Do barriers to candidacy reduce political competition? Evidence from a bachelor's degree requirement for legislators in Pakistan

September 2013

Madiha Afzal*

Abstract

In the 2002 election, candidates for Pakistan's federal legislature had to possess at least a bachelor's degree. This policy disqualified 60 out of the 207 incumbent legislators from running for election again. Using a difference-in-differences approach with panel data on all electoral constituencies in Pakistan, I find that this ballot access restriction does not affect political competition across all constituencies with disqualified incumbents equally. Stronger political competition is defined as a larger number of candidates contesting election, a smaller vote share and vote margin for the winning candidate, and a less concentrated candidate field, as measured by a Herfindahl-Hirschman index (HHI) of vote shares. Competition declined significantly in constituencies where the disqualified incumbent belonged to a small party and where literacy levels were lower (signifying a smaller pool of substitute candidates). However, political competition increased in areas where the disqualified incumbent was stronger in terms of his winning vote margin.

*Assistant Professor, University of Maryland, School of Public Policy. Email: mafzal@umd.edu.

1 Introduction

Before the October 2002 general elections, Pakistan's (then) President, General Musharraf, invoked a constitutional amendment that mandated a minimum education requirement on all candidates running for the office of Member of National Assembly (the lower house of Parliament in Pakistan).[1] In order to contest elections in 2002, all candidates for the office of Member of National Assembly (MNA) had to file proof of graduating with a bachelor's degree or higher with their candidacy papers.[2] This policy change had the effect of disqualifying 29%, or 60, of the 207 MNAs elected in 1997 from running again.[3] It restricted 97% of the country from running for Parliament, leaving only 3% of the population who were college graduates as eligible to contest national elections.

This paper examines how this unprecedented ballot access restriction affected political (or electoral) competition,[4] using four measures of competition: winning vote margin and share, a Herfindahl-Hirschman index of political competition, and the number of candidates standing for election. To my knowledge, no country other than Pakistan has ever imposed an education requirement for legislators. Most democratic countries have simple age and citizenship requirements for political candidates, along with the condition that candidates have no criminal convictions. Filing fees (or deposit requirements in some countries, where the deposit is forfeited unless the candidate gets a certain minimum number of votes) and petition requirements (also called signature requirements, in which a minimum number of signatures is required to nominate the candidate to run) are the typical ballot access laws around the world as well as in the United States.[5]

I quantify the impact of incumbent disqualification resulting from the policy change

[1] Pakistan has a parliamentary system of government. Its political system is described in detail in the next section.

[2] In Pakistan, a bachelor's has generally been a 14-year degree, even though both private and public colleges and universities have increasingly introduced 15- or 16-year degrees in order to make them compatible with foreign degrees.

[3] The 1997 National Assembly was suspended until 2002 without replacement when Musharraf took power in a coup in October 1999.

[4] Political competition and electoral competition will henceforth be treated as interchangeable terms.

[5] Voters must also meet minimum age and citizenship qualifications. In addition, certain requirements for voter registration, such as IDs, as well as the location of registration sites, may de facto disqualify many voters to whom these are not easily accessible. Besley, Persson, and Sturm (2010) show that the reduction in voting restrictions such as poll taxes and literacy tests in the US South (through the Voting Rights Act of 1964) increased political competition.

on a comprehensive array of measures of electoral competition: the number of candidates contesting election, the vote share and the vote margin of the winning candidate, and a Herfindahl-Hirschman political competition index. I also examine the effects of important variations across constituencies wherein incumbents were disqualified in terms of literacy rates, the size of the incumbent's party and the incumbent's electoral prowess.

This paper provides the first major evidence of the effects of a ballot access restriction not only in a developing country, but also one with an unstable and flawed democracy. Pakistan has had a democratically elected parliament but a military ruler as the head of state since 1999, who imposed a number of constitutional changes, including this.

The empirical strategy employed in this paper is essentially a generalized difference-in-differences approach. I use data on all electoral constituencies in the 1997 and 2002 elections, the closest elections before and after the policy change. As mentioned earlier, I measure the effect of disqualification of the incumbent from being able to run for re-election on a number of political competition outcomes, and my empirical strategy allows me to control for year and constituency fixed effects. Looking at disqualification of the incumbent identifies the constituencies directly hit by the education requirement, i.e., those constituencies represented by a legislator not having earned a bachelor's degree.

One would expect this policy change to lead to a reduction in political or electoral competition, as in a standard barriers-to-entry argument. The parallel with industrial organization theory has also been drawn previously (Drometer and Rincke 2009; Stratmann 2005; Ansolabehere and Gerber 1996). However, in Pakistan, one could expect the opposite result as well. If this policy change disqualified strong but uneducated incumbent legislators whose presence had traditionally discouraged entry of new but electorally weaker educated candidates, political competition could increase with the influx of educated candidates into the political arena. Indeed, I see both of these effects manifested, depending on the underlying constituency characteristics.

While the overall effect of incumbent disqualification on political competition is statistically insignificant, there is evidence of important heterogeneity in the effect of disqualification on political competition. There was a significant decline in political competition in those areas where the incumbent was disqualified and it was harder to find a substitute candidate

(I call these vulnerable constituencies). I measure this substitutability by examining literacy rates in the constituency and political party sizes, reasoning that less literate constituencies and smaller political parties find it harder to field substitutes for disqualified candidates. On the other hand, I find that in constituencies where a stronger incumbent (as measured by a higher initial vote margin) was disqualified, political competition increased. The central finding of this paper is that the policy change weakened political competition in vulnerable constituencies affected by incumbent disqualification, but strengthened it where the disqualified incumbent had been strong.

This research is related directly to three papers that examine the impacts of US ballot access laws, namely filing fees and petition requirements, on certain dimensions of political competition, as well as to a larger literature on entry barriers in politics (Tullock 1965). Ballot access restrictions vary across US states and over time. Ansolabehere and Gerber (1996) examine the effects of these laws on the probability of an uncontested seat, the frequency of legislator retirements, and vote shares of congressional election winners. The authors find that more stringent ballot qualification rules raise the probability of an uncontested seat and decrease the frequency of legislator retirements. Stratmann (2005) focuses on state lower house elections, and studies the impact of filing fees and signature requirements on the number of candidates running. The paper concludes that filing fees reduce the number of both major- and minor-party candidates, the latter in particular, while signature requirements reduce the number of major-party candidates. Drometer and Rincke (2008) argue that changes in ballot access restrictions across US states and over time are a response by the incumbent politicians to changes in electoral competition. Therefore, such legal restrictions cannot be treated as exogenous. Given this, in a follow-up paper, the same authors (Drometer and Rincke 2009) use the natural experiment created by a Supreme Court ruling in Ohio to examine the effect of lowering ballot access restrictions, and find a resulting increase in the number of third party and independent candidates.

