On the Returns to Holding Political Office (Is It Worth It?)*

Heléne Berg†

August 1, 2020

Abstract

The paper estimates causal effects of being elected in a local election on monetary returns. The claim for causality is made in a research design where the income of some candidate who just barely won a seat is compared to that of some other candidate who was close to winning a seat *for the same party*, but ultimately did not. This research design is made possible thanks to a comprehensive, detailed data set covering all Swedish politicians who have run for office in the period 1991–2006. The analysis establishes that monetary returns from local politics are absent both in the short and long run. By relating this null result to effects on the future political career, the paper provides suggestive evidence that local politicians are not primarily motivated by money.

Keywords: Returns to politics, incumbency effects, regression discontinuity design
JEL codes: C23, D72, J44

*I thank the editor and two reviewers for their very constructive comments. Likewise, I thank two anonymous reviewers for insightful comments and suggestions. I also thank Philippe Aghion, Matz Dahlberg, Per-Anders Edin, Olle Folke, Kaisa Kotakorpi, Eva Mörk, Torsten Persson, Erik Plug, Johanna Rickne, Michael Smart, David Strömberg, Björn Öckert, Robert Östling, participants at the 2[nd] National Conference of Swedish Economist in Uppsala September 2011, at the CESifo Area Conference on Public Sector Economics in Munich April 2012, at the 68[th] Annual Congress of the IIPF in Dresden August 2012, at the 27[th] Annual Congress of the EEA in Malaga August 2012, as well as participants at several workshops and invited seminar talks for helpful discussions, suggestions and comments. Financial support from Handelsbanken's Research Foundation (Wv12-0192) is gratefully acknowledged. Declarations of interest: none.

†Previously Heléne Lundqvist. Department of Economics, Stockholm University, SE-106 91 Stockholm, Sweden; CESifo; UCFS. helene.berg@ne.su.se

1 Introduction

For an activity to take place, the benefits must outweigh the costs. This notion is key in many political economy models. In the seminal model in Downs (1957), a politician is "some agent" whose main objective is to maximize votes and win elections in order to reap some unspecified benefits from being in office. Despite the key theoretical role, there is a lack of empirical evidence of what the returns to politics are. This paper studies just that.

To this aim, I first look at monetary returns from politics by estimating causal effects of being elected in a local election on income shortly after being elected as well as up to 15 years later. This is made possible thanks to an extensive data set covering all Swedish politicians who have run for office at any level (local, regional or national) in the period 1991–2006.[1]

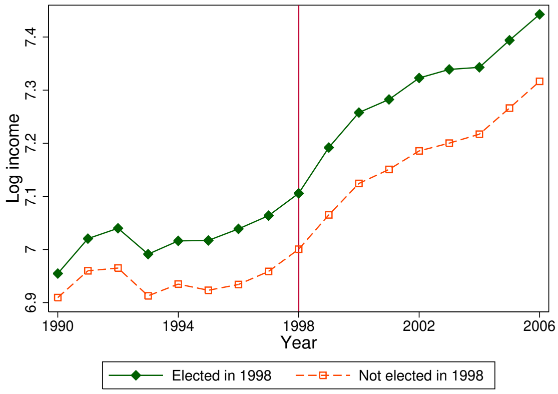
To get a first idea of what these monetary returns could be, Figure 1 displays the income profiles for all candidates who ran for a municipal council in the 1998 election, separately by whether or not they were elected. Although those elected clearly have higher income than those who were not, the gap is almost as large before the election as after. These differences can potentially be the result of selection—i.e., that elected candidates would have earned more than non-elected candidates even in the absence of being elected. It can also illustrate different political histories—i.e., that elected candidates in 1998 are more likely to have been elected also in previous elections. While it is possible to partly control for these and other confounding factors, the figure illustrates quite well the difficulty in identifying the causal effect of being elected.

Instead, the claim for causality is made in a simple yet compelling research design which, to my knowledge, this paper was the first to apply. It fits into the class of identification strategies that rely on stochastic features of close elections (e.g., Lee et al., 2004 and Folke, 2014), but differs in that identification comes from within-party discontinuities rather than between. The idea is to compare the income of some candidate who just barely won a seat to that of some other candidate who was close to winning a seat *for the same party*, but ultimately did not. Because elections result in a fixed final ranking of each party's candidates,[2] the discontinuity between these candidates—whom I refer to as the *borderline elected* and *borderline defeated*—is well-defined. Moreover, other candidates than these two can be used to detect and control for any possible direct effects of being more

[1] The majority of local politicians in Sweden hold regular jobs and, at least partly, devote their spare time to politics. This means that monetary returns from politics can stem both directly from official perquisites and remuneration as well as from a better paid private job, even in the short run.

[2] Which to a large extent corresponds to the party's own ballot paper rankings of candidates; see Section 2.1.

Figure 1: Disposable income among candidates running for a municipal council in 1998



Note: The figure plots average disposable income among candidates who were elected into a municipal council in 1998 and among candidates running for a municipal council in 1998 without getting elected. Income is measured in logs of 100 SEK deflated to 2000 year values.

Source: Statistics Sweden & The Swedish Election Authority.

highly ranked on income.

Technically, the identification strategy is a type of regression discontinuity (RD) design where the forcing variable is the difference between a candidate's rank and the rank of the borderline elected. The identifying assumption is that the direct effect of rank on the outcome is "smooth" for ranks around the borderline elected.

Applying this RD design, I show graphically and econometrically that monetary returns from local politics are absent irrespective if one considers the period right after the election or up to 15 years later. This result holds for different income measures such as disposable income, labor income and income from the largest source. It is also true on average as well as when considering heterogeneous effects across various dimensions of parties, councils and candidates.

Quantifying what the returns to politics are can provide indirect evidence of what motivates some people to engage in politics. Given that monetary returns from politics are absent, one may be inclined to conclude that this is *not* what motivates politicians. On the other hand, it is possible that the mere hope of positive returns is what motivates people to engage in politics, and that the absence of returns will be a disappointment

I argue that if *non-monetary* returns are what matters, the willingness to pursue a political career should be independent of any monetary returns. I therefore apply the same RD design as for income and estimate causal effects of being elected into a municipal council on the probability of running and being elected in subsequent local elections. Comparing the effects on income with the effects on the future political career provides indirect evidence of what motivates politicians. That being said, the paper does directly study what motivates politicians.

The analysis shows a non-negligible effect of being elected once on being reelected. A possible interpretation of this incumbency effect is that those elected want to continue in politics up to eight years (two election periods), even after learning that monetary returns are absent. This thus suggests that local politicians are not primarily motivated by money.[3]

As in most RD designs, a potential drawback is that the effect is only locally identified at the threshold. In particular, one may worry that returns to office are smaller for borderline elected candidates than for elected candidates further up the ranking. I provide evidence that this is not the case. The absence of any substantial income effects seems to pertain also for the average politician. Having said that, Swedish politics *can* probably be lucrative for *some*. Indeed according to Berg (2020), there are substantial income effects of being elected into the Swedish national parliament. But this is so rare, so that it arguably would be unrealistic to hope for and be motivated by that.

The method in the paper is applicable thanks to high-quality data. Lack of proper data is probably the main reason why, for a long time, the returns to politics were more or less a black box. However more recently, a few studies have overcome the data limitations and produced pieces of convincing empirical evidence; notably Berg (2020), Eggers and Hainmueller (2009), Diermeier et al. (2005), Fisman et al. (2014), Kotakorpi et al. (2017) and Querubin and Snyder (2013).[4] This paper adds to the still scant evidence on what types of returns that motivate politicians in two main ways. First, the RD application applied constitutes a methodological innovation. As already noted, the identification strategy is clearly related to studies that rely on discontinuities in votes share in elections where some party won with a small margin. Here, I instead rely on the discrete discontinuity in candidate ranks resulting from the fact that each party will assign only as many seats as were won in the election. This makes the paper one of few to focus on within-party discontinuities rather than between, and unique in the sense that it exploits the discrete candidate ranking.[5]

---

[3] This notion is confirmed by Swedish local politicians themselves expressing ideological goals and learning about politics and society as the main motivation. In contrast, few think that political engagement is rewarded on the labor market or that it increases their status. This is according to a study done for a Governmental report; see Swedish Ministry of Integration and Equality, 2001.

[4] Folke et al. (2017) extend the concept of political returns to the politicians' children, and Folke and Rickne (2018) to non-pecuniary returns (or costs) as captured by divorce rates.

[5] The identification strategy was first introduced in an early version of this paper (Lundqvist, 2011). Subsequently, at least one more paper has used within-party RD variation, but then with a continuous forcing variable; see Fiva and Røhr (2018). Other

Second, unlike the above mentioned papers on the returns to politics, this paper has an explicit focus on the local rather than the national political arena.[6] I argue that local politics is the relevant context when thinking about (the indirect) question of what motivates politicians. This is where the vast majority of people engaged in politics are found. Furthermore, local politics deals with issues affecting the everyday life of citizens, making its actors an important group to study. The local arena is also where those who do advance to national politics normally start off. For example, among the 349 members of the Swedish parliament in 2006, 75% had previously held a municipal council seat during at least one election period.[7]

Another merit of the paper is its high-quality data. It covers all candidates who have run for office at any level (local, regional or national) in any of the five elections held during the period 1991–2006.[8] Two crucially important features are, first, that it contains the same information on all candidates irrespective of whether or not they were elected. Second, for most of the elections, it contains sufficiently detailed information to reproduce the final ranking of candidates resulting from the election. This is what makes it possible to determine who is the borderline elected. These two features, alone, make the data unique in its kind. Furthermore, rich register-based information on characteristics such as age, sex, foreign background, educational attainment, labor market status, occupation and various income measures is matched to all these candidates using a unique person identifier. The registers are in annual form and cover the years 1990–2006 for all candidates. This makes it possible to (i) follow candidates over a long time period; (ii) verify the identifying assumption with many pre-determined candidate characteristics; and (iii) study heterogeneous treatment effects across candidate characteristics.

Aside from complementing the papers listed above on returns to politics, the paper relates to the extensive literature on political incumbency effects, reviewed in e.g. Eggers and Spirling (2017). However, only a few recent papers have studied systems with, as here, proportional representation; see Dahlgaard (2016), Fiva and Røhr (2018), Fiva and Smith (2018), Golden and Picci (2015) and Hyytinen et al. (2018). The nature and mechanisms of incumbency effects in such a setting are likely quite different than in "traditional" majoritarian systems.[9]

---

potential applications for RD designs with a discrete forcing variable are, e.g., admissions based on discrete test scores or discrete grades.

[6]Kotakorpi et al. (2017) apply their analysis to both national and local politics.

[7]By studying marginal or average local politicians, the paper complements work on political careers of local top politicians (e.g., mayors). See for example Gagliarducci and Paserman (2012), who study this from a gender perspective.

[8]Versions of the same data are used in a number of concurrent papers on a variety of important political economy questions; see Besley et al. (2017), Dal Bó et al. (2017), Folke et al. (2016, 2017) and Folke and Rickne (2018).

[9]The paper also spurs interest in the literature asking the follow-up questions how

The remainder of the paper is structured as follows: Section 2 describes the key features of local politics in Sweden and the procedure for ranking candidates within parties. Section 3 presents the regression discontinuity strategy and the identifying assumptions, as well as the data. Section 4 discusses what the treatment—being elected into a municipal council vs. being close to being elected—is likely to capture, as well as how to measure returns. The main results on monetary returns are presented in Section 5. Section 6 presents results on political careers and discusses interpretation. Preceding the final and concluding section, Section 7 discusses the external validity of the results.

2 Swedish local politics

This section provides an overview of key features of Swedish local politics and municipal elections. There are 290 municipalities in total. These are responsible for a range of public sector goods and services, including primary and secondary education, child care and care for the elderly. Each municipality is governed by a municipal council elected every fourth year (every third year before 1994) in proportional elections. Local elections are held on the same day as elections to the national parliament. Voter turnout is high from an international perspective; usually around 80%.

Around two thirds are single-constituency municipalities. Municipalities with a larger electorate have multiple constituencies. In the case of two constituencies or more, candidates are elected separately from each constituency. The municipal council decides on the total number of council seats, subject to minimum restrictions set by the Municipal Law ranging between 31 for municipalities with up to 12,000 eligible voters to 101 for the municipality of Stockholm. The median council size is 41. Seats are distributed between parties based on vote shares via the so-called "modified odd-number method". There is no formal vote threshold for a seat.[10] All seven major parties in the national parliament (eight after the 2010 election) operate and have separate organizations at the national, regional and local level.[11] In some municipalities, there are additional local parties.

payoffs from politics matter for the selection of politicians and for the resulting policies, see e.g. Caselli and Morelli (2004), Mattozzi and Merlo (2008) and Messner and Polborn (2004) for theoretical contributions, and Ferraz and Finan (2009), Gagliarducci and Nannicini (2013) and Kotakorpi and Poutvaara (2010) for empirical evidence.

[10]These and other regulations surrounding elections are mainly stipulated in the Municipal Law and the Elections Act. A new Municipal Law was implemented in 2018, but the numbers given refer to the regulations in place during the studied period.

