Why Primaries in Latin American Presidential Elections? *Journal of Politics* 71, 339-352.
- Lee, David, Enrico Moretti, and Matthew Butler 2004. Do Voters Affect or Elect Policies? Evidence from the U.S. House, *Quarterly Journal of Economics* 119, 807–59.
- Lijphart, Arend 1984. *Democracies: Patterns of Majoritarian and Consensus Government in Twenty One Countries*, New Haven: Yale University Press.
- McCrary, Justin 2008. Manipulation of the Running Variable in the Regression Discontinuity Design: a Density Test," *Journal of Econometrics*, 142, 698-714.
- Meirowitz, Adam and Kenneth Shotts 2009. Pivots Versus Signals in Elections, *Journal of Economic Theory* 144, 744-771.
- Montabes, Pereira and Carmen Ortega 2002. Position Effects and Party Nomination Strategies under the Limited Vote: The 2000 Spanish Senate Elections, *Representation* 38, 304-316.
- Mueller, Walter 2005. Austria: A Complex Electoral System with Subtle Effects, in Gallagher, Michael and Paul Mitchell (eds.), *The Politics of Electoral Systems*, Oxford: Oxford University Press.
- Pettersson-Lidbom, Per 2008. Do Parties Matter for Economic Outcomes? A Regression-Discontinuity Approach, *Journal of the European Economic Association* 6, 1037–56.
- Persson, Torsten and Guido Tabellini 2000. *Political Economics: Explaining Economics Policy*, Cambridge, MA: MIT Press.
- Piketty, Thomas 2000. Voting as Communicating, *Review of Economic Studies* 67, 169–91.
- Powell, Bingham 1984. *Contemporary Democracies: Participation, Stability and Violence*, Cambridge, MA: Harvard University Press.
- Razin, Ronny 2003. Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates, *Econometrica* 71,1083–1119.
- Serra, Giles 2006. *Primary Elections or Smoke-Filled Rooms: A Theory of Party Democratization in Latin America*. Ph.D. Thesis.
- Snyder, James, Shigeo Hirano, and Olle Folker 2012. Partisan Imbalance in Regression Discontinuity Studies Based on Electoral Thresholds, Mimeo, Columbia University.
- Svaleryd, Helena and Jonas Vlachos 2009. Political Rents in a Non-Corrupt Democracy, *Journal of Public Economics* 93, 355-372.
- SOU 1999:92. *Premiär för personval* (Premier for Preferential Voting).
- SOU 2007:68. *Ett decennium med personval* (Lessons from a Decade of Preferential Voting).
- SOU 2008:125. *En reformerad grundlag* (A Reformed Constitution).
- Taagepera, Rein and Matthew Shugart 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems*, New Haven: Yale University Press.

Appendix

A1. Data Description

In Figure A1, we show the density of observations as a function of the absolute win/loss margin. We show this for the full sample and for the sample where the relative margin between the two candidates is less than 50 percent. In both graphs we use 100 bins. We do not include negative as well as positive margins, because these would simply mirror each other (i.e. the loss margin of the runner up is identical to the win margin of the winner).

Figure A1 – The Density of Observations by Win Margin



Notes: The figures illustrate the density of observations as a function of the absolute win/loss margin. In the left-side graph we show the full distribution, while the right-side graph shows our estimation sample with an absolute margin of less than 50 percent. Each figure contains 100 bins.

The left-hand figure shows a clear concentration of observation within a win margin of 50 percent. Outside of the 50 percent margin, the density of observations goes down as we move away from the threshold. When examining our estimation sample in the right-hand graph, we see that the observations are evenly spread out within the 50 percent window.