A fourth related paper, Linden (2005), looks at the effect of ballot access restrictions in India, an electoral context that bears similarities to Pakistan (minus the military interventions). He examines the larger deposit and stricter nomination requirements for candidates that were instituted after the 1996 elections in order to discourage frivolous candidates, and

documents a resulting decline in the number of candidates. The related literature, then, finds that barriers to candidacy of various forms have resulted in reductions in certain dimensions of competition, whereas I find that this ballot access restriction in Pakistan strengthened political competition in some constituencies affected by incumbent disqualification, and weakened it in others. This paper also contributes to a broader literature that studies the impact of political institutions on electoral competition (Rae 1971; Lipjhart 1990).

The remainder of this paper is organized as follows. I first discuss Pakistan's political context and the law. Following that, I describe the data and provide basic summary statistics, and then present the econometric analysis and discuss the empirical results.

2 Pakistan's political context and the constitutional amendment

2.1 Political context

Pakistan has a parliamentary system of government. The national parliament consists of the Senate (upper house), whose 87 members are elected indirectly by the four provincial assemblies, and the National Assembly (lower house), 272 of whose 342 members are elected directly from single seat constituencies using plurality rule. Of the other 70 seats, 60 are reserved for women and 10 for minorities, but women and minorities can run for election from the 272 general seats as well.[6,7]

Each party can field one candidate per constituency, and various independent candidates (not aligned with any party) contest elections as well. Candidates can run for election, and win, from more than one constituency. In practice, there are a few strong candidates in every election who win from multiple constituencies (such as major party leaders including former Prime Minister Benazir Bhutto and current Prime Minister Nawaz Sharif). When that occurs, the candidate must resign from all winning seats but one, and by-elections are

[6] These numbers are for 2002, after the delimitation of constituencies in response to the population census of 1998. For the 1988-1997 elections, the National Assembly consisted of 207 general seats. I fully account for this redistricting, as explained in detail later.

[7] I do not study reserved seats in my analysis given that they are not based on direct election but on nominations.

held in the constituencies with newly vacant seats.

Pakistani politics is vigorously competitive. The average number of candidates running for election from a single constituency was 8.07 in the 1997 election and 7.54 in the 2002 election (see Table 1 and the Data and Variables section for further discussion). The average candidate won with a vote margin of 26% in the 1997 election and 17% in the 2002 election. Nine parties were represented in the National Assembly in the 1997 elections, and seventeen in the 2002 elections (a number of these were splinters of the nine parties represented earlier).

2.2 The constitutional amendment

Musharraf's Chief Executive Order No.7/2002 mandated that in the 2002 election, all candidates for the federal legislature, the National Assembly, had to have a bachelor's degree. This was one of a number of constitutional amendments issued by executive order in the run-up to the election.[8] It is not the purpose of this paper to evaluate the rest of these constitutional changes, but I will discuss their relation, if any, to my analysis of the effects of the education requirement. Among these were reservations of parliamentary seats for women and minorities and a lowering of the voting age from 21 to 18 years.

The executive order generated a great deal of controversy. Opposition parties and human rights advocates argued that the requirement was undemocratic and exclusionary (Haven 2002). It disqualified not only many new potential entrants into politics, but also prevented some experienced politicians from running for reelection. The law certainly ended any notion of equal access to the political process based on educational attainment. To the extent that more highly educated people tend to belong to wealthier families, it excluded the poor.

[8] The executive order was abolished by a seven-member bench of the Supreme Court on April 21, 2008, following Musharraf's fall from power, in response to a petition filed by two members of the Jamiat Ulema-i-Islam Fazl-ur-Rehman Group (JUI (F)) political party, on account of inconsistency with articles 17 (freedom of association) and 25 (equality of citizens) of the Constitution. In fact, previous drafts of this paper had been written before the law was struck down by the Supreme Court, and the statistic I cited (that this policy prevented 97% of the population from contesting elections) was mentioned as a reason for the dissolution of the requirement.

The February 2008 elections were held with the education requirement in place, therefore the legislators elected that year all possess bachelor's degrees as well; only for those contesting by-elections did the law no longer hold. In the May 2013 election, the law no longer applied, and therefore the candidates and legislators elected this year are not required to have a bachelor's degree. In the future, it would be useful to study the effect of going from a more educated Parliament elected in 2008 to one elected without this requirement in 2013.

In addition, it is likely that women, minorities, and potential candidates from rural areas were affected adversely (Haven 2002). Critics of the law also pointed out anecdotally that relatives replaced many of the disqualified incumbents in Parliament, acting as 'puppets' for the disqualified parliamentarians.

Proponents of the bachelor's degree requirement argued that the policy would improve political selection, and that better educated politicians tend to be less prone to corruption. Musharraf also declared that the law would ensure a more efficient and less corrupt legislature, and his Law Minister claimed that only the powerful opposed it. Interestingly, smaller parties also spoke out in favor of the law since they felt that as some major party political heavyweights were disqualified, it would improve their candidates' chances (Bhatty 2002). However, as my analysis will show, small parties were adversely affected by the requirement.