[11]Since the founding of the Green Party in 1981, national politics has been dominated by seven parties; besides the Green, there is the Left Party, the Social Democrats, the Center Party, the Liberal Party, the Moderate Party and the Christian Democratic Party. In the 1991 election, the populist party the New Democrats made a short appearance, and in the 2010 election the right-wing extremist party the Sweden Democrats—which had so

The municipal council is the highest decision-making body in the municipality. Its tasks are regulated in the Municipal Law; it *must* appoint members and replacements for committees, the most important of which is the executive board[12] (i.e., the "government" of the municipality); it *must* decide on issues that are of first-order relevance to the municipality such as the budget, the rate of the proportional income tax, organizational forms for the executive branch, remunerations to elected representatives and local referenda; it *can* delegate decisions on issues that are of second-order relevance to the executive board and to working committees.

Hence, the power of the council as stated in the Municipal Law is quite high. However, a parliamentary report with the purpose of considering measures for improving local democracy suggested that the council's power ought to increase (Swedish Ministry of Integration and Equality, 2001). This suggestion was motivated by an increasing trend in delegations to the executive board and to the chairmanships of committees. Some viewed the council as merely being a formal decision-making institution on issues that have in practice been settled much earlier in the political process.

Part of the explanation for the more widespread delegations is the fact that the majority of local politicians have other occupations and devote their spare time to politics. Less than 3% of all elected representatives and around 8% of the politicians elected into the council receive full-time or part-time compensation.[13] According to a survey of local politicians conducted in 1999, the hours per week devoted to politics are around 18 among chairs, 8 among regular council members and 5 among council replacements (Hagevi, 2000). This system implies that time constraints can be significant obstacles. Despite of this, it is generally viewed as desirable, because it also has the benefit of sustaining close connections between politicians and voters.

Below follows a description of the process of getting elected into the municipal council, a process which forms the basis for the identification strategy of the paper.

2.1 Assignment of seats within parties

Candidates can only be elected to the municipal council via parties. Parties running for election nominate and subsequently rank candidates on ballot papers. Somewhat generalized, the procedure is as follows (Bäck and Möller, 2003):

1. All party members can nominate candidates. At this stage, special-

far only been locally successful–entered the national parliament.

[12]The executive board is appointed such that the resulting distribution of seats between parties mirrors the seat distribution in the council.

[13]This is according to Öhrvall (2004), Öhrvall and Persson (2008) and own data. At least 40 but less than 100% of full-time pay are classified as part-time, although this is a rough classification since it is not always clear what constitutes a full-time assignment.

interest politics plays a role in that youth organizations, women's organizations, unions etc. nominate their preferred candidates. Anyone who has the right to vote in the municipal election can be nominated for their municipality's council.

2. An appointed election committee ranks the nominated candidates who have agreed to run. Naturally, overall popularity plays a role in the ranking but also representativity in terms of gender, age, experience and political standpoints. Some parties hold internal trial elections to assist in the ranking.

3. The ballot paper rankings are fixed. This normally occurs around six months before the election.

A party can run with several ballot papers in a single constituency and/or with one ballot paper in several constituencies. This means that there can be several *ballot paper rankings* in a single constituency and/or one *ballot paper ranking* for several constituencies. Because the seats are assigned separately for each constituency, there is, however, always one single *final ranking* per constituency. Given the total number of seats that each party has won in the constituency, it is according to this final ranking that seats are distributed within parties.

Starting with the 1998 election, voters can mark *one* preferred candidate on the ballot paper (so-called preference voting). When determining the final ranking, the top is set based on the ranking of such preference votes. During the period studied, the threshold for being elected via preference votes was 5% of the party's votes in the constituency, though this must be at least 50 votes. For candidates who do not reach this threshold, so-called comparison numbers are calculated, which are then ranked.

How the ballot paper ranking translates into the final ranking can be a complicated matter, for example when there are multiple ballot papers per constituency or when candidates run in several constituencies. These complications only arise in a minority of cases, and the details of the procedure are described in the Appendix. For the majority of cases, the final ranking mirrors the ballot paper ranking, except that candidates who have reached the preference vote threshold are put at the top.[14] The following section describes how this final candidate ranking is used for identification of the effect of being elected into a municipal council.

[14] In the three elections since the introduction of the preference vote covered by the data, around 15–20% of the candidates reached this threshold. However, considerably fewer were elected *because of* their preference votes, as the majority of those who reached the threshold were also sufficiently highly ranked on their party's ballot paper. Thus, the difference between the ballot paper ranking and the final ranking induced by moving candidates elected via preference votes to the top is, in practice, very small.

3 Identification and data

This section describes the identification strategy and data used to estimate the effects of being elected into a local council on future income and on electoral outcomes in subsequent elections.

3.1 RD in ballot paper ranking

The task is to measure how being elected into a local council affects future outcomes in terms of income and success in subsequent elections. The fundamental identification problem that needs to be solved is that who gets elected is not random. Indeed, successful politicians are successful for a reason, and this reason may very well be correlated with their potential for future income gains and political successes.

The proposed solution is to exploit the stochastic features of elections that imply that for some candidates, it is more or less random whether or not they are elected. Specifically, I will use the variation in treatment status between candidates running for the same party, given the number of seats won by that party. The idea is to reproduce the final ranking of candidates (as laid out in Section 2.1 and the Appendix) of a party that won n seats in some constituency. I then compare the outcome (income, say) of the treated n^{th} candidate to that of the untreated $(n+1)^{th}$ candidate. I refer to these two candidates as the *borderline elected* and the *borderline defeated*, respectively.

It is possible that the final ranking is systematically related to the outcome of interest. In other words, it is possible and even likely that there is a systematic difference in the innate "quality" even between the borderline elected and the borderline defeated. Other candidates than the borderline elected and defeated can help detect such direct effects. To this aim, visual inspection of the data is particularly illustrative; the treatment effect will be seen graphically as the difference between the borderline elected and defeated that is above and beyond differences between any other two candidates.

Technically, the identification strategy is a type of regression discontinuity (RD) design where the forcing variable is the difference between a candidate's rank and the rank of the borderline elected, $rank^{\star}$. The identifying assumption is that the direct effect of rank on the outcome is "smooth" for ranks around the borderline elected. Put differently, the chosen control function must capture the direct effect of $rank^{\star}$ between the borderline defeated and the borderline elected, so that the only remaining difference is the treatment effect of being elected. Conceptually, this requires that the quality of candidates evolve smoothly around the borderline elected, which is more likely to be the case if parties cannot perfectly anticipate how many seats they will win.[15] In the current setting, I argue that this is most likely to hold in a sample as close to the borderline candidates as possible, yet where it is still possible to control for the direct effect of $rank^\star$. I return to this below.

Following Lee et al. (2004), the random variation in who gets elected induced by close elections has been exploited in numerous papers estimating "party effects". My approach is similar in spirit but differs in a few ways. Firstly, I exploit discontinuous variation within parties rather than between parties. Secondly and most importantly, the forcing variable is discrete.[16]

The discreteness of the forcing variable has implications for the identifying assumption as well as for inference. Sekhon and Titiunik (2017) broadly categorize RD designs into two types, and state the respective identifying assumptions; the continuity-based approach and the random assignment approach. My approach fits into the former, although the discreteness implies that some modifications to the identifying assumption are needed. Instead of assuming that the potential outcomes are continuous in the forcing variable close to the threshold, I need to assume that a certain functional form captures the direct effect of the forcing variable close to the threshold. Furthermore, what is "close" differs from in a continuous setting. Indeed, Lee and Card (2008) point out that non-parametric identification as formalized in Hahn et al. (2001) is not even feasible when the forcing variable is discrete. That is, one is required to assume some functional form for the direct effect of the forcing variable.

Lee and Card (2008) focus on the inference problem when the forcing variable is discrete.[17] They show that when the assumed parametric form differs from the true parametric form by some error that is identical irrespective of treatment status, the treatment effect is still identified, although the standard errors need to be inflated. Inflating the standard errors is then done by clustering at the level of the discrete values of the forcing variable. Yet another special feature of the current setting is that the forcing variable can only take a limited number of values. This implies that the solution proposed by Lee and Card (2008) is not feasible.

The maximum number of values the forcing variable $rank^\star$ can take is (somewhat simplified) the number of candidates listed on a given ballot

[15] As should be clear from Section 2.1, there is a considerable amount of internal democracy within the parties in setting the ranking. This suggests that the quality of the (borderline) defeated candidates matters even when there is little uncertainty about how many seats the party will win. A related requirement is that the *candidates* cannot perfectly anticipate how many seats their party will win and take this into account when deciding whether or not to accept a certain rank nomination. In practice, the the ballot paper ranking is ultimately decided by the party, and especially so around the borderline candidates.

[16] The first application using the RD design, Thistlethwaite and Campbell (1960), was also based on a discrete forcing variable.

[17] Deke and Dragoset (2012) further discuss this.

paper. In the data, the mean is 20 and the mode is 15. But in the preferred specification from an identification-point of view, the forcing variable only takes three values. Specifically, I assume that the direct effect of $rank^{\star}$ is linear for a limited sample consisting of the n^{th}, $(n+1)^{th}$ and $(n+2)^{th}$ ranked candidates. I refer to such a set of candidates per party and constituency as the *borderline group*.[18] Hence, the borderline group consists of the borderline elected, the borderline defeated and one additional non-elected candidate. As argued above, by limiting the estimation sample to candidates close to the borderline candidates, the identifying assumption is more likely to hold. Yet with three candidates, it is still possible to control for a linear direct effect of $rank^{\star}$.

Furthermore, the reason for one additional non-elected rather than one additional elected candidate is to have the sample as representative as possible. In the latter case, the sample needs to be restricted to parties where at least two candidates were elected via comparison numbers. Now, instead, the only restriction is that there is at least one candidate elected via comparison numbers.[19] Figure 2 illustrates the importance of this point. It shows the number of borderline groups that includes candidates with the respective ranks. The main reason why many borderline groups lack candidates with positive ranks is that they only have 1–2 seats in the council. In other words, these parties are so small so that they do not have any candidates ranked higher than the borderline elected. Analogously, some parties do not have enough candidates on their ballot papers further down the list that would be assigned a low rank. Figure 2 suggests that it is likely that the lack of candidates with high ranks that follow from being a small party is systematically related to potential outcomes. This is thus the rationale for estimating the direct effect of $rank^{\star}$ using additional non-elected candidates. Note, however, that I do conduct a robustness analysis where also additional, higher-ranked elected candidates are included.

The main regression to be estimated on the borderline groups of of candidates ranked n^{th}–$(n+2)^{th}$ is then:[20]

$$Y_{i,g,t+j} = \beta_0 + \beta_1 elected_{i,g,t} + \beta_2 rank^{\star}_{i,g,t} \ (+\mathbf{\Gamma}'\mathbf{X}_{i,g,t-1}) + \varepsilon_{i,g,t+j}, \qquad (1)$$

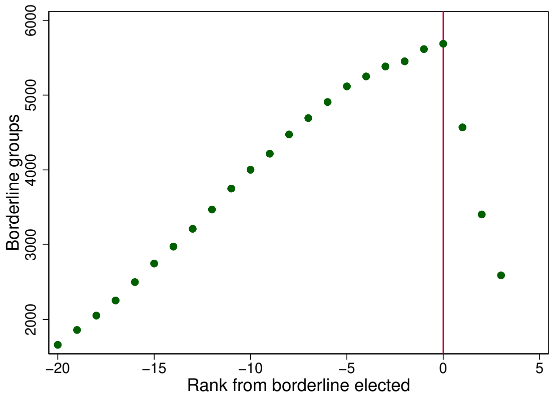
where $Y_{i,g,t+j}$ is the outcome for candidate i in borderline group g running in election year t, j periods ahead. The forcing variable $rank^{\star}_{i,g,t}$—the difference between the rank of candidate i in group g and the rank of the

[18]The majority of borderline groups are at the constituency level. But when a ballot paper overlaps several constituencies, the group is at the municipality level; see the Appendix.

[19]See the Appendix for details.

[20]For the income outcomes, the estimated model will be a log-linear. For the future election outcomes, a linear probability model will be estimated.

Figure 2: Number of borderline groups that include candidates with a certain *rank**



Source: Statistics Sweden & The Swedish Election Authority.

borderline elected in group g—is defined such that it equals 0 for the borderline elected and -1 and -2 for the candidates who would have been elected had the party gained one or two more seats, respectively. The term in parenthesis represents effects of a vector of individual characteristics measured one year prior to the election that will be controlled for in most of the estimations (although they should be redundant for identification purposes).[21] Finally, $\varepsilon_{i,g,t+j}$ is an error term that is clustered on municipality in the main specification. As noted above, the solution to the inference problem proposed by Lee and Card (2008) is not applicable in the current setting. Instead, I will investigate the robustness of the standard errors by varying the the level of clustering in other dimensions, such as party-by-county and county-by-election.

Both the graphical analysis and the estimations of equation (1) will consider short-, medium- and long-run outcomes. For income outcomes, the time index $t+j$ will be the average over 1–3, 6–8 and 13–15 years after election t.[22] For election outcomes, $t+j$ will be the first, second and fourth subsequent election.[23]

The treatment parameter of interest is β_1 and the condition for the causal

[21]Table 1 lists the control variables. To control for past political experience, a set of dummies indicating whether the candidate ran for/was elected into a municipal council in the past three elections are also included as controls. Because the earliest election in the data is 1991, these dummies are censored or partly censored (set to zero) for the 1991 and 1998 elections.

[22]Short-run income 1–3 years after the election measures immediate effects, whereas income after 13–15 years is as long as the data goes. Medium-run income after 6–8 years measures intermediate effects some time after the first election period.