A2. Robustness Checks for the Baseline Results

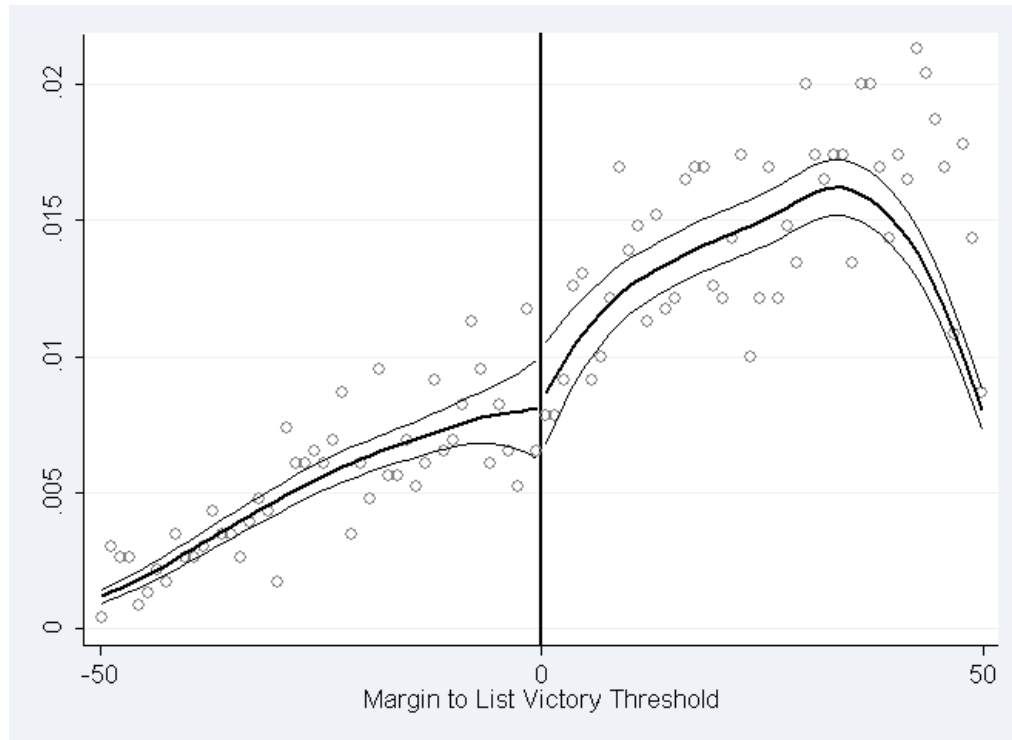
Manipulation around the Threshold

An important condition for an RDD design to yield unbiased results is that the density of observations is smooth around the threshold defining the treatment. In our case, this is not clear a priori. For example, the current leader could have more resources to draw preferential votes. S/he could also have more information about the expected distribution of the votes and use his or her resources to tilt the odds of winning. If such attempts were successful, this would show up as a higher density of observations to the right of the threshold, the cases where the current leader narrowly won the preferential vote.

To test for such manipulation, we rely on the two-step procedure proposed by McCrary (2008). In the first step, the forcing variable is partitioned into equally spaced bins and frequencies are computed within those bins. In the second step, the frequency count within each bin is used as the outcome variable in a local linear regression. Our main analysis will always have balanced frequency counts around the threshold, by definition, since we include both the winner and the runner up. To

examine possible manipulation around the threshold, we therefore restrict the sample to the current party leaders to see if they are more likely to win close races. We present the results of this test graphically, with the regression lines as well as the raw density of observations.

Figure A2 – Test for Balance around the Threshold



Notes: The graph illustrates a McCrary test for balance in the density of observations around the thresholds for being the list winner. The sample is restricted to the current party leaders. The estimated discontinuity jump is 0.0433 with a standard error of 0.1705 – thus the jump at the density is not statistically significant.

Figure A2 shows the distribution of observations, by their margin to the list-winning treatment threshold. Over the full support, the distribution is skewed to the right, which is expected given the advantage of current leaders in the preferential vote (recall Figure 2). However, neither the raw data nor the local linear regression lines suggest any systematic sorting around the threshold. Thus, there is little indication that the current party leader can affect his/her preferential vote tally relative to the top challenger, when the competition is neck to neck.

Placebo Outcomes

We now turn to some placebo analysis. This is organized in the same way as the baseline analysis, except that we consider different outcomes variables and exclude all control variables. We examine four outcomes: (i) being the current leader, (ii) current list placement, (iii) female gender, and (iv) years of education. These outcomes are strong predictors of receiving preferential votes and of being selected as party leader in the next election. The placebo outcomes are also determined prior to treatment. Therefore, they should not be affected by passing the threshold if our analysis is correct. A graphical analysis is presented in Figure A3 and regression results in Table A1.

Figure A3 – Graphical Analysis of Being the List Winner on Placebo Outcomes



Notes: The figure contains binned averages of four placebo outcomes as a function of the distance to the list-winning threshold. Each bin contains 50 observations. The lines are fitted third-order polynomials.

The results from the graphical analysis and the regressions are clear: there is no evidence for treatment status being significantly related to any of the predetermined outcomes. The graphical analysis shows no indication of a shift as we pass the threshold for any of the four outcomes; nor do the regressions provide any evidence of sorting into treatment. All estimates for the specifications that include some sort of control function are very close to zero and all but one is statistically insignificant. The few significant estimates are found for list placement and being the current leader in the “close 10 percent” and “close 5 percent” specifications. This can be explained by the strong relationship between the forcing variable (i.e. preferential votes as a share of the party vote) and list placement.¹⁵ The steep slope of this relationship leads to an imbalance in sample means between the two sides of the threshold as we expand the window but do not include any control function for this slope. Also, for years of education we find a weakly significant effect in some of the specifications with control functions, which seems to be a product of the control functions adjusting to outliers close to the threshold. In sum, we find no evidence of violations to the identifying assumptions.