Under the law, equivalency of non-standard degrees had to be determined by the University Grants Commission of Pakistan (which has since been renamed the Higher Education Commission). Islamic degrees were generally granted equivalence to a bachelor's degree by the commission, a controversial practice. Given this, opponents of the education requirement argued that it was unfairly beneficial to religious parties. The equivalence of Islamic degrees to a B.A. was contested in the Supreme Court by a few lawyers, but the case has not been resolved. The petitioners argued that equivalence had been granted only to the extent that those possessing religious degrees could teach religious courses, and not for other purposes (Rehman 2006). Critics argued that this put the newly formed alliance of religious parties, the Muttahidda Majlis-e-Amal or MMA, at an advantage. In fact, the MMA won 45 out of the 272 seats in the National Assembly in 2002, a majority of which were in the North West Frontier Province (NWFP), where it also went on to form the provincial government. In contrast, only one of its constituent parties won two out of 207 seats in the National Assembly in 1997. The highest degree held by 40% of the MMA legislators elected in 2002 (18 out of 45) was a religious one, compared to only 4.41% of non-MMA legislators (10 out of 227). In separate work, I show that the MMA did, in fact, benefit directly from the education requirement, but only in the NWFP.[9]

[9] This is at least partly attributable to the anti-American stance of the MMA which resonated within the NWFP, the province which borders Afghanistan, in the wake of the US war there.

What about disqualified candidates acquiring fake degrees, or even possibly a real degree, in the time between the announcement of the law and the filing of candidacy papers? The latter was essentially impossible for the 2002 election, and the former difficult, given the timing of events; the education law was announced on June 24, 2002, and the elections were held in October of that year. However, there were some allegations of fake degrees after the 2002 election.[10] Using data on the educational attainments of all MNAs in 2002 and the intuition that there is little incentive to acquire a fake degree higher than the minimum required, that is, a bachelor's, I ran a partial test for the prevalence of fake degrees. I compared the average number of years of education for constituencies where the incumbent was disqualified relative to those where the incumbent was not disqualified: the averages are similar and appear higher than a bachelor's degree, supporting the hypothesis that there were not many fake degrees in 2002.[11] In any case, fake degrees do not really affect my analysis, since the education requirement still basically functioned as a ballot access restriction even if candidates could acquire fake degrees: after all, even getting a fake degree functions as a barrier to entry.

2.3 Delimitation of constituencies (redistricting)

Pakistan had 207 national electoral districts, referred to as constituencies, for the four elections held between 1988 and 1997. The 1998 national population census, the first in 17 years, necessitated a delimitation, or redistricting, of these constituencies in order to account for the large increases in population which took place in the intervening period. Therefore, in 2002, the Election Commission of Pakistan (ECP) drew new constituency boundaries and expanded the number of parliamentary seats to 272 in accordance with the Delimitation of Constituencies Act of 1974. Maps of the 1997 and 2002 constituencies, by province, are in Online Appendix A.[12]

In order to use panel data, and at an even more fundamental level, in order to determine

[10] However, more allegations surfaced after the 2008 elections, a number of which were later substantiated, showing that given more time, the acquisition of fake degrees turned out to be more of a problem.

[11] The results are available on request.

[12] All appendices are available online at this website: `https://sites.google.com/site/madihaafzal/research/PublicChoiceAppendix.pdf?attredirects=0`

which constituencies in 2002 were affected by disqualified incumbents, I matched the 1997 constituencies to the 2002 constituencies using Geographic Information Systems (GIS) as well as a population-weighted matching method.[13] Only 13 constituencies were unaffected by redistricting. Since the number of constituencies increased as a result of redistricting, constituencies were typically divided into two (or sometimes into three). My population-weighted matching method, described in detail in Online Appendix B, creates 207 virtual constituencies in 2002 using the electoral data from the 272 actual constituencies in 2002, yielding a panel of 207 constituencies for the regression analysis.

3 Data and variables

3.1 Political or electoral competition: outcome measures

The political competition variables are constructed using data from the Election Commission of Pakistan, which publishes the number of votes received by each contesting candidate, along with his or her party affiliation, in each constituency after every general election. I also separately obtained the population of each constituency in 1997 and 2002 from the Election Commission. The following four variables are defined to represent various dimensions of political, or equivalently electoral, competition:

1) The number of *candidates* contesting election, which reflects the pool of candidates and therefore the extent of choice available to voters. A larger number of candidates signifies more vigorous political competition.

2) The *vote share* of the winning candidate, which is equal to the votes received by the winning candidate divided by the total votes polled in the constituency. This refers to the absolute strength of the winning candidate. A smaller vote share signifies stronger political competition.

[13] The Election Commission does not have the maps in GIS formats required for this matching. Therefore, I obtained the maps in graphic formats from the Election Commission for 2002. For 1997, I got the maps from a private consulting company, ECIL, which had been hired as a consultant firm by the ECP. I then manually constructed the shape files for both the 1997 and the 2002 maps in GIS by first georeferencing the maps to have the same projection, and then digitizing the maps to have constituency boundaries (by drawing polygon lines through the maps). I used GIS to intersect the 1997 and 2002 boundaries, and then calculated the exact area weights of these intersected areas relative to the 1997 constituency and 2002 constituency to which they belonged. Details of the construction steps in GIS are available upon request.

3) The *vote margin* of the winning candidate, equal to the votes received by the winning candidate minus the votes received by the first runner-up candidate, divided by the total votes polled in the constituency. This reflects the closeness of the election and the strength of the winning candidate relative to the first runner up. A smaller vote margin signifies more political competition.

4) The *Herfindahl-Hirschman political competition* index, defined as one minus the political Herfindahl index, which equals the sum of squares of the vote shares of each candidate running for election in a constituency, or $1 - \sum VS_i^2$, where VS_i = vote share of candidate i. This incorporates all the candidates standing for election and their strength. If only one candidate runs and wins, the Herfindahl-Hirschman index is 0. The larger the number of candidates running and the smaller their relative vote shares to each other, the closer the Herfindahl-Hirschman index will be to 1. A higher HHI signifies stronger political competition.

I use four measures of various dimensions of political competition since one index alone cannot tell us whether a constituency is competitive owing to many candidates on the ballot or because of close races between fewer candidates. The four measures are related to each other. Vote share and vote margin often run hand in hand, since vote margin is equal to the winner's vote share minus the runner up's vote share. The number of candidates running and the vote share and vote margin measures are not necessarily related, since candidates can run but get no votes. However, it is true that with many candidates running and all else equal, vote shares for the winner will be smaller if all of the candidates capture at least some votes. Finally, since the Herfindahl-Hirschman index incorporates both the number of candidates running for election and each candidate's vote share, it incorporates the other three measures: the higher the vote share and the vote margin, the lower the index, and the larger the number of candidates, the higher the HHI. The Herfindahl-Hirschman measure goes beyond the three measures combined, since it incorporates each candidate's vote share, not just the winner's and the runner up's shares. In examining all four dimensions of political competition as outcome variables, I check to see whether disqualification affects the number of candidates in the playing field, or the closeness of the race, or both.