[23]Four elections ahead is as far as the data allows the analysis to go. The reason for not studying the third subsequent outcome is simply to keep the number of outcomes down.

effect to be identified in equation (1) is that the direct effect of rank relative to the borderline elected is captured by β_2, meaning, once more, that it must be (at most) of order one for candidates ranked n^{th}–$(n+2)^{th}$.

More than three candidates per borderline group are required for the treatment effect to be identified if the correct functional form for direct effect of $rank^{\star}$ is of higher order than one.[24] As a complement to the main specification in (1), I allow other functional forms for the forcing variable. Specifically, the following regression on the borderline elected and several defeated candidates is estimated:

$$Y_{i,g,t+j} = \beta_0 + \beta_1 elected_{i,g,t} + \sum_{p=1}^{\bar{p}} \beta_{2p} (rank^{\star}_{i,g,t})^p + \varepsilon_{i,g,t+j}. \tag{2}$$

In equation (2) the term summing over order of polynomial p represents the direct effect of $rank^{\star}$ and $\bar{p}$ is the highest order of polynomial included in the regression. Several versions of equation (2) will be estimated by varying $\bar{p}$ between 1 and 3 and the number of defeated candidates included (i.e, the bandwidth) between 5 and 10.

With the empirical setup represented by equations (1) and (2), controlling for group-specific characteristics or a group fixed-effect (or some other more aggregate fixed-effect) is, for identification purposes, more or less redundant. To see this, note that the estimation samples consist of a nearly-balanced panel with borderline groups of candidates with the same $rank^{\star}$ values. The only exceptions are those groups where there are too few defeated candidates so that it is not possible to assign low values of $rank^{\star}$ to anyone (see Figure 2). Therefore, unless these exceptions are systematic, any group characteristics must be uncorrelated with $rank^{\star}_{i,g,t}$ and hence, also with the treatment variable $elected_{i,g,t}$ since this is simply an indicator variable $1(rank^{\star}_{i,g,t} = 0)$.[25]

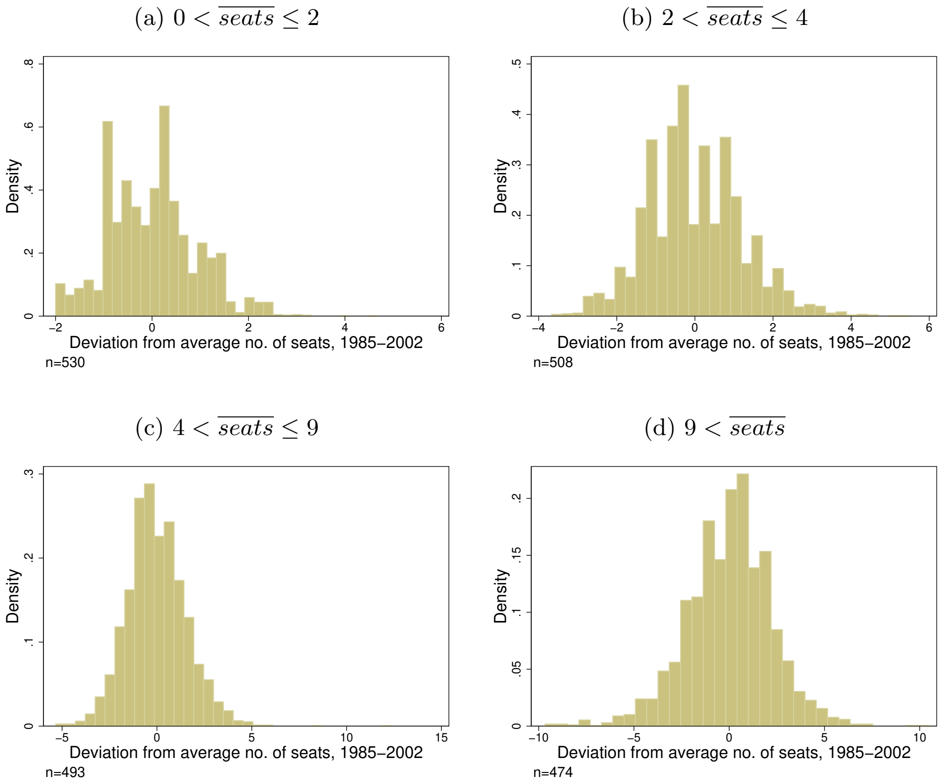
As explained above, the identifying assumption more or less relies on that parties cannot perfectly anticipate which candidates that will be elected. This, in turn, may be more likely to hold for some groups than for others. Specifically, parties that have repeatedly won n seats may anticipate that they will do so also in the next election and, consequently, may not care about the quality of the $(n+1)^{th}$ candidate. Figure 3 assesses whether this is likely to be a problem. Separately by party size, it shows the variability of seats for a given party in a given council over elections 1985–2002. Variability is measured as the deviation in the number of seats in a particular

[24]Analogously, a simple mean comparison of the borderline elected and defeated identifies the treatment effect if there is no direct effect of $rank^{\star}$.

[25]One may still want to include group fixed-effects to increase the precision of the estimates. However, it turns out that doing this neither affects the point estimates nor the standard errors (the results are available upon request).

election from the mean number of seats over the entire period.

Figure 3: Variability in parties' number of seats



Note: The figures show the distribution of the deviation in the number of seats in a particular election between 1985 and 2002 from the mean number of seats over the entire period, $\overline{seats}$. *Source:* Statistics Sweden.

Reassuringly, Figure 3 shows substantial variation even for parties that on average have two seats or less (top left plot). To further investigate the validity of the identifying assumption, the empirical analysis will contain robustness checks where I mimic a group-specific unanticipated shock that affects who the borderline elected is. Specifically, the estimation sample will be restricted to only include (i) groups whose total number of seats changed from the previous election; (ii) groups that won their n^{th} seat or lost their $(n+1)^{th}$ seat with narrow vote margins; and (iii) the combination of (i) and (ii). For this exercise, the definition and calculation of minimum changes in votes to win or lose an additional seat in proportional elections as developed by Folke (2014)[26] will be used.[27]

[26] I thank him for generously sharing his STATA code.

[27] As pointed out by a referee, another type of potential shock is when the election result turned out to differ substantially from polls close to the election. Unfortunately, polls are typically conducted only for the national election (which is held on the same day as local

Moreover, to strengthen the notion that β_1 really captures the effect of being elected, placebo regressions in which each group is assigned one or two additional seats so that the $(n+1)^{th}$ or the $(n+2)^{th}$ candidate is the "borderline elected" will be estimated. These estimations will serve as complements to the graphical analysis where such placebo effects can be directly detected.

3.2 Data

Detailed data over political candidates is a necessity for applying the research design described above. The data used in this paper covers all candidates who have run for office to a Swedish municipal council or to the national parliament in any of the five elections held during the period 1991–2006.[28] The elections to municipal councils in 1991, 1998 and 2002 define the population under study for short-run outcomes. For medium-run outcomes, the elections in 1991 and 1998 define the population. Because the data ends in 2006, only the 1991 election defines the population for long-run outcomes. The number of borderline groups is around 1800–1900 in each of these three elections. Data from the 1994 election is of poorer quality and could not be used to define borderline groups. However, data from all elections between 1994 and 2006 will be used to construct outcome variables (see below for details). The 2006 data also contains information that is useful for descriptive purposes. The analysis is done only on the seven parties that dominated national politics during the studied period. Local parties are thus excluded.[29]

Two crucially important features of the data are, first, that it contains the same information on all candidates irrespective of whether they were elected or not. Second, except for the 1994 election, it contains all ballot paper rankings so that the final ranking that identifies the borderline groups can be calculated.[30] These two features, alone, make the data unique in its kind. Furthermore, rich register-based information on characteristics such as age, sex, foreign background, educational attainment, labor market status, occupation and various income measures is matched to all candidates using a unique person identifier. The registers are in annual form and cover the years 1990–2006 for all candidates. This enables an empirical analysis that (i) follows candidates over a relatively long time period; (ii) can verify the

elections).

[28]The data is obtained from Statistics Sweden and The Swedish Election Authority. Candidates running for a county council are also covered, but this data will not be used in this paper.

[29]The main reason for excluding local parties is that they are very diverse and would therefore be likely to introduce unnecessary noise.

[30]Because the 1991 and 1998 election data contains somewhat less information than the 2002 election data, some assumptions were needed to find borderline groups in these two elections. See the Appendix for details.

identifying assumptions using pre-determined covariates; and (iii) looks at heterogeneous treatment effects across characteristics such as age and level of education.

The outcome variable to measure income effects is disposable income. It is individualized but measured at the household level. Disposable income is the sum of numerous types of after-tax income of the family, including, e.g., labor income, capital income, pensions and unemployment and sickness benefits. The income measure is meant to capture all types of monetary returns from politics (see, however, the discussion in Section 4). As explained above, short- medium- and long-run effects on income are defined as the average log income over 1–3, 6–8 and 13–15 years after the election, respectively. Averaging over three-year periods reduces some of the noise that often plagues income data. And to avoid result driven by outliers, the three-year averages are censored at the 1^{st} and 99^{th} percentiles. The analysis will be performed on logs of the three-year averages.

To measure effects on future political careers, two outcomes are considered; the probability of begin nominated for and elected into a local council. Short- medium and long-run effects are defined in the first, second and fourth subsequent election, respectively. Descriptive statistics of all outcomes in the sample of candidates in borderline groups with $rank^{\star} = \{-2, -1, 0\}$ are provided in Table 1 in the Appendix.

3.3 Balance of covariates

The register data includes numerous variables measuring the candidate's characteristics. These can be used to test the validity of the identifying assumption, for example by testing the robustness of the regressions results to including them as control variables. A more direct test is to run the regression model in (1) with these variables as outcomes. Recall that the identifying assumption requires that the quality of candidates evolve smoothly around the borderline elected. In particular, the preferred model assumes that a linear control function captures any direct effect of $rank^{\star}$. This implies that there should be no "treatment effect" on these variables. In other words, the estimate of β_1 should not differ from zero.

Column 1 of Table 1 provides t-statistics of the β_1 estimate from running these regressions.[31] Indeed, all t-statistics are small enough to support the validity of the model. There is no additional change in characteristics among the borderline elected, above and beyond any direct linear rank effect.[32] Sim-

[31] All time-variant covariates in Table 1 are set one year before the election. All variables are therefore pre-determined which is a prerequisite and should thus not be affected by the treatment. Note that this also refers to disposable income, which will serve as outcome variable but is then measured after rather than before the election.

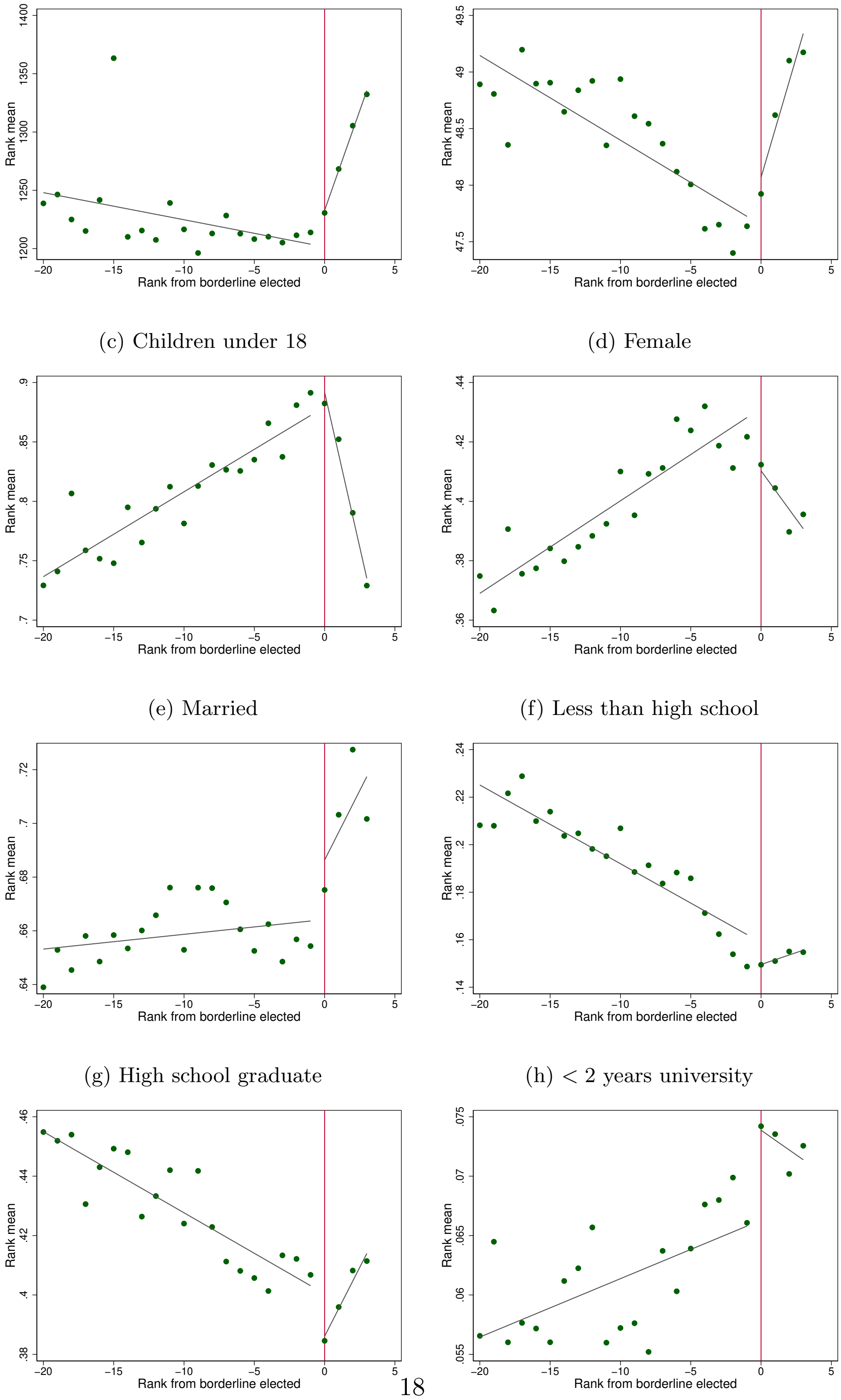
[32] An analogous test is to run a regression of the binary variable *elected* on $rank^{\star}$ and all covariates in Table 1 and test for joint significance of the covariates. Doing this, the obtained F-statistic is 0.80 (p-value 0.71), thus strengthening the confirmation of no