¹⁵ Snyder et al. (2012) show that this type of imbalance is not a problem in RDD specifications. If one controls for the relationship between the outcome and the forcing variable, the imbalance will also be controlled for.

Table A1 – Regression Analysis of Being the List Winner on Placebo Outcomes

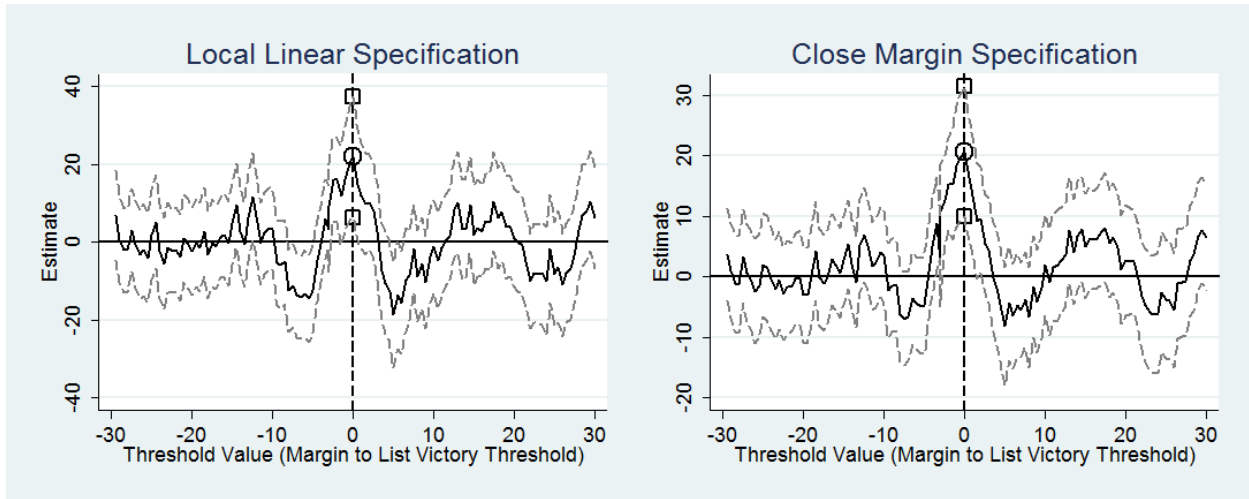
	OLS	Flexible Polynomials		Local Linear		Close Margin	
		3 rd Pol	4 th Pol	I-K test	10% window	10 % window	5% Window
Party leader current election							
Treatment Effect	45.92*** (1.78)	7.67 (8.20)	-5.17 (10.46)	3.71 ^a (7.97)	3.73 (9.39)	14.86*** (4.42)	13.79** (6.56)
Observations	4,486	4,486	4,486	1,236	888	888	406
Current list rank							
Treatment Effect	-0.94*** (0.05)	-0.35 (0.22)	-0.03 (0.30)	-0.08 ^b (0.19)	0.01 (0.26)	-0.33*** (0.11)	-0.19 (0.17)
Observations	4,425	4,425	4,425	1,451	881	881	403
Woman							
Treatment Effect	-9.32*** (1.58)	4.14 (6.64)	7.63 (8.35)	2.75 ^c (5.46)	3.17 (7.44)	-0.68 (3.57)	0.49 (5.26)
Observations	4,486	4,486	4,486	1,652	888	888	406
Years of education							
Treatment Effect	0.11 (0.08)	-0.90*** (0.34)	-0.54 (0.44)	-0.48* ^d (0.29)	-0.56 (0.39)	-0.27 (0.18)	-0.28 (0.27)
Observations	4,483	4,483	4,483	1,556	888	888	406

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered at the party-group level are reported in parentheses. Bandwidths ^a13.5%; ^b16.4%, ^c28.0%; ^d17.2%.

Placebo Thresholds

By varying the placement of the treatment threshold, we can ascertain that our analysis is not picking up one of many jumps. In Figure A4, we let the treatment threshold vary and plot the estimated treatment effects along with a 95 percent confidence interval. The figure shows the placebo analysis for two specifications with local- linear regression and an estimation window of 10% (graph to the left), and the specification without any control function and a window of 5% (graph to the right). These figures reassure us that the spikes in the estimated treatment effects are found at the true threshold values – these spikes are also the only ones that reach statistical significance.

Figure A4 – Placebo Analysis of Estimated Treatment Effects



Notes: The graphs show the estimated treatment effect on becoming party leader in the next election of being the list winner in the current election for different values of the treatment threshold. Zero corresponds to the true threshold value, while positive and negative numbers correspond to different false placebo values (for different percentages away from the true value). The black line shows the point estimates, and gray lines show a 95 percent confidence interval.