Table 1 presents summary statistics for electoral outcomes for the 1997 and 2002 elec-

tions, as well as for the matched 'virtual' 2002 constituencies. First, note that constituency 28 (in the Tribal Areas), which had 107 candidates contesting election in 1997, is an outlier.[14] I have dropped constituency 28 in all of my empirical analyses.[15] Analyzing the unmatched data, we can see that the average number of candidates running in 2002 was 7.54, a slight reduction from 8.07 candidates in 1997. Similarly, the average of the vote share of the winning candidate went down from 55% in 1997 to 49% in 2002. The mean of the vote margin of the winning candidate also decreased from 0.26 in 1997 to 0.17 in 2002. Finally, the mean of the Herfindahl-Hirschman political competition index increased from 0.57 to 0.62, although the distribution shifted towards zero as both the minimum and the maximum decreased. Overall, therefore, political competition seems to have weakened on some dimensions and strengthened on others between 1997 and 2002. Note that although the number of constituencies increased from 207 in 1997 to 272 in 2002 owing to Pakistan's growing population, the average size of the constituencies also increased from approximately 406,000 in 1997 to 489,000 in 2002.

Comparing the matched 2002 data to the actual election results from 2002, we can see that the matching does not substantially change the structure of the data (since we constructed population-weighted averages), which is reassuring.

3.2 Independent variables

3.2.1 Incumbent disqualification

To define the main independent variable, disqualification, I collected data on the educational attainments of the 207 MNAs elected in 1997. I assign the values of the variable **after** matching the 2002 constituencies to the 1997 constituencies, so the variable is binary. For the 1997 constituencies, incumbent disqualification equals 0 across all constituencies. For the matched 2002 constituencies, I define "disqualified" as a dummy set equal to 1 if the MNA elected from the corresponding constituency in 1997 did not have a Bachelor's degree or

[14] This constituency was an outlier in terms of candidates standing for election, but not in terms of other political competition outcomes. The winning candidate in that constituency had 29% of the total votes polled and his winning margin as a proportion of the total votes was 11%.

[15] The maximum number of candidates running across constituencies in 1997 after excluding constituency 28 was 35.

higher. Table 1 contains the summary measures for incumbent disqualification. As the table shows, 59 out of 198 (legislator education data is missing for nine constituencies in 1997) incumbent legislators were disqualified, while 139 were not. Note that I have information on the education levels of the 272 MNAs elected in 2002, but the incumbent disqualification variable is defined using the 1997 data alone. For the MNAs elected in 2002, the education information obtained was part of the application for candidacy filed with the Election Commission. Before that election, there were no public educational records for Pakistani legislators. Therefore, for each of the MNAs elected in 1997, I requested the education levels of all MNAs from the respective District Coordination Officers (DCOs). An officer from the DCO's office obtained this information either through phone, or by personally going to the (ex-)MNA's house.

While disqualification of the incumbent is not the only effect of the education requirement, this is the effect of the policy which I will measure, and my main independent variable in the analysis. Ideally, we would want to identify all potential candidates who wanted to run for election in 2002 but could not because of the policy change, but this is obviously impossible to do. A solution to this problem could have been to look at the 1762 candidates contesting election in 1997 and see which ones were disqualified in 2002, but since education levels of the candidates were not recorded before 2002, this information was collected only for the 207 winners in 1997. This was partly because of difficulties in collecting education information for 1555 more people, along with the fact that many former losing candidates were more difficult to track down than former legislators.

Sixty out of 207, or 29% of, MNAs elected in 1997 were disqualified in 2002. On average, over the 1988-1997 period, during which four elections were held, 68% of National Assembly incumbents ran for election again from the same constituency in the next election. Of these, on average, 34% of incumbents won the next election from the same constituency (38%, 34%, and 31% in the 1990, 1993, and 1997 elections respectively).[16] Of course, these statistics abstract from the fact that incumbents may choose to run from a different constituency, so the percentage rerunning or winning across constituencies would be higher if one looks across all constituencies and not just at the incumbent's own. Given the substantial numbers

[16]In terms of parties, 47% of parties were reelected from the same constituency.

of candidates running for reelection, combined with the fact that there were even more constituencies to run from in 2002 following redistricting, the disqualification of 29% of incumbents is a significant number and imposed a significant constraint. In Online Appendix C, I check for whether disqualified incumbents were replaced by relatives in 2002.

In 2002, of the 147 incumbents who were not disqualified, 23 were reelected from the same constituency (to be precise, from one of the constituencies the 1997 constituency split into post-redistricting). Of the 60 constituencies where the incumbent was disqualified, 27 elected a new candidate from the incumbent's party in 2002 (this number is based on the largest constituency which emerged out of the corresponding 1997 constituency post-redistricting),[17] suggesting that quite a few of these constituencies may have reelected the incumbent had he been able to run.

Constituencies where the incumbent was disqualified are significantly less urban than those where the incumbent was not disqualified (26.4% versus 33.6%) but not significantly less literate (41.5% versus 43.9%). As long as the urban population proportion stays relatively constant within constituencies over time, constituency fixed effects take care of this difference between disqualified versus non-disqualified constituencies. The only problem that might arise is if constituencies hit by incumbent disqualification change their urban/rural composition differently than constituencies represented by incumbents who were not disqualified, but it is hard to imagine why this would be the case. An examination of incumbents who were disqualified by party affiliation does not suggest any revealing patterns (Online Appendix Table D3[18]); indeed, as I show later when discussing small parties, there is no correlation between the size of the party and disqualification of its candidates.

Why should incumbent disqualification matter for political competition? We want to see whether new candidates enter the race or whether there is a reduction in the number of candidates with the incumbent forced off the ballot. In addition, does the winner's vote share and margin change given that the winner is definitely different this time? While we don't have direct evidence for this, there might also be a correlation between the fact that the incumbent was disqualified and the types of candidates in 1997: perhaps constituencies

[17] This number increases to 39 if we consider any of the constituencies the corresponding 1997 constituency split up into in 2002.