Table 1: Balance and representativity of pre-determined characteristics of candidates in the borderline groups

	β_1	Sample		
	t-stat.	Borderline groups	All elected	All non-elected
Disposable income	0.88	1204.6	1345.9	1189.4
		(522.8)	(574.0)	(514.9)
Age	0.22	47.7	49.3	47.9
		(12.1)	(10.8)	(12.9)
Children under 18	-0.46	0.88	0.75	0.81
		(1.18)	(1.10)	(1.14)
Female	-1.08	0.41	0.40	0.40
		(0.49)	(0.49)	(0.49)
Married	1.50	0.66	0.71	0.66
		(0.47)	(0.45)	(0.47)
Less than high school	0.57	0.15	0.16	0.20
		(0.36)	(0.37)	(0.40)
High school graduate	-1.13	0.40	0.40	0.43
		(0.49)	(0.49)	(0.49)
$<$ 2 years university	1.41	0.070	0.072	0.061
		(0.26)	(0.26)	(0.24)
$\geq$ 2 years university	0.15	0.37	0.36	0.30
		(0.48)	(0.48)	(0.46)
Graduate studies	-0.92	0.0093	0.0094	0.0083
		(0.096)	(0.097)	(0.091)
Born in Sweden	-0.47	0.94	0.95	0.94
		(0.25)	(0.22)	(0.25)
Born in other Nordic country	-0.79	0.030	0.026	0.029
		(0.17)	(0.16)	(0.17)
Born in non-Nordic Europe	1.09	0.018	0.017	0.020
		(0.13)	(0.13)	(0.14)
Born in North America	-0.08	0.0023	0.0011	0.0021
		(0.048)	(0.033)	(0.045)
Born elsewhere	0.94	0.014	0.0091	0.014
		(0.12)	(0.095)	(0.12)
Both parents foreign-born	-0.15	0.010	0.0068	0.0087
		(0.100)	(0.082)	(0.093)

Note: Column 1 reports the t-statistic of the estimate of β_1 from running equation (1) on each of the variables on the sample of candidates in the borderline groups. Columns 2–4 report the mean and standard deviation (in parentheses) in the different samples. The borderline groups include candidates with $rank^{\star} = \{-2, -1, 0\}$. All variables are measured one year before the election. Income is measured in 100 SEK deflated to 2000 year values (9 SEK≈1 USD). The education variables indicate highest completed level. Born elsewhere equals one for individuals born in Africa, Asia, Oceania, Russia or S. America. Both parents foreign-born equals one for individuals born in Sweden but with both parents foreign-born. All variables but Disposable income, Age and Children under 18 are binary.
Source: Statistics Sweden.

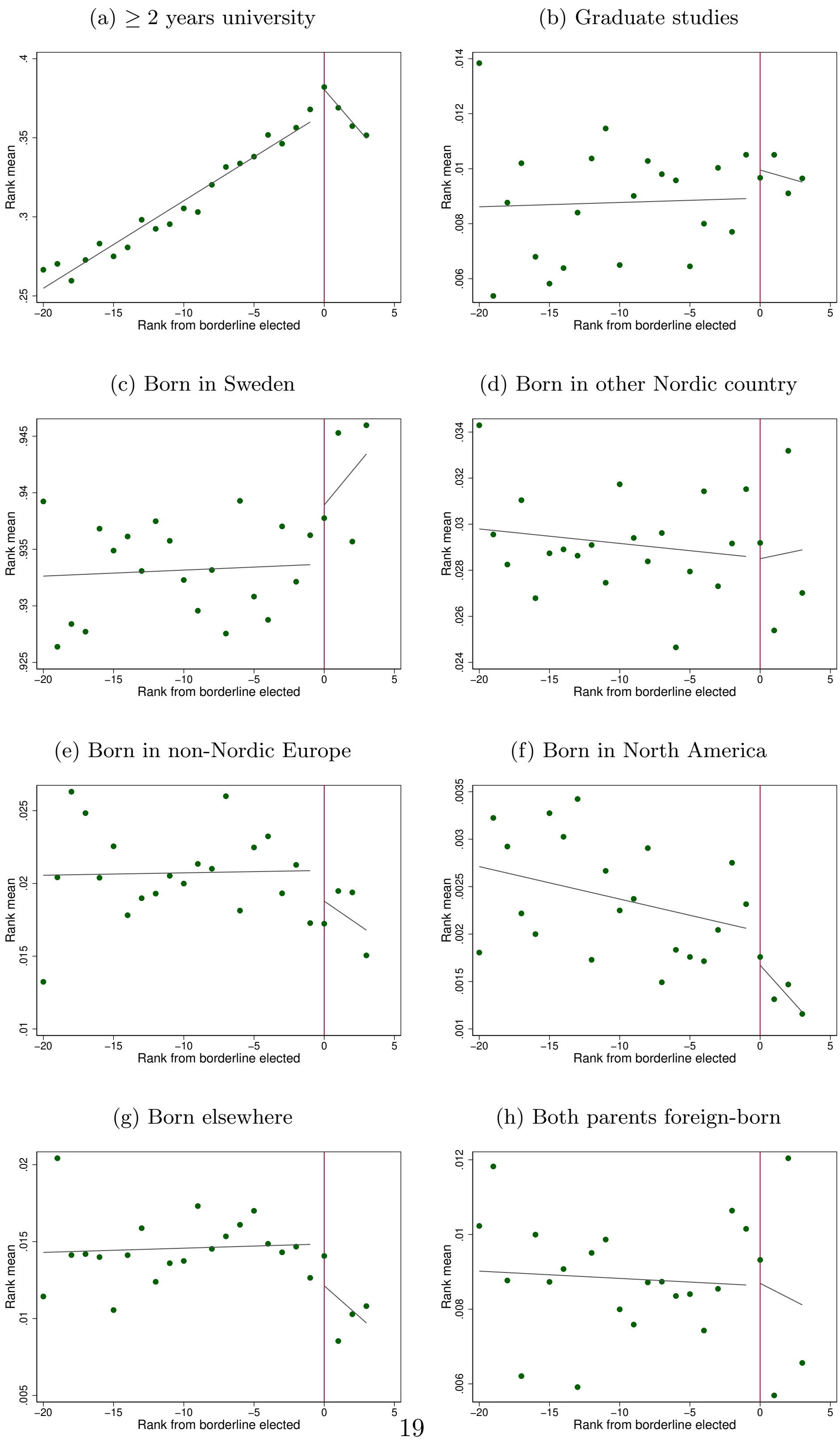
Figure 4: Balance of covariates



(a) Disposable income (b) Age (c) Children under 18 (d) Female (e) Married (f) Less than high school (g) High school graduate (h) < 2 years university

Note: The figures plot (i) covariate means by rank from borderline elected, measured one year before the election; and (ii) fitted regression lines on each side of the threshold. See Table 1 for variable definitions.
Source: Statistics Sweden & The Swedish Election Authority.

Figure 5: Balance of covariates (continued)



Note: The figures plot (i) covariate means by rank from borderline elected, measured one year before the election; and (ii) fitted regression lines on each side of the threshold. See Table 1 for variable definitions.
Source: Statistics Sweden & The Swedish Election Authority.

ilarly, Figures 4–5 illustrate the covariate balance in a broader sample than the borderline groups. Some of the characteristics vary smoothly on both sides of the threshold at the borderline elected. But for some, the pattern changes somewhat to the right of the threshold. Recall from the discussion of Figure 2 above that borderline groups that do include candidates with positive ranks may be selective. Again, this is the rational for estimating the linear control function using non-elected rather than elected candidates.

To see how the candidates in the borderline groups differ from other candidates, Table 1 also shows the mean and standard deviation of the individual characteristics for three different samples.[33] Column 2 includes candidates in the borderline groups, column 3 includes all elected candidates, and column 4 includes all non-elected candidates. In terms of age and marital status, the borderline groups are more similar to the non-elected sample, whereas in terms of education they are more like the elected sample. In general, the representativity is hence quite good.

4 What do the treatment and outcome variables capture?

This section discusses potential drawbacks of the RD design, with a focus on external validity. It also discusses how to measure monetary returns, how to (if at all) measure non-monetary returns, and ultimately what can be concluded about what motivates politicians.

4.1 How marginal is the borderline elected?

A potential drawback with RD designs is that the effect identified often has strong internal validity at the expense of external validity. The reason is that, by definition, the treatment effect is only locally identified. In the current setting, this may have consequences for the interpretation of the results. Specifically, one may worry that returns to office are different (presumably smaller) for borderline elected candidates than for elected candidates further up the ranking. This pertains both to monetary as well as non-monetary returns.

As in most RD designs, the current setting is not able to circumvent the drawback of a locally identified effect. But I do conduct a set of analyses to get a sense of the external validity of the results.[34] One such analysis is to study effects separately for small parties with only 1–2 seats. For these

non-linearities in the direct rank effect.

[33]The descriptive statistics refer to the 1991, 1998 and 2002 elections, which are studied in the paper.

[34]For one thing, we know from Table 1 that the borderline group constitute a more or less representative sample of candidates.

parties, the borderline elected are more like the average politician. Furthermore, I introduce an alternative identification strategy. In this approach, *all* elected candidates are compared to non-elected candidates in a difference-in-difference setting. The external validity is likely stronger than in the main RD approach, but perhaps instead at the expense of the internal validity.[35] Ultimately, when interpreting the results, we still need to keep in mind that the estimated RD effects are local.

Another aspect of the RD effect is whether the treatment is sharp or fuzzy at the threshold at $rank^{\star} = 0$. Specifically, the borderline defeated are quite likely to serve as replacements. A replacement can stand in for several regular members, and the total number of replacements to be appointed is decided by the council prior to the election (as a share below half of the total seats won). If any regular council member resigns and is replaced by a borderline defeated candidate (or if the borderline elected resigns), the variation in treatment status will thus be fuzzy at the threshold.

Fortunately, for the 2002 and 2006 elections, there is information on early resignations and effective replacements that can tell the extent to which the treatment effects obtained from running the regression in (1) underestimate effects of being de facto treated (i.e., actually having served in the council). Let the borderline elected candidates be defined as treated only if they did not resign in their first year. And let defeated candidates instead be defined as treated if they replaced someone at least 300 days before the next election. Then, according to the 2002 and 2006 data, 95% of all borderline elected are treated, and 40% of all borderline defeated. The corresponding percentage among candidates with $rank^{\star} = -2$ is around 20%.If this information were available for all elections, a fuzzy RD design would be ideal. As revealed by these percentages, running the corresponding first stage on the 2002 and 2006 data on the borderline groups yields an estimate of around 0.30 (with a t-statistic of 18.5). Although the treatment of having actually served in the council is not sharply determined by $rank^{\star}$, there is thus still substantial discontinuous variation at the threshold at $rank^{\star} = 0$.[36]

4.2 How to measure returns to politics?

To fully capture monetary returns in the data is not easy. I use disposable income because it is a broad measure that captures as much of monetary returns as possible. With the available data, it is also possible to check the sensitivity of the results to alternative income measures.

The effect of being elected on disposable income will be positive if indi-

[35] Yet another drawback of the difference-in-difference strategy is that it cannot be used to estimate effects on future political careers. See Section 6.3.

[36] Another aspect is that there are other forums for gaining political experience outside of the council in committees, but also this type of work is more common among the borderline elected than among the borderline defeated (see Lundqvist, 2011).

viduals acquire certain skills that are rewarded in the labor market, if there is a positive signaling effect or if the individuals develop closer ties to certain firms or organizations. Note that such returns could be retained while still in politics, since the majority of local politicians hold regular jobs and, at least partly, devote their spare time politics. While still in politics, there is also the direct effect of official perquisites and remunerations. There is, however, also the possibility of mechanisms operating in the opposite direction: political engagement may require foregone earnings because of time and effort constraints.[37]

Some forms of monetary returns cannot be captured in disposable income or other available measures in the data. These include improved access to higher-quality housing for example via building permits, and lower costs for other types of goods and services. The same is true for monetary returns in the form of outright bribes, as these are unlikely to show up in official income registers. But to the extent that politicians attempt to hide parts of their (illegitimate) income by transferring *official* income within the household, such returns will show up in their disposable income. All in all, disposable income likely captures most but not all types of monetary returns.

Capturing non-monetary returns in the data is much more challenging. In fact, I make no attempt to measure this directly. Instead, I argue that if they are important drivers, politicians should want to continue their political career after having been elected. I therefore investigate how being randomly elected into a municipal council affects future political career prospects. Furthermore, if non-monetary returns are what matters, the willingness to pursue a political career should be independent of any monetary returns. Comparing the effects on income with the effects on the future political career can therefore provide indirect evidence of what motivates politicians. That being said, the paper cannot directly study what motivates politicians.

Two variables measure how being borderline elected affects the political career prospects; being nominated for and being elected in subsequent elections to municipal councils.[38] The idea is that the results will indicate if being elected has a positive encouragement effect on continuing in politics through increased likelihood of accepting a nomination. Alternatively, being elected may imply learning and being disappointed by what local politics entails, which would then discourage future political engagement. Such encouragement and learning effects can involve both monetary and non-monetary aspects. Of course, the results can also be interpreted

[37]The Municipal Law (4 Ch. 12§) states that elected representatives have the right to be "reasonably compensated" for foregone earnings due to their political assignments.

[38]An earlier version of the paper also studies the effect of being borderline elected locally on the probability of advancing to national politics. Because the parliament only has 349 seats, being elected is a very rare event. Hence, only the probability of getting nominated to the national parliament are studied. These effects are positive in the first and second subsequent elections. See Lundqvist (2011).

as incumbency effects operating via voters and via the candidate selection process in the party organization.[39]

In practice, both monetary and non-monetary returns may be important motivational factors. They may even be dependent on each other, in the sense that non-monetary returns such as power may give rise to opportunities for monetary benefits. For analytical purposes, there are still good reasons to study them separately. Again however, non-monetary returns will only be indirectly measured via future political career prospects. Neither is disposable income a perfect measure of monetary returns. This needs to be kept in mind when interpreting the results.

5 Monetary returns from being elected

This section quantifies the monetary returns from politics by analyzing the effect of being elected into a municipal council on income. Specifically, effects are estimated on disposable income 1–3, 6–8 and 13–15 years after being elected, respectively. The three different periods are referred to as short, medium and long run. The analysis combines graphical presentations with econometric methods as described in Section 3.

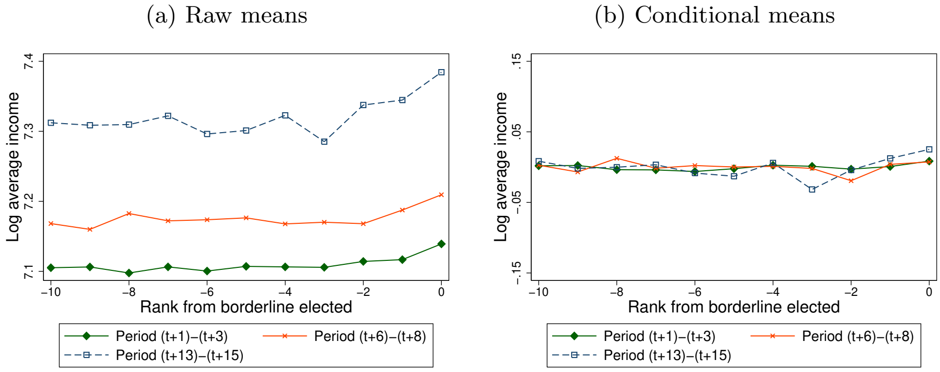
Let us first look at the graphics in Figure 6. It plots the $rank^{\star}$-specific means of disposable income in the three different periods. The plot to the left shows raw means, whereas the plot to the right shows conditional means obtained from a regression of the outcome variable on a set of individual controls measured one year before the election; the number of children aged below 18 and a set of dummies for age, gender, marital status, income quantile, highest completed education, foreign background and past political experience. Recall that the variable $rank^{\star}$ is defined as the difference between a candidate's final rank and the final rank of the borderline elected, so that it takes the value zero for the borderline elected and negative values for non-elected candidates.

Direct effects of $rank^{\star}$ on the outcome are represented by the overall slope of the lines connecting the $rank^{\star}$-specific means. Conceptually, the treatment effect is the difference between the borderline elected ($rank^{\star} = 0$) and the borderline defeated ($rank^{\star} = -1$) that is above and beyond the difference between any other two candidates. Visually, a treatment effect therefore corresponds to a jump at $rank^{\star} = -1$. The raw means to the left thus reveal small or zero effects on income from being elected.[40] This is particularly clear for medium-run income, where any jump at $rank^{\star} = -1$

[39]For example, Cirone et al. (2020) present evidence of the importance of seniority for candidate progression in closed-list systems.

[40]Not only are the treatment effects absent, but what might be somewhat surprising is that also the direct effects of $rank^{\star}$ are negligible. Thus, to the extent that income is a proxy for ability (in some broader sense), candidates around the borderline elected are not ranked according to this.

Figure 6: Short-, medium- and long-run disposable income



Note: The figures plot means of disposable income by rank from borderline elected in election year *t*. Income is deflated to 2000 year values and measured as logs of three-year averages in the short run (years t+1 to t+3), medium run (years t+6 to t+8) and long run (years t+13 to t+15). Conditional means are the residuals obtained from a regression of the outcome variable on the following individual controls measured one year before the election: the number of children aged below 18 and a set of dummies for age, gender, marital status, income quantile, highest completed education, foreign background and past political experience.
Source: Statistics Sweden & The Swedish Election Authority.

is completely absent. For short- and long-run income, a slight jump can be detected. For the latter, however, as there is a considerably more distinct jump at $rank^\star = -3$, this is more likely to be due to random variation than to a treatment effect.

Comparing the left and the right plots, the main difference is that there is a mean-adjustment to zero for all income periods (as these are residuals). Although this adjustment makes the plot less clear, it is suggestive of the same pattern as in the raw means. Hence, this suggests that to extensively control for pre-determined characteristics would not alter the results.

The econometric counterpart to the plots in Figure 6 is given in Table 2. The table thus shows results from estimating equation (1) on candidates with $rank^\star = \{-2, -1, 0\}$. Note that the parameter β_2 is the marginal effect of $rank^\star$ and thus corresponds to the overall slopes in the plots, whereas β_1 is the main parameter of interest that captures the additional effect of having $rank^\star = 0$, or the effect of being elected.

The results in column 1 are for short-run income without any further controls. Column 2 controls for the same set of individual controls as in the right plot with the conditional means. Equivalent results for medium- and long-run income are given in columns 3–4 and 5–6, respectively. As seen in the table, none of the estimated treatment effects are statistically significant. The point estimates are very close to zero either with controls (for the short and long run) or without controls (for the medium run). Note especially that the suspected jumps in short- and long-run income seen graphically are not statistically significant. Qualitatively, the inclusion

Table 2: Effects of being elected on disposable income

	Period (t+1)–(t+3)		Period (t+6)–(t+8)		Period (t+13)–(t+15)	
	(1)	(2)	(3)	(4)	(5)	(6)
elected	0.0199	0.00412	0.00271	-0.0184	0.0327	-0.00823
	(0.0124)	(0.00888)	(0.0176)	(0.0154)	(0.0282)	(0.0246)
rank⋆	0.00260	0.00409	0.0192*	0.0230***	0.00713	0.0192
	(0.00714)	(0.00495)	(0.0105)	(0.00883)	(0.0165)	(0.0145)
Observations	16673	16673	10915	10915	5283	5283
Individual controls	no	yes	no	yes	no	yes

Note: The table reports effects of being elected into a municipal council on disposable income measured as logs of three-year averages 1–3 (columns 1–2), 6–8 (columns 3–4) and 13–15 (columns 5–6) years after the election. Individual controls measured one year before the election are: number of children aged below 18 and a set of dummies for age, gender, marital status, income quantile, highest completed education, foreign background and past political experience. Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

of controls makes no difference, and—although the size of the estimates changes when controls are included—estimates with and without controls are within the 95% confidence interval of one another. In all regards, the econometric results thus confirm the graphical inspection.

5.1 Returns while in vs. after exiting politics

Some candidates elected in a particular election are still active politicians 6–8 and 13–15 years later, while others are not. The candidates' medium- and long-run income should be seen as the result of optimizing behavior, which may lead to political careers of different length for different people. But it is also of interest to see whether returns to politics kick in after leaving politics. In general, however, looking at income conditional on exiting politics is problematic since exit is endogenous. For example, some politicians may exit because they expect it to be profitable. Others may exit because they were unsuccessful incumbents.

A way of circumventing this problem is to look at the income profile of candidates who were not only elected by chance, but who also left politics for exogenous reasons. An exogenous source of variation in exit rates that lies close at hand is being borderline defeated. To this end, Figure 7 plots the income profile of candidates who were borderline elected in 1998 *and* borderline defeated in 2002. This is to be compared with the income profile of candidates who were borderline defeated in 1998 and not elected in 2002 either, whose income profile is also seen in the figure. These candidates were neither elected in previous elections in the data (1991 or 1994), nor in the next election (2006). 975 candidates of the borderline defeated in 1998 satisfy these conditions, but only 59 of the borderline elected in 1998.

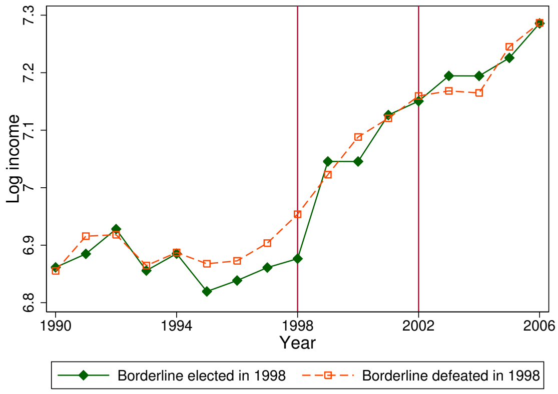
As seen in the figure, the income levels are very similar several years before the 1998 election, while being in office as well as after exiting. The exceptions are the few years preceding the 1998 election, when the income of those who get elected later is lower. Because of the small sample size, one should be careful about reading too much into this pattern. But the figure does suggest that there is no income gain following the exogenous exit from politics in 2002.

5.2 Robustness of the effects on monetary returns

The main result from above is that monetary returns from local politics are absent. This section investigates the robustness of this result and thus the internal validity of the RD estimates.

First, consider the inference problem which follows from the forcing variable being discrete. As explained in Section 3, the solution suggested by Lee and Card (2008) to cluster at the level of the forcing variable is not applicable here. Instead, to get a sense of the robustness of the estimated standard

Figure 7: Disposable income while in vs. after exiting politics



Note: The figure plots average disposable income among candidates who were borderline elected in 1998 *and* borderline defeated in 2002, and among candidates who were borderline defeated in 1998. Income is measured in logs of 100 SEK deflated to 2000 year values.

Source: Statistics Sweden & The Swedish Election Authority.

errors and the associated t-statistics, the model is estimated with different levels of clustering. The resulting t-statistics are provided in Table 3. The top panel shows t-statistics from the original model where the standard errors are clustered at the municipality level. Moving down the table, it is clear that these t-statistics are very robust to alternative levels of clustering. This is even true for the long-run estimate clustered at the county-by-election, which only results in 21 clusters.

Next, consider the implication of the identifying assumption that parties should not be able to perfectly anticipate which candidates that will be elected. To investigate whether this is likely to hold, I conduct a robustness analysis by re-estimating the baseline results in Table 2 using only cases when it indeed is unlikely that the parties could have known who would be the borderline elected. The idea is that the ranking of candidates would be different if it was a priori certain who would actually be elected.

I propose a number of instances when there are reasons to believe that the election outcome was more uncertain. The results from these regressions are presented in Table 4 (where column 1 reproduces the baseline results with controls in column 2 of Table 2): (i) the party's number of seats changed from the previous election, see column 2; (ii) the party won their last seat or were close to winning an additional seat with a vote margin of less than 1 and 0.5%, see columns 3–4; and (iii) a combination of (i) and (ii), see column 5.[41] The table shows estimates that are statistically insignificant and close to zero across all the different specifications. That is, the results

[41] As vote margins, I use the minimum changes in votes to win or lose an additional seat in proportional elections as defined and calculated by Folke (2014).

Table 3: Robustness of t-statistics of effects on disposable income

	Period (t+1)–(t+3)	Period (t+6)–(t+8)	Period (t+13)–(t+15)
Cluster	(1)	(2)	(3)
Municipality	0.46	-1.19	-0.33
$N_{clusters}$	290	289	285
Borderline group	0.46	-1.21	-0.33
$N_{clusters}$	5596	3718	1880
Party-by-county	0.55	-1.22	-0.37
$N_{clusters}$	147	147	147
County-by-election	0.45	-1.22	-0.50
$N_{clusters}$:	63	42	21
Observations	16673	10915	5283

Note: The table reports t-statistics and the number of clusters with different levels of clustering for estimated effects of being elected into a municipal council on disposable income measured as logs of three-year averages 1–3 (column 1), 6–8 (column 2) and 13–15 (column 3) years after the election. All regressions include individual controls (see Table 2).

from these set of elections which arguably were more uncertain are very similar to the baseline results. This can be taken as evidence that parties in general cannot perfectly anticipate how many votes they will win and rank their candidates accordingly.