[18] Online Appendix Tables D1 and D2 list the party affiliations of MNAs elected in 1997 and 2002.

that had elected an uneducated incumbent were also more likely to have had candidates who were uneducated, since that could represent voter preferences or the pool of available candidates in the constituency. That would imply that we are partly capturing candidate disqualification in the incumbent disqualification variable as well.

3.2.2 Constituency level variables

District-level *literacy* data were obtained from the Federal Bureau of Statistics. We will examine whether the effect of incumbent disqualification differs by the literacy level of the constituency, which acts as a proxy for the pool of available educated candidates in the constituency in 2002. The literacy data are based on the 1998 census, and are matched to the 1997 constituencies. It signifies the percentage of the constituency population, aged 10 and older, which is literate. Because it was obtained for one year only, it does not change over time for each constituency. The average literacy rate across constituencies is 43.24% (see Table 2).

Party size is used as a proxy for the availability of alternatives for the disqualified incumbent, that is, a measure of how difficult it is to find a replacement for him. I define *small party* based on the number of candidates put up for election by that party in all of the national constituencies in Pakistan in 1997. Online Appendix Table D4 illustrates this definition. In the limit, the smallest party is an independent candidate. Focusing on the number of candidates fielded by the party instead of the winners reflects on the actual capabilities of the party rather than the election results which are determined endogenously. I define a party as small if it fields fewer than 50 candidates for election. Alternatively, this definition means that a party is small if it fields candidates in approximately less than one-fourth of the total National Assembly constituencies, which number 207 in 1997. This seems to be a natural cutoff in the data, as can be seen in the table, and the difference between a regional (small) and a national (large) party. Nineteen percent of incumbents elected in 1997 belonged to a small party using this definition (Table 2). However, I used three alternate cutoffs for small party to check for robustness as well - whether the number of candidates fielded is less than 10, 20, or 100. I also define a continuous party size measure by using the actual number of candidates fielded by each party in 1997 instead of an indicator variable

for small party size. The mean number of candidates put up for election by the incumbent's political party in 1997 was 136 candidates. Finally, I also consider an alternate small party measure, namely the number of candidates in the constituency who belong to a small party instead of using just the winner belonging to a small party. That number is the same for both election years for each constituency, and averaged at 5.11 (Table 2).

The final measure I use is the initial vote margin of the winner in the 1997 election, that is, the votes of the winning candidate minus the votes of the runner up, divided by total votes polled. It is meant to capture the strength of the incumbent elected in 1997, and is therefore the same for each constituency for both election years. The average winning vote margin in 1997 was 0.26 (Table 2).

4 Empirical analysis

4.1 Effect of disqualification

The empirical strategy employed in this paper is essentially a difference-in-difference approach. I use data on all electoral constituencies in the National Assembly for the elections before (1997) and after (2002) the education requirement. The coefficient of interest measures the effect of disqualification of the incumbent on a number of electoral competition outcomes, controlling for constituency population, and year and constituency fixed effects. Looking at disqualification of the incumbent identifies the constituencies hit by the education requirement, i.e., those constituencies that had elected a legislator without a degree prior to the policy change; these are the areas where the law has a binding effect. In particular, the initial empirical specification is (the coefficient of interest is γ):

$$P_{ct} = \alpha_c + \beta_t + \gamma D_{ct} + Pop_{ct} + \varepsilon_{ct},$$

where P_{ct} : Electoral outcome in constituency c at time t. As described earlier, I use four different dimensions of political competition: the *vote share* of the winning candidate; the *number of candidates*; the *vote margin* of the winning candidate as a proportion of total votes; and the *Herfindahl-Hirschman index of political competition*. The number of candidates measures the number of political players in the constituency whereas the vote

share and vote margin get at the strength of the major political players, and the political competition index incorporates both the number of candidates as well as each candidate's strength.

α_c : Constituency fixed effects.

β_t : Year effects. The *year* dummy in each regression captures anything specific about the 1997 or the 2002 elections that affects all constituencies equally.

D_{ct} : *Disqualified* is defined as a dummy equal to 1 for a constituency in 2002, if the legislator elected from that constituency in 1997 did not have a Bachelor's degree or higher; it equals 0 for all constituencies in 1997. It measures incumbent disqualification due to the education requirement.

Pop_{ct} : Constituency population for 1997 and 2002.

t : 1997, 2002 (election years).

Therefore, γ identifies the change in political competition due to the education requirement for those constituencies where the incumbent was disqualified, adjusting for any change in political competition that would occur without the requirement, which is identified using those constituencies wherein the incumbent was not disqualified.

Constituency-level time varying controls are not readily available in Pakistan. The only constituency-level information other than electoral results that is available is constituency population, which I include as an independent variable in the regression. Controlling for constituency population helps us tackle other electoral changes that might affect constituencies of different sizes differently, such as the change in voting age from 21 to 18. Data for Pakistan generally are more readily available at the administrative district level, which I can match to the constituency level. But such data (including literacy rates and urban/rural data) are available only for the year 1998, the census year, and therefore are time-invariant for the 1997-2002 time period. Given that I include constituency fixed effects, these automatically take care of all constituency-level time-invariant characteristics. I also include time dummies to capture any election year changes affecting all constituencies equally.

With this empirical approach, any level differences between constituencies wherein the incumbent was disqualified (where the law binds) and those where the incumbent was not disqualified (where the law does not bind) are taken care of by constituency fixed effects.

However, difference-in-differences does assume that in the absence of the policy intervention, political competition in the constituencies where the incumbent was disqualified would have grown at the same rate as in constituencies where the incumbent was not disqualified. I tested this indirectly before the policy intervention, by comparing growth rates for the political competition variables between 1993 and 1997 in constituencies hit by disqualification in 2002 with constituencies not hit by disqualification in 2002.[19] I cannot reject equality of means; therefore, there were no strong pre-existing differential trends in political competition between these two types of constituencies.[20]

4.1.1 Redistricting

The fact that redistricting of constituencies and the education requirement both occurred at the same time raises possible concerns that these changes were not independent of one another. This can be a problem since it could imply endogeneity of my independent variable, disqualification. To test for orthogonality of disqualification and delimitation, I created a variable that signifies the extent of redistricting in the constituency. Online Appendix B contains a detailed discussion of this variable. When this redistricting variable is greater than 1, it signifies that the constituency is split. In my data, the mean of this variable is 1.30, and the standard deviation is 1.00. Finally, the median is 1.18, signifying a right-skewed distribution.