As argued in Section 3, the quality of candidates is more likely to evolve smoothly just around the borderline elected. This is the rationale for restricting the borderline group to three candidates and the direct effect of rank to be linear. The robustness of this baseline specification is investigated by estimating versions of equation (2). In these regressions, five or ten defeated candidates are included and the direct effect effect of rank is specified as linear, quadratic or qubic. Thus, in RD language, these regressions vary the bandwidth and the order of polynomial of the control function.

The result from this exercise is given in Table 5. Ten defeated candidates are included in columns 1, 3 and 5, and five defeated candidates are included in column 2, 4 and 6. Each column contains results from three different regressions with different order of polynomial $\hat{p}$, and the $\hat{p}$ preferred by the Akaike information criterion (AIC) is in bold. The table shows somewhat mixed results, with some significant estimates of which one is even negative. In general however, most estimates are close to zero. Indeed, none of them imply any economically significant monetary returns (around 1%; see columns 1 and 3). The overall conclusion from this exercise is hence that the RD model is somewhat sensitive to exact specification, but that the baseline result that monetary returns from local politics largely are absent holds.

Recall that the borderline group only includes the borderline elected

Table 4: Effects of being elected on disposable income in uncertain elections

	(1)	(2)	(3)	(4)	(5)
		Period $(t+1)$–$(t+3)$			
elected	0.00412	0.0110	0.00986	0.00985	0.0179
	(0.00888)	(0.0103)	(0.0102)	(0.0139)	(0.0120)
Observations	16673	12637	13218	7699	10037
		Period $(t+6)$–$(t+8)$			
elected	-0.0184	-0.00246	-0.00724	0.00768	0.0112
	(0.0154)	(0.0177)	(0.0171)	(0.0210)	(0.0199)
Observations	10915	8437	8706	5129	6749
		Period $(t+13)$–$(t+15)$			
elected	-0.00823	0.00440	-0.00849	0.0285	0.0106
	(0.0246)	(0.0282)	(0.0288)	(0.0379)	(0.0331)
Observations	5283	4086	4258	2558	3308
Vote margin (%)	no restr.	no restr.	1	0.5	1
$\lvert\Delta\text{seats}\rvert \geq 1$	no	yes	no	no	yes
Individual controls	yes	yes	yes	yes	yes

Note: The table reports effects of being elected into a municipal council on disposable income measured as logs of three-year averages 1–3 (top panel), 6–8 (mid panel) and 13–15 (bottom panel) years after the election. Column 1 reproduces the baseline results in columns 2, 4 and 6 of Table 2. All regressions include individual controls (see Table 2). Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Table 5: Effects of being elected on disposable income; allowing non-linear effects of $rank^{\star}$

	Period $(t+1)$–$(t+3)$		Period $(t+6)$–$(t+8)$		Period $(t+13)$–$(t+15)$	
	(1)	(2)	(3)	(4)	(5)	(6)
elected, $\bar{p}=1$	**0.00924****	**0.00877**	**0.0129***	0.0140	0.0311***	0.0175
	(0.00445)	**(0.00547)**	**(0.00765)**	(0.00902)	(0.0120)	(0.0145)
elected, $\bar{p}=2$	0.00522	0.0111	0.0147	-0.00129	**0.0141**	**-0.0122**
	(0.00591)	(0.00948)	(0.00971)	(0.0154)	**(0.0156)**	**(0.0231)**
elected, $\bar{p}=3$	0.0116	-0.00885	0.00292	**-0.0639****	-0.00717	-0.0736
	(0.00884)	(0.0184)	(0.0142)	**(0.0307)**	(0.0218)	(0.0496)
Observations	54525	32350	35407	21031	16808	10076
rank* $\geq$	-10	-5	-10	-5	-10	-5

Note: The table reports effects of being elected into a municipal council on disposable income measured as logs of three-year averages 1–3 (columns 1–2), 6–8 (columns 3–4) and 13–15 (columns 5–6) years after the election. The AIC-preferred polynomial is in bold. All regressions include individual controls (see Table 2). Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

(and two non-elected) because fewer groups even have elected candidates with higher $rank^{\star}$ (see Figure 2). This may create a selection problem. Baring this in mind, I conduct another robustness analysis using only the sample of groups that indeed have candidates with at least $rank^{\star} = 3$. These are thus parties with at least four seats. Figure 8 shows rank averages of disposable income in these groups, and Table 6 presents RD estimates of the jump at the threshold at $rank^{\star} = 0$. A linear control function is used, and is estimated separately on each side of the threshold. The bandwidth is either ± 2 (columns 1, 3 and 5) or ± 3 (columns 2, 4 and 6). As the table shows, the null result from above is confirmed also in this specification.

A common test of the identifying assumption in RD studies is so-called placebo regressions. It is not possible to directly test if the direct effect of the forcing variable is smooth at the threshold, since the assumption refers to the counterfactual situation when there is no treatment at the threshold. The idea with placebo regressions is instead that the RD estimate truly represents the causal treatment effect if the direct effect is smooth close to—but not at—the threshold. That is, if there is no "treatment effect" as one moves away from the threshold. Such placebo regressions are more relevant in applications where the estimated treatment effect at the threshold is non-zero, but can still provide a sense of validity of the RD model estimated here.

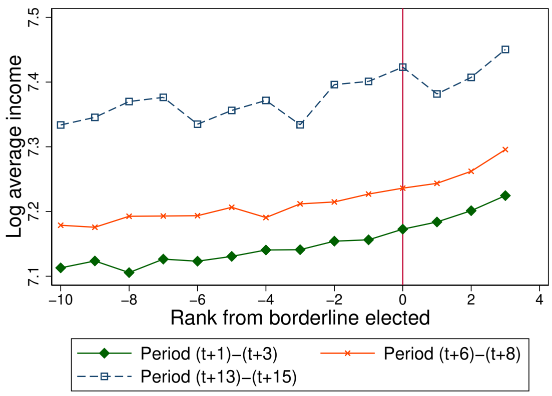
The placebo regressions falsely assign the borderline elected status to candidates with $rank^{\star} = -1$ or $rank^{\star} = -2$. The results can more or less be inferred from Figure 6, while Table 7 gives the exact estimates and their

Table 6: Effects on disposable income in groups with with more elected candidates

	Period $(t+1)$–$(t+3)$		Period $(t+6)$–$(t+8)$		Period $(t+13)$–$(t+15)$	
	(1)	(2)	(3)	(4)	(5)	(6)
elected	-0.00567	0.000625	-0.00864	0.00357	0.0238	0.0127
	(0.0104)	(0.00822)	(0.0165)	(0.0129)	(0.0261)	(0.0191)
rank*	0.0102*	0.00661**	0.0116	0.00316	-0.00652	0.0100
	(0.00538)	(0.00272)	(0.00950)	(0.00456)	(0.0154)	(0.00744)
rank$^\star \times$ elected	-0.0129	-0.00308	-0.0165	0.00564	0.0117	-0.00486
	(0.0106)	(0.00551)	(0.0176)	(0.00826)	(0.0290)	(0.0138)
Observations	12800	17933	8627	12081	4678	6548
Included rank*:	± 2	± 3	± 2	± 3	± 2	± 3
Individual controls	yes	yes	yes	yes	yes	yes

Note: The table reports effects of being elected into a municipal council on disposable income measured as logs of three-year averages 1–3 (columns 1–2), 6–8 (columns 3–4) and 13–15 (columns 5–6) years after the election. All regressions include individual controls (see Table 2). Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Figure 8: Effects on disposable income in groups with more elected candidates



Note: The figure plots means of disposable income by rank from borderline elected in election year t. Income is deflated to 2000 year values and measured as logs of three-year averages 1–3 years after the election.

Source: Statistics Sweden & The Swedish Election Authority.

significance. As can be seen, there are no placebo effects in the short-run. But there is a significant 4% jump in medium-run income at $rank^{\star} = -1$, and a 7% jump in long-run income at $rank^{\star} = -2$. Had there been a positive effect also at the true threshold, these results would have put the causality into question. Now, the interpretation is simply instead that there is some variation between candidates that is not captured by a linear control function. This is however not to say that any true positive effects of being elected are likely.

Recall that the argument for studying effects on disposable income is that it is a broad measure that captures as much of monetary returns as possible. The main result from above is that there are no significant effects on disposable income of being elected, and thus that monetary returns from local politics are largely absent. Tables 8 and 9 present results for two alternative income measures; labor income and income from largest source.[42] All estimates are small and statistically insignificant, thus providing additional support to the conclusion from above.

6 Is it worth it?

Given that monetary returns from politics are absent, one may be inclined to conclude that this is *not* what motivates politicians. On the other hand, it is

[42]These measures contain zeroes and are therefore not logged. The unit of measurement is three-year averages in 100 SEK, and the sample averages are approximately 1600 SEK for both income measures; see Table 1 in the Appendix for descriptive statistics.

Table 7: Placebo estimates of the effects on disposable income

	Period (t+1)–(t+3)		Period (t+6)–(t+8)		Period (t+13)–(t+15)	
	(1)	(2)	(3)	(4)	(5)	(6)
elected^{placebo}	0.00657	-0.00181	0.0403**	-0.0131	-0.0123	0.0701**
	(0.00856)	(0.00948)	(0.0155)	(0.0162)	(0.0262)	(0.0271)
rank*	-0.00368	-0.00201	-0.0171*	-0.00364	0.0292*	-0.0392**
	(0.00531)	(0.00570)	(0.00914)	(0.00922)	(0.0150)	(0.0154)
Observations	16361	15998	10645	10335	5077	4867
Cut-off at rank*:	-1	-2	-1	-2	-1	-2
Individual controls	yes	yes	yes	yes	yes	yes

Note: The table reports placebo estimates of being elected into a municipal council on disposable income measured as logs of three-year averages 1–3 (columns 1–2), 6–8 (columns 3–4) and 13–15 (columns 5–6) years after the election. All regressions include individual controls (see Table 2 in the paper). Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

possible that the mere hope of positive returns is what motivates people to engage in politics, and that the absence of returns will be a disappointment.

This section studies how being elected into a municipal council affects the future political career. The idea is that if politicians are motivated by something else than monetary returns, they should want to pursue a political career—despite the fact that (they learn that) there are no monetary returns. Using the same RD model as above, I estimate causal effects of being elected into a municipal council on the probability of running for and being elected into a municipal council in subsequent elections. In addition to studying average effects, I also conduct a thorough heterogeneity analysis. The paper cannot *directly* study what motivates politicians, but comparing the effects on income with the effects on the future political career across several subgroups provides *indirect evidence.*

6.1 Effects on future political careers

I start by assessing if being elected into a municipal council in election year t has an effect on the probability of *running* in future elections to a municipal council in the first, second and fourth subsequent election, respectively. The results are presented in Figure 9 and Table 10. The figure and the table are to be read in the same way as the baseline results from income in the previous section (except that the outcome is now the probability of being nominated for election instead of income).

Also these graphs are quite illustrative. Not surprisingly, there is a positive direct relationship between *rank** and the probability of running in future elections to a municipal council, as seen from the overall positive slopes. However, there is little evidence of any treatment effect of being

Table 8: Effects of being elected on labor income

	Period $(t+1)$–$(t+3)$		Period $(t+6)$–$(t+8)$		Period $(t+13)$–$(t+15)$	
	(1)	(2)	(3)	(4)	(5)	(6)
elected	27.13	-32.31	24.66	-10.13	-53.95	-26.25
	(38.10)	(31.31)	(58.45)	(47.41)	(99.07)	(75.38)
rank*	50.30**	58.18***	28.50	45.33*	57.78	63.59
	(22.38)	(17.69)	(33.82)	(26.95)	(60.30)	(46.65)
Observations	16673	16673	10915	10915	5283	5283
Individual controls	no	yes	no	yes	no	yes

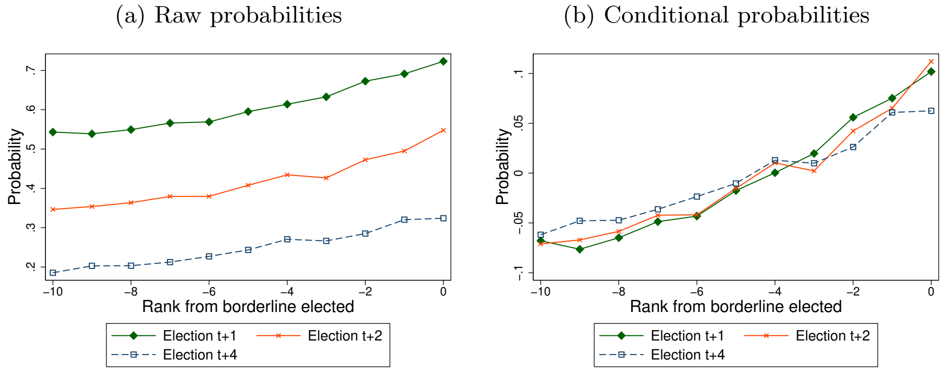
Note: The table reports effects of being elected into a municipal council on labor income measured as three-year averages 1–3 (columns 1–2), 6–8 (columns 3–4) and 13–15 (columns 5–6) years after the election. Individual controls measured one year before the election are: number of children aged below 18 and a set of dummies for age, gender, marital status, income quantile, highest completed education, foreign background and past political experience. Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Table 9: Effects of being elected on income from largest source

	Period (t+1)–(t+3)		Period (t+6)–(t+8)		Period (t+13)–(t+15)	
	(1)	(2)	(3)	(4)	(5)	(6)
elected	13.75	-40.21	6.733	-26.37	-4.781	19.46
	(35.57)	(28.92)	(53.22)	(43.42)	(89.51)	(66.27)
rank⋆	20.25	28.11*	8.946	25.20	-1.767	4.559
	(21.15)	(16.75)	(30.91)	(24.82)	(54.61)	(40.77)
Observations	16673	16673	10915	10915	5283	5283
Individual controls	no	yes	no	yes	no	yes

Note: The table reports effects of being elected into a municipal council on income from largest source measured as three-year averages 1–3 (columns 1–2), 6–8 (columns 3–4) and 13–15 (columns 5–6) years after the election. Individual controls measured one year before the election are: number of children aged below 18 and a set of dummies for age, gender, marital status, income quantile, highest completed education, foreign background and past political experience. Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Figure 9: Probabilities of being nominated in future elections to a municipal council



Note: The figures plot the probability of being nominated in future elections to a municipal council by rank from borderline elected in election year t. Conditional probabilities are the residuals obtained from a regression of the outcome variable on the following individual controls measured one year before the election: the number of children aged below 18 and a set of dummies for age, gender, marital status, income quantile, highest completed education, foreign background and past political experience.
Source: Statistics Sweden & The Swedish Election Authority.

elected, as there is no jump between the borderline elected and defeated (at $rank^{\star} = -1$), except maybe in the second subsequent election. This pattern is confirmed by Table 10, where all estimates are statistically insignificant except the one for the election in $t + 2$ (at the 10% level and only without individual controls). Controlling for individual characteristics barely affects the point estimates and—just as for income—the only graphical differences are in the intercepts.