I ran a simple correlation between incumbent disqualification and this redistricting variable. This yields a correlation coefficient of 0.01, and a p-value of 0.85, implying that these two variables are not correlated, thus validating my analysis. The redistricting variable is also not related to other correlates at the constituency level: the percentage of the constituency that is urban (the correlation coefficient is 0.04 and the p-value is 0.44) and the percentage that is literate (the correlation coefficient is -0.00 and the p-value is 0.95), which is reassuring.

This empirical strategy also assumes that any other changes, policy or otherwise, occur-

[19] Note that I look at political competition growth rates for the vote fraction, vote margin and candidate variables.

[20] The difference in means t-test for vote fraction growth has a value of: $|p|>t=0.19$; for number of candidates growth the value is: $|p|>t=0.39$; and for vote margin growth the value is: $|p|>t=0.16$.

ring at the same time as disqualification did not affect constituencies with disqualified and qualified incumbents differently. In Online Appendix E, I discuss, one by one, the other policy changes that occurred at the same time as this education requirement, and show that this is a reasonable assumption.

4.1.2 Results

The basic panel regression across all Pakistani constituencies (Table 3) shows an insignificant overall effect of disqualification on political or electoral competition. However, the signs of the coefficients on all political competition variables are consistent with reduced political competition in the constituencies where the incumbent was disqualified as a result of this policy change. In terms of magnitudes, these numbers are moderate. Vote share and vote winning margin increase by 4% to 5% as a proportion of total votes in constituencies where the incumbent was disqualified; this suggests that the absolute and the relative strength of the winner both increase in these constituencies. Similarly, the effect on the number of candidates is sizeable: exactly one less candidate runs from constituencies where the incumbent was disqualified. One explanation is that once the incumbent is disqualified and cannot stand for reelection, another candidate from his party may not take his place. It could also be that in such constituencies, other candidates may also have been disqualified and one of them on average is not replaced by another candidate from his party. Finally, the Herfindahl-Hirschman political competition index decreases by 0.04, a small amount.[21] The regressions all include constituency population, but the coefficients are not shown in this table or the following tables, since they are very small (0.00) and insignificant across all specifications.

[21] What factors can introduce a bias in these results? First, if the education of legislators in 1997 is misreported, this can bias my estimates downward, so that they place a lower bound on the true effect. Second, the GIS matching may not be perfect, since I constructed the GIS maps based on image files. However, this was the only way this analysis could be done, and it can only cause possible measurement error in the dependent political competition variables which were matched for 2002. It is important to clarify that this possible measurement error would be because of my GIS matching and not as a result of strategic redistricting.

4.2 Variations across constituencies with disqualified incumbents

Constituencies in which the incumbent was disqualified vary in important ways, and we expect the effect of disqualification to vary depending on the characteristics of the constituency. To account for this, the main empirical specification is:

$P_{ct} = \alpha_c + \beta_t + \gamma D_{ct} + \phi V_c + \delta D_{ct} * V_c + Pop_{ct} + \varepsilon_{ct}$, where the new variable, V_c, signifies constituency level variables. The variables V_c I examine are literacy, small party, and initial vote margin. The hypotheses and results for each of these are discussed in turn below. The variables V_c do not vary over time, so ϕ, the coefficient on V_c, cannot be identified separately from the constituency fixed effect α_c, and the estimated coefficients therefore capture some part of a constituency fixed effect. The interacted terms, $D_{ct} * V_c$, do vary over time, and the coefficient of interest, δ, can be interpreted as the change in political competition for a one unit change in V_c in constituencies where the incumbent is disqualified. Note that I do not argue that V_c or $D_{ct} * V_c$ is exogenous: in the case of small party and initial vote margin, it is likely not. The analysis below is meant to document empirically the differences in political competition in constituencies with disqualified incumbents which vary on certain dimensions rather than to argue for causal interpretations.

4.2.1 Literacy

I hypothesize that in areas where it is harder to find a substitute for the disqualified incumbent, political competition should decrease significantly. A constituency level variable that can shed light on substitutability is the percentage of the constituency that is literate. It would be ideal to use the percentage of the constituency that holds a bachelor's degree, but getting an accurate statistic on that is difficult. As long as the percentages of bachelor's degree holders and literate citizens within constituencies are positively correlated, literacy works as a proxy for the availability of alternative candidates. I hypothesize that when the incumbent is disqualified, fewer candidates will run in less literate constituencies, and through that effect, the HHI will also decrease in these constituencies.

Table 4 shows the results of the regression with the constituency literacy-incumbent disqualification interaction. In constituencies where the incumbent is disqualified, a 10%

reduction in the literate population is associated with a 0.07 increase in vote share, a 0.07 increase in the vote margin, a 0.97 decrease in the number of candidates running and a 0.06 decrease in political competition. Therefore more vulnerable areas, which had a less literate population and therefore fewer possible educated candidates, suffered a decline in political competition when their incumbent was disqualified. The coefficient on number of candidates is marginally insignificant, but the ones on the other political competition measures are significant. This is surprising, and it implies that the effect may not operate by reducing the number of candidates alone: one explanation is that the disqualification of an uneducated incumbent in less literate constituencies leads not only to a reduction in candidates running but much weaker slates of candidates, leading the new winner to win by a larger margin. These new educated politicians may not be as aware of the needs of the constituents in these particular constituencies, or may have policy preferences that are different from those of their voters since the legislator and voters have very different education levels, resulting in a welfare loss for the constituents. Understanding this further is important for future research.

4.2.2 Small party

An alternative electoral proxy for the availability of substitutes for disqualified incumbents is party size. The assumption underlying this proxy is that if the membership shares of educated candidates are the same in small and large parties, the former will by definition have fewer potential candidates to field; this will reduce political competition.[22] Small party and disqualification are not correlated, both according to the small party indicator and the continuous party size measure. Therefore, incumbents in 1997 who belonged to a small party were not more likely to be uneducated. Why does that make sense and how is this consistent with my assumption that the proportion of educated candidates in small and large parties is equal? Because while it implies that the winners in 1997 were of similar education levels across small and large parties (which is sensible), my assumption relates to what

[22] We could define a more direct measure of the availability of substitutes for disqualified incumbents than the size of the party but this would be prohibitive in terms of data requirements. I use small party as a proxy for a smaller pool of educated candidates in a party, but we could look for the actual pool of educated candidates in a party in lists of party members, and therefore use that measure for substitutability. However, this would require not only gathering together historical party lists, but also inquiring into the education levels of all the party members.

happens once a candidate has been disqualified: that is, while incumbents from small and large parties are equally likely to be disqualified, once they are disqualified, the assumption states that a small party is going to have a harder time finding an educated candidate to replace its incumbent relative to a large party. Constituencies that were represented by small party candidates in 1997 were significantly less urban than constituencies not so represented (21.97% versus 34.41%) and less literate (32.04% versus 45.62%) according to the small party indicator measure; and also significantly less literate according to the continuous party size measure. Small parties may be ethnic or regional parties, such as the Awami National Party, or parties with smaller memberships and supporting bases for other reasons.

The regression results (Table 5) confirm the substitutability hypothesis: constituencies wherein the 1997 winner was from a small party or was an independent candidate and was disqualified face a subsequent reduction in political competition, in terms of winning vote share and winning vote margin, as well as the number of candidates and the Herfindahl-Hirschman political competition index. These results are robust to using the three alternate cutoffs for small party.[23] This reveals that there was indeed a real effect of this education requirement on political competition: competition declined significantly in those areas where the incumbent was disqualified and it was harder to find a substitute candidate. Moreover, these numbers are large in terms of magnitude; they reveal that in those constituencies in which the incumbent was disqualified as a result of the education requirement and he belonged to a small party (relative to a large one), winning candidates secured 19% more of the total vote and won a 25% larger margin over the runner-up candidates. In addition, five fewer candidates ran for election, and the Herfindahl-Hirschman political competition index fell by 0.14.

As an additional robustness check, I define two alternate measures for small party: continuous party size and the total number of candidates in the constituency who belonged to smaller parties. The results (available on request) show that in constituencies where the incumbent was disqualified and belonged to a smaller party, electoral competition declined along all dimensions. In areas where the incumbent was disqualified and more of the can-

[23]The results are also robust to excluding independent candidates from the definition of small party, although a little weaker. They are available upon request.

didates in his constituency belonged to a small party, political competition also weakened significantly along all dimensions. The results are therefore robust to a number of alternate definitions for substitutability of a disqualified candidate: for the small party variable defined for the winner of each constituency in 1997, both as an indicator and as a continuous variable, as well as small party defined over the entire pool of candidates running from the constituency in 1997.

In order to confirm the assumptions underlying the analysis above, we can check whether a small party candidate who was disqualified is replaced by a candidate who runs (and wins) from the same small party. If he is not, that lends support to the assumption that smaller parties have fewer educated candidates available to replace their disqualified incumbents. To do this, I define two new dependent variables: an indicator for whether a candidate from the same party as the 1997 winner runs for election in 2002 (*incumbent party run)*, and another for whether or not a candidate from the same party wins in 2002 (*incumbent party win).*[24] I regress these on the independent variables, disqualification and the interaction of disqualification with small party, controlling for year and constituency fixed effects as before. The coefficient of interest is on the interaction of small party and disqualification, and it measures whether a disqualified incumbent from a small party was replaced with someone from his party as a candidate (and eventually as a winner), relative to a disqualified incumbent from a large party. A negative coefficient implies that a disqualified incumbent from a small party was harder to replace with another candidate from the same small party, relative to a disqualified incumbent from a large party being substituted for by another candidate from his own (large) party: the results in the first two columns of Table 6 thus confirm the assumption above that small parties have lesser supplies of educated candidates to put up for election. In terms of sizes, we see that where an incumbent from a small party was disqualified, the incumbent party was 50% less likely to run in the 2002 election (Table 6, column 1), and correspondingly, 35% less likely to win the 2002 election (column 2), than where an incumbent from a large party was disqualified. This effect is robust to the two

[24]To do this, I need to match parties from the 1997 constituencies to the 2002 constituencies but I cannot use the population-weighted matching method given that parties are a qualitative variable. I therefore use three alternative methods for matching the 1997 constituencies to the 2002 constituencies: the largest area overlap, any area overlap, and an overlap of a certain minimum area (0.38). The results presented here are for the largest area overlap, but those for the other two matching methods are very similar.

alternate definitions of small party, and remains negative and strongly significant.[25]

Is this effect confined to the small party of the disqualified incumbent or does it extend to other small parties as well? I examine this in Table 6, column 3. To do this analysis, I define a new variable called *small party win,* which is an indicator variable for whether or not the winner belongs to a small party, defined separately for 1997 and 2002.[26] Note that this is different from the small party variable used in the interaction term, which defines small party as the same for both years, based on the party of the winner in 1997. I find that small parties overall are also less likely to win in those constituencies where the incumbent in 1997 was from a small party and was disqualified, a result all the more important because those are the very constituencies that have preferences for small party candidates as revealed by their incumbent choices. In addition, this shows that in contrast to what their representatives thought (Bhatty 2002), small parties actually were weakened by the education requirement. In fact, Table 6, column 3 shows that not only are all small parties significantly less likely to win in constituencies where a small party incumbent was disqualified, they are also no more likely to win in constituencies where a large party incumbent was disqualified.

These results are important, because small parties play a significant role in Pakistan: they can help in forming the coalition government and can give real competition to the ruling party. We know that small parties won 18 out of 207 seats in 1997 and independent candidates won another 21. The literature also shows that small parties matter across different contexts, such as by making the major parties take more divergent policy positions (Adams and Merrill 2006). The results complement the evidence from ballot access restrictions in the United States which affect minor party candidates and independents in particular.

4.2.3 Initial vote margin

Third, I examine how the effect of incumbent disqualification varies by the political strength of the disqualified incumbent. The vote margin by which an incumbent was elected in 1997 can serve as a measure for his strength. I hypothesize that in constituencies where a stronger

[25] Results available upon request.