Figure 9 shows that there is a high degree of persistence in running in subsequent elections, especially the first (the left plot shows large average probabilities of running). A possible interpretation of the null result above is therefore that neither the borderline elected nor the borderline defeated are discouraged to run again. But what is relevant is ultimately how being elected once affects the probability of being elected, nut only running, in subsequent elections.

These reelection probabilities are assessed in Figure 10 and Table 11.[43] Focusing on candidates with $rank^{\star} = \{-2, -1, 0\}$, from the graphics one can detect a positive short-run treatment effect (i.e., being elected in the

[43]Note that these are unconditional reelection probabilities, in the sense that they are not conditional on running. The reason for this is that the decision to run in future elections can conceptually be an outcome due to the treatment, which means that a causal interpretation of the conditional effects on being reelected would not be valid. In practice, because of the previous result that there are no large effects on running probabilities, conditioning on running only scales up the future reelection probabilities without making a qualitative difference.

Table 10: Effects of being elected on the probability of being nominated in future elections to a municipal council

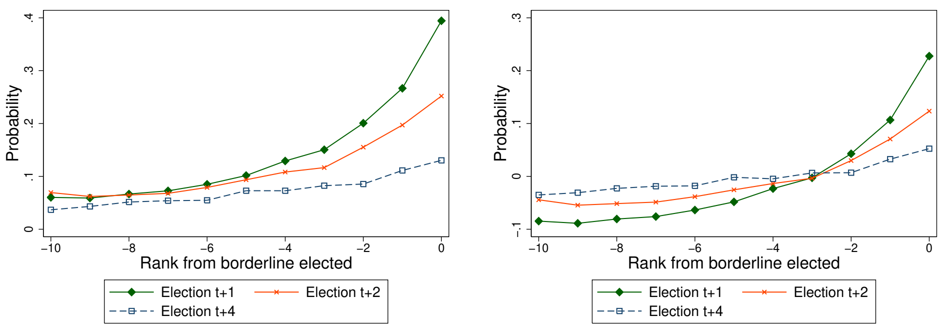
	Election t+1		Election t+2		Election t+4	
	(1)	(2)	(3)	(4)	(5)	(6)
elected	0.0133	0.00960	0.0302*	0.0251	-0.0316	-0.0328
	(0.0133)	(0.0133)	(0.0183)	(0.0180)	(0.0264)	(0.0267)
rank*	0.0186**	0.0198**	0.0224**	0.0234**	0.0353**	0.0348**
	(0.00778)	(0.00782)	(0.0109)	(0.0107)	(0.0157)	(0.0159)
Observations	16754	16754	11208	11208	5710	5710
Individual controls	no	yes	no	yes	no	yes

Note: The table reports effects of being elected into a municipal council on the probability of being nominated in the first (columns 1–2), second (columns 3–4) and fourth (columns 5–6) subsequent election to a municipal council. Individual controls measured one year before the election are: number of children aged below 18 and a set of dummies for age, gender, marital status, income quantile, highest completed education, foreign background and past political experience. Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

first subsequent election), as there is a jump between the borderline elected and defeated. According to columns 1–2 in Table 11, this effect is a statistically significant 6 percentage points and it is unaffected when controlling for individual characteristics. As suggested by Figure 10 and as confirmed in columns 3–6 in Table 11, there are, however, no effects of being elected in election t on also being elected in elections $t + 2$ and $t + 4$.

Figure 10: Probabilities of being elected in future elections to a municipal council

(a) Raw probabilities (b) Conditional probabilities



Note: The figures plot the probability of being elected in future elections to a municipal council by rank from borderline elected in election year t. Conditional probabilities are the residuals obtained from a regression of the outcome variable on the following individual controls measured one year before the election: the number of children aged below 18 and a set of dummies for age, gender, marital status, income quantile, highest completed education, foreign background and past political experience.
Source: Statistics Sweden & The Swedish Election Authority.

In terms of magnitude, 6 percentage points amount to about the same size as the direct effect of $rank^{\star}$, and around 20% of the mean election rate in election $t + 1$ among candidates in the borderline group (see the descriptive statistics in the Appendix).[44] In other words, there is a non-negligible effect of being elected once on being reelected. A possible interpretation is that those elected indeed want to continue in politics up to eight years (two election periods), even after learning that monetary returns are absent. Section 6.3 extends this idea by looking not only at whether this holds on average but also across various subgroups.

[44]However, recall from the discussion in Section 4 that a fair share of the borderline defeated who initially were council replacements in fact overtook a permanent council seat. If treatment is instead defined as actually having served in the council, then treatment status is fuzzy at the threshold at $rank^{\star} = 0$. Evaluating the magnitude of such a treatment effect requires scaling up the coefficient by around three (since the first stage is estimated to around 0.30). Thus, the obtained result means that having served in the council in the previous election period explains as much as 60% of the probability of being elected in the next election.

Table 11: Effects of being elected on the probability of being elected in future elections to a municipal council

	Election t+1		Election t+2		Election t+4	
	(1)	(2)	(3)	(4)	(5)	(6)
elected	0.0619***	0.0583***	0.0131	0.0112	-0.00664	-0.00651
	(0.0139)	(0.0138)	(0.0157)	(0.0155)	(0.0179)	(0.0177)
rank⋆	0.0659***	0.0663***	0.0418***	0.0417***	0.0257**	0.0264**
	(0.00728)	(0.00732)	(0.00858)	(0.00853)	(0.0102)	(0.0102)
Observations	16754	16754	11208	11208	5710	5710
Individual controls	no	yes	no	yes	no	yes

Note: The table reports effects of being elected into a municipal council on the probability of being elected in the first (columns 1–2), second (columns 3–4) and fourth (columns 5–6) subsequent election to a municipal council. Individual controls measured one year before the election are: number of children aged below 18 and a set of dummies for age, gender, marital status, income quantile, highest completed education, foreign background and past political experience. Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

6.2 Robustness of the effects on future political careers

Before turning to the heterogeneity analysis, this section briefly discusses the robustness of the estimated reelection effects. The same robustness analysis is conducted as was done for the effect on disposable income, but to save on space, the results are referred to the Appendix.

First, consider the inference problem. Appendix Table 2 shows the t-statistics to be very robust across the different levels of clustering. Again, although it is unfortunate that the solution suggested by Lee and Card (2008) is not applicable, this suggest that the problem is not severe.

Second, consider the specifications that restrict the sample to cases when it is more unlikely that the parties could have known who would be borderline elected. Table 3 in the Appendix shows estimates that are quite robust across the different specifications. For example, the estimate hardly changes even when the sample size is cut in half as the vote margin of the last seat is restricted to 0.5%. Again, this is evidence that the identifying assumption of the model holds, since it suggests that parties cannot perfectly anticipate which candidates that will be elected. Consequently, they do not base their ranking of candidates on this.

Third, consider the RD regressions that vary the bandwidth and order of polynomial of the control function; see Table 4 in the Appendix. The resulting point estimates as well as the significance levels are somewhat sensitive to different bandwidths and order of polynomial. But restricting the attention to the AIC-preferred specifications, the baseline result in Table 11 is quite robust. Being borderline elected in election t increases the chances of being elected in election $t+1$ by around 5 percentage points (compared to 6 in Table 11). The election probabilities in later elections are not affected, supporting the baseline results.

Fourth, consider the RD analysis that includes elected candidates ranked higher than the borderline elected. Figure 1 and Table 5 in the Appendix again largely confirm the baseline results. The estimated short-run reelection effect is a bit smaller. But note that this specification uses a selective sample of larger parties. And as the analysis below will show, there is quite some heterogeneity in the effects across party size.

Finally, consider the placebo regressions that falsely assign the borderline elected status to non-elected candidates further down the ranking. Appendix Table 6 shows significant placebo estimates in the first subsequent election. This is also where there are significant treatment effects at the true threshold, casting some doubt on whether these can be interpreted as causal. The size of the placebo effects are however only 30–40% of the effect size for the true borderline elected. Moreover, a possible interpretation of the significant placebo estimates is that there are non-linear direct effects of $rank^{\star}$ on future election probabilities. Indeed the results from the previous robustness analysis (see Table 4 in the Appendix) show that higher order of

Table 12: Short-run effects on disposable income and reelection probabilities; across election periods

	Disposable income	Local reelection
elected× 1991	0.00180	0.0511**
	(0.0138)	(0.0255)
elected× 1998	0.00458	0.0864***
	(0.0152)	(0.0254)
elected× 2002	0.00398	0.0310
	(0.0173)	(0.0249)
Observations	16673	16754

Note: The table reports effects of being elected into a municipal council on disposable income measured as logs of three-year averages 1–3 years after the election (column 1) and on the probability of being elected in the first subsequent election to a municipal council (column 2). All regressions include individual controls (see Table 2). Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

polynomials are preferred by the AIC criterion when the bandwidth allows for that. And again, these preferred estimates are very close to the baseline estimate of 6 percentage points. Taken together, the conclusion from these analyses is that the effect of being elected into a municipal council on average increases reelection probabilities with around 5–6 percentage points.

6.3 Heterogeneity and interpretation

I argue that the fact that politicians want to continue in politics up to eight years even after learning that monetary returns are absent can be interpreted as them not being primarily motivated by money. This section lends further support to this argument. In particular, the learning mechanism is much more likely if the result from above holds not only on average but also across several subgroups. Consequently, I conduct an extensive heterogeneity analysis on short-run income and short-run reelection probabilities to see how general the result is that there is no effect on income, but still a non-negligible effect on reelection probabilities.

Heterogeneous effects are studied in different dimensions. Table 12 separates the treatment effect across the three election periods t. Table 13 separates the treatment effect across different types of parties and councils; governing vs. opposition party, party size and council size. Table 14 separates the treatment effect across different individual characteristics; ed-

Table 13: Short-run effects on disposable income and reelection probabilities; across parties and councils

	Disposable income	Local reelection
elected× opposition party	0.0160	0.0652***
	(0.0123)	(0.0186)
elected× governing party	-0.00995	0.0500**
	(0.0133)	(0.0216)
Observations	16673	16754
elected× 1–2 seats	0.0111	0.0939***
	(0.0177)	(0.0273)
elected× 3–4 seats	-0.00926	0.0391
	(0.0181)	(0.0280)
elected× 5–9 seats	0.0179	0.0666**
	(0.0182)	(0.0297)
elected× 10+ seats	-0.00583	0.0265
	(0.0173)	(0.0290)
Observations	16673	16754
elected× small council	0.00582	0.0946***
	(0.0188)	(0.0292)
elected× small-medium council	-0.00216	0.0459**
	(0.0131)	(0.0201)
elected× medium-large council	-0.0261	0.0958**
	(0.0294)	(0.0428)
elected× large council	0.0386**	0.0292
	(0.0171)	(0.0311)
Observations	16673	16754

Note: The table reports effects of being elected into a municipal council on disposable income measured as logs of three-year averages 1–3 years after the election (column 1) and on the probability of being elected in the first subsequent election to a municipal council (column 2). All regressions include individual controls (see Table 2). Council size is defined as follows (the number of seats is always odd): small councils have at most 39 seats, small-medium councils have 41–49 seats, medium-small councils have 51–59 seats and large councils have 61 or more seats. Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Table 14: Short-run effects on disposable income and reelection probabilities; across individuals

	Disposable income	Local reelection
elected× < high school	0.0166	0.0357
	(0.0255)	(0.0357)
elected× < 2 years university	0.00269	0.0455**
	(0.0127)	(0.0211)
elected× ≥ 2 years university	-0.000976	0.0792***
	(0.0149)	(0.0244)
Observations	16633	16688
elected× age 18–39	0.0172	0.0722**
	(0.0187)	(0.0298)
elected× age 40–49	0.000634	0.00809
	(0.0167)	(0.0274)
elected× age 50–59	0.0156	0.0897***
	(0.0160)	(0.0285)
elected× age 60+	-0.0265	0.0661**
	(0.0200)	(0.0274)
Observations	16673	16754
elected× no prev. experience	-0.00148	0.0458
	(0.0211)	(0.0309)
elected× with prev. experience	0.00844	0.0668***
	(0.0136)	(0.0223)
Observations	10990	11044
elected× female	0.00199	0.0477**
	(0.0132)	(0.0225)
elected× male	0.00375	0.0661***
	(0.0120)	(0.0181)
Observations	16673	16754

Note: The table reports effects of being elected into a municipal council on disposable income measured as logs of three-year averages 1–3 years after the election (column 1) and on the probability of being elected in the first subsequent election to a municipal council (column 2). All regressions include individual controls (see Table 2). Educations is defined as the highest completed level in year t. Previous experience is an indicator for whether the candidate ran in any of the previous two local elections. Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

ucation, age, previous political experience and gender.[45]

The three tables contain in total 24×2 estimates. From these, it is clear that there is no income effect of being elected. The estimates are overall close to zero, and only one is statistically significant. In contrast, there are overall positive, significant reelection effects of around 5–10 percentage points. Interestingly, in the one case where there is a significant income effect (of around 3%), there is no reelection effect.

As discussed above, the paper cannot directly study what motivates politicians. But the analysis provides quite strong evidence that there are no monetary returns from local politics. The consistent absence of monetary returns across the various subgroups suggests that there are little reasons for politicians to even hope for significant income gains. Despite this, those elected stay in office quite some time. This may be interpreted as indirect evidence that they are not primarily motivated by monetary returns.

7 External validity

The RD design by construction estimates effects for marginal candidates. A potential concern is that these are so marginal so that the results do not hold on average. It is possible that there are indeed monetary returns for the average local politician. This section provides evidence against this.

A first thing to note is that borderline candidates in small parties with only one or two seats are more like the average politician than a marginal candidate. And the heterogeneity analysis showed that the result is the same in these small parties; there is no income effect but there is an effect on reelection probabilities (see Table 13).

To get a further sense of the external validity, I complement the RD design with an analysis on others than the borderline elected. This alternative identification strategy is a difference-in-difference model where *all elected* candidates constitute the treatment group. A set of non-elected candidates constitute the control group. In particular, for parties with n elected candidates, the top n^{th} ranked non-elected candidates are included as controls. The following regression is estimated:

$$Y_{i,t+j} - Y_{i,t-1} = \beta_0 + \beta_1 elected_{i,t} (+\mathbf{\Gamma}'\mathbf{X}_{i,t-1}) + \varepsilon_{i,t+j}, \tag{3}$$

where $Y_{i,t+j}$ is log disposable income for candidate i measured as the average

[45]The heterogeneous effects are estimated as interactions of the treatment variable $elected_{i,g,t}$ in the main regression specification (1). Controls for level effects are included. When studying heterogeneity across individuals, borderline group fixed effects are included to account for the fact that individual characteristics may be correlated with group characteristics. Each panel in the three tables corresponds to a separate regression.

over 1–3, 6–8 and 13–15 years after the election in year t. $Y_{i,t-1}$ is the previous three-year average, i.e. over years $t-4$–$t-1$. As in the RD model, the term in parenthesis denotes a vector of characteristics that are added to check the robustness of the results.[46] β_0 captures the overall change over time, whereas β_1 is the DiD estimate capturing the additional change among the elected.[47]

Table 15 presents the resulting DiD estimates. In contrast to the RD estimates, these tend to be statistically significant. Note though, that the number of observations is much higher. In terms of magnitude, the results suggest an income effect of around 2%. Although statistically significant, this does not imply any substantial monetary gains. In fact, some of the RD estimates are of the same size.

Given that the DiD is a very different type of specification as compared to the RD, it is interesting that the results are so similar. In particular, the DiD treatment effects in Table 15 refer to all elected candidates. The similarity of the results suggests that the local treatment effect identified in the RD design not only applies to marginal, borderline elected.

Having said that, entering Swedish politics *can* probably be lucrative for *some*. Indeed according to Berg (2020), there are substantial income effects of being elected into the Swedish national parliament. But this is so rare, so that it arguably would be unrealistic to hope for and be motivated by that.

8 Concluding remarks

The paper estimates causal effects of being elected in a local election on monetary returns. The claim for causality is made in a research design where the income of some candidate who just barely won a seat is compared to that of some other candidate who was close to winning a seat *for the same party*, but ultimately did not. The paper is one of few to focus on within-party discontinuities rather than between, and unique in the sense that it exploits the discrete candidate ranking.

The analysis establishes that monetary returns from local politics are absent both in the short and long run. This holds for different income measures such as disposable income, labor income and income from largest source. It is also true on average as well as when considering heterogeneous effects across various dimensions of parties, councils and candidates.

This null result may lead one to conclude that monetary returns are not what motivates some people to engage in local politics. On the other hand, it is possible that the motivation is the mere hope of positive returns, and that

[46]Unlike in the RD regressions, controls for income quantile in $t-1$ is not included in the DiD regressions, since that effectively would make a triple-difference specification.

[47]Reelection probabilities are not estimated with DiD model. The reason is that $Y_{i,t-1}$ (whether or not the candidate was elected also in the previous in election) is highly correlated with both Y_{t+j} and the treatment variable $elected_{i,t}$.

Table 15: Effects of being elected on disposable income estimated with a DiD model

	Period (t+1)–(t+3)		Period (t+6)–(t+8)		Period (t+13)–(t+15)	
	(1)	(2)	(3)	(4)	(5)	(6)
elected	0.0241***	0.0131***	0.0194***	0.0203***	-0.00378	0.0130**
	(0.00200)	(0.00223)	(0.00389)	(0.00610)	(0.00698)	(0.00604)
Observations	76714	76714	50308	24984	24108	24108
Individual controls	no	yes	no	yes	no	yes

Note: The table reports effects of being elected into a municipal council on income from largest source measured as three-year averages 1–3 (columns 1–2), 6–8 (columns 3–4) and 13–15 (columns 5–6) years after the election estimated with a DiD model. Individual controls measured one year before the election are: number of children aged below 18 and a set of dummies for age, gender, marital status, highest completed education, foreign background and past political experience. Standard errors clustered on municipality are in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

the absence of returns will be a disappointment. The paper cannot directly study politicians' motivation. But by relating the null result on income to positive effects on the future political career, the paper provides indirect, suggestive evidence that local politicians are not primarily motivated by money. And indeed, it can be argued that local politics is the relevant context when thinking about the question of what motivates politics. That being said, it is possible that returns to local politics and local politicians' motivations may be different in other countries.

Another aspect of external validity is the potential worry with the RD design, that returns to office are smaller for borderline elected candidates than for elected candidates further up the ranking. This does however not seem to be the case; the absence of any substantial income effects pertain also for the average politician. Yet, Swedish politics *can* probably be lucrative for *some*. Indeed according to Berg (2020), there are substantial income effects of being elected into the Swedish national parliament. But this is so rare, so that it arguably would be unrealistic to hope for and be motivated by that. Thus in conclusion, according to the notion that 'political service is a calling and that money is a distraction' (Besley, 2004), this paper delivers good news.

References

Bäck, M. and T. Möller (2003): *Partier och organisationer*, Stockholm: Norstedts Juridik, 6 ed.

Berg, H. (2020): "Politicians' Payments in a Proportional Party System," *European Economic Review*, forthcoming.

Besley, T. (2004): "Joseph Schumpeter Lecture: Paying politicians: Theory and evidence," *Journal of the European Economic Association*, 2, 193–215.

Besley, T., O. Folke, T. Persson, and J. Rickne (2017): "Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden," *American Economic Review*, 107, 2204–42.

Caselli, F. and M. Morelli (2004): "Bad politicians," *Journal of Public Economics*, 88, 759–782.

Cirone, A., G. Cox, and J. Fiva (2020): "Seniority-based nominations and political careers," *American Political Science Review*, forthcoming.

Dahlgaard, J. O. (2016): "You just made it: Individual incumbency advantage under Proportional Representation," *Electoral Studies*, 44, 319–328.

Dal Bó, E., F. Finan, O. Folke, T. Persson, and J. Rickne (2017): "Who becomes a politician?" *The Quarterly Journal of Economics*, 132, 1877–1914.

Deke, J. and L. Dragoset (2012): "Statistical power for regression discontinuity designs in education: Empirical estimates of design effects relative to randomized controlled trials," Working paper, Mathematica Policy Research.

Diermeier, D., M. Keane, and A. Merlo (2005): "A political economy model of congressional careers," *American Economic Review*, 95, 347–373.

Downs, A. (1957): *An economic theory of democracy*, New York: Harper and Row.

Eggers, A. and J. Hainmueller (2009): "MPs for sale? Returns to office in postwar British politics," *American Political Science Review*, 103, 513–533.

Eggers, A. and A. Spirling (2017): "Incumbency effects and the strength of party preferences: Evidence from multiparty elections in the United Kingdom," *The Journal of Politics*, 79, 903–920.

Ferraz, C. and F. Finan (2009): "Motivating politicians: The impacts of monetary incentives on quality and performance," Working Paper 14906, NBER.

Fisman, R., F. Schulz, and V. Vig (2014): "The private returns to public office," *Journal of Political Economy*, 122, 806–862.

Fiva, J. and H. Røhr (2018): "Climbing the ranks: Incumbency effects in party-list systems," *European Economic Review*, 101, 142–156.

Fiva, J. and D. Smith (2018): "Political dynasties and the incumbency advantage in party-centered environments," *American Political Science Review*, 112, 706–712.

Folke, O. (2014): "Shades of brown and green: Party effects in proportional election systems," *Journal of the European Economic Association*, 12, 1361–1395.

Folke, O., T. Persson, and J. Rickne (2016): "The primary effect: Preference votes and political promotions," *American Political Science Review*, 110, 559–578.

——— (2017): "Dynastic political rents? Economic benefits to relatives of top politicians," *The Economic Journal*, 127, 495–517.

Folke, O. and J. Rickne (2018): "All the single ladies: Job promotions and the durability of marriage," Uppsala University, mimeo.

Gagliarducci, S. and T. Nannicini (2013): "Do better paid politicians perform better? Disentangling incentives from selection," *Journal of the European Economic Association*, 11, 369–398.

Gagliarducci, S. and M. D. Paserman (2012): "Gender interactions within hierarchies: Evidence from the political arena," *The Review of Economic Studies*, 79, 1021–1052.

Golden, M. and L. Picci (2015): "Incumbency effects under proportional representation: Leaders and backbenchers in the postwar Italian chamber of deputies," *Legislative Studies Quarterly*, 40, 509–538.

Hagevi, M. (2000): "Professionalisering och deltagande i den lokala representativa demokratin. En analys av kommunala förtroendeuppdrag 1999," Report 13, CEFOS.

Hahn, J., P. Todd, and W. Van der Klaauw (2001): "Identification and estimation of treatment effects with a regression-discontinuity design," *Econometrica*, 69, 201–209.

Hyytinen, A., J. Meriläinen, T. Saarimaa, O. Toivanen, and J. Tukiainen (2018): "When does regression discontinuity design work? Evidence from random election outcomes," *Quantitative Economics*, 9, 1019–1051.

Kotakorpi, K. and P. Poutvaara (2010): "Pay for politicians and candidate selection: An empirical analysis," *Journal of Public Economics*, 95, 877–885.

Kotakorpi, K., P. Poutvaara, and M. Terviö (2017): "Returns to office in national and local politics: A bootstrap method and evidence from Finland," *the Journal of Law, Economics, & Organization*, 33, 413–442.

Lee, D. and D. Card (2008): "Regression discontinuity inference with specification error," *Journal of Econometrics*, 142, 655–674.

Lee, D., E. Moretti, and M. Butler (2004): "Do voters affect or elect policies? Evidence from the US House," *Quarterly Journal of Economics*, 119, 807–859.

Lundqvist, H. (2011): "Empirical essays in political and public economics," Ph.D. thesis, Department of Economics, Uppsala University.

Mattozzi, A. and A. Merlo (2008): "Political careers or career politicians?" *Journal of Public Economics*, 92, 597–608.

Messner, M. and M. Polborn (2004): "Paying politicians," *Journal of Public Economics*, 88, 2423–2445.

Öhrvall, R. (2004): *Hel- och deltidsarvoderade förtroendevalda*, Stockholm: Statistics Sweden.

Öhrvall, R. and J. Persson (2008): *Elected representatives in municipalities and county councils 2007: A report on the number of politicians and representativity (English summary)*, Stockholm: Statistics Sweden.

Querubin, P. and J. M. Snyder (2013): "The Control of Politicians in Normal Times and Times of Crisis: Wealth Accumulation by US Congressmen, 1850–1880," *Quarterly Journal of Political Science*, 8, 409–450.

Sekhon, J. and R. Titiunik (2017): "On interpreting the regression discontinuity design as a local experiment," *Advances in Econometrics*, 38, 1–28.

Swedish Ministry of Integration and Equality (2001): "Att vara med på riktigt—demokratiutveckling i kommuner och landsting," Swedish Government Official Report 2001:48.

Thistlethwaite, D. and D. Campbell (1960): "Regression-discontinuity analysis: An alternative to the ex post facto experiment," *Journal of Educational Psychology*, 51, 309–317.