[26] That is, separately for each year, I define a party as small if it fields 50 or less candidates in that year, and then the small party win variable is an indicator variable for whether the winner from that constituency belongs to a small party in that year.

incumbent was disqualified as a result of the education requirement, political competition increased along all four dimensions I study. Disqualification of a political heavyweight can encourage new candidates, who otherwise may have thought they did not have any chance of winning, to enter the political arena. In addition, disqualifying a stronger incumbent also likely reduced both the absolute and the relative strength of the new winner.

Table 7 shows that in those areas where a stronger incumbent, as measured by a larger initial vote margin, was disqualified, political competition increased on all four dimensions. Specifically, with a 10% increase in the initial vote margin for incumbents who were later disqualified, the new winner's vote share decreased by 0.09, his vote margin decreased by 0.12, the number of candidates running increased by 1.46, and the Herfindahl-Hirschman index increased by 0.07. These effects are both large and significant.

These results provide evidence that where a political heavyweight was disqualified as a result of this law, political competition actually increased by encouraging more candidates to run, by reducing the absolute and relative strength of the winner, and by making the race more competitive. This result runs contrary to the expected outcome of a ballot access restriction: it shows that this law actually served to strengthen political competition in some areas. It is important to reiterate here that the welfare effects of stronger political competition are ambiguous. I examine the effects of this policy change on legislator experience, one dimension of legislator quality, in Online Appendix F.

5 Conclusion

This paper provides the first evidence of the electoral effects of an unprecedented ballot access restriction based on education in a developing country governed by a flawed political system, characterized by a functioning and democratically elected Parliament, but with a military dictator at its helm. Electoral rules, such as those related to candidate eligibility, are usually changed so significantly only in circumstances where a dictator can exercise control over them, in this case General Musharraf, and the effects of such policy changes are important to quantify. This paper has provided the first rigorous evaluation of the effects of the legislator education requirement imposed in Pakistan in 2002 on political competition.

I show that incumbent disqualification as a result of this requirement had varying effects on political competition, depending on the underlying political and socio-demographic constituency characteristics. The paper shows that both party and candidate characteristics (party size, and the strength of the incumbent measured by his initial vote margin) matter in explaining the effect of disqualified incumbents, which shows that this law affected both candidates and parties.

Specifically, competition declined sharply in those areas where the incumbent was disqualified and there was a smaller pool of available substitute candidates, both when low substitutability is a consequence of the incumbent belonging to a small party or of low constituency literacy levels. We know that these two measures are correlated: that constituencies which had elected a candidate from a small party in 1997 are significantly less literate. However, there is a fundamental difference between the long-term effects of substitutability as measured by a small party incumbent versus low constituency literacy. Small parties can ultimately adjust to the new policy by actively recruiting more educated members and candidates. On the other hand, literacy in the constituency cannot change as easily, and therefore the weaker political competition which I find in low literacy constituencies with disqualified incumbents may persist in the long run.

Second, I find that political competition was invigorated where stronger political incumbents were disqualified, as measured by a larger initial vote margin, because the disqualification of political heavyweights led to entry of new candidates into the political arena. This lends credence to the establishment's claim that the politically powerful opposed the law. Therefore, while the fact remains that this education requirement disqualified the vast majority of the nation's citizens from contesting elections, it is important to note, contrary to theory and popular opinion, that political competition did not decline everywhere, even across constituencies where the incumbent was disqualified, and actually strengthened in certain areas.

To interpret the effects I find in this paper in terms of economic and policy outcomes, it is useful to look to a large theoretical and empirical literature that relates political competition to economic outcomes. Notably, Besley and Burgess (2002) show that larger voter turnout rates and stronger political competition is associated with better government responsiveness

in the form of more generous public food distribution and more effective responses to natural disasters in India. Besley, Persson, and Sturm (2010) show that more vigorous political competition led to higher incomes for the United States. Higher political competition is also theoretically thought to help in the alignment of voter and representative preferences, but it can increase information costs for voters in making their electoral choices. In theoretical work, Bardhan and Yang (2004) show that greater political competition leads to tradeoffs of economic costs and benefits, and that it can pose a threat to long term investments because of uncertainty that extends across election cycles. The authors argue that information asymmetries, distributional conflicts, and the characteristics of public investment opportunities can play a role in mapping electoral competition into good or bad economic outcomes. Hence, more political competition does not necessarily translate into higher welfare, although much of the empirical evidence does point in that direction.[27] I am examining the policy outcomes of this ballot access restriction in Pakistan in separate research.

Acknowledgments

I would like to thank the editors and three anonymous referees for their excellent comments and suggestions.

[27] In addition to the literature referenced above, see Besley and Case (1995).

6 References

Adams, J., & Merrill, S. (2006). Why small, centrist third parties motivate policy divergence by major parties. *American Political Science Review, 50(3),* 403-417.

Ansolabehere, S., & Gerber, A. (1996), The effects of filing fees and petition requirements on US House elections. *Legislative Studies Quarterly, 21(2)*, 249-64.

Bardhan, P., & Yang, T. (2004). Political competition in economic perspective. University of California, Berkeley, Department of Economics Working Paper No. E04-341.

Besley, T., Persson, T., & Sturm, D. (2010). Political competition and economic performance: evidence from the United States. *Review of Economic Studies, 77(3)*, 1329-1352.

Besley, T., & Burgess, R. (2002). The political economy of government responsiveness: theory and evidence from India. *Quarterly Journal of Economics, 117(4),* 1415-1451.

Besley, T., & Case, A. (1995). Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits. *Quarterly Journal of Economics, 110(3),* 769-798.

Bhatty, M. (2002). Positive electoral reforms. The Dawn Newspaper - Opinion, January 28.

Drometer, M., & Rincke, J. (2009). The impact of ballot access restrictions on electoral competition: evidence from a natural experiment. *Public Choice, 138,* 461-474.

Drometer, M., & Rincke, J. (2008). The design of political institutions: electoral competition and the choice of ballot access restrictions in the United States. CESifo working paper No. 2406.

Haven, P. (2002). Pakistan candidates must have degree. The Washington Post, October 7.

Lijphart, A. (1990). The political consequences of electoral laws, 1945-85. *American Political Science Review, 84,* 481-497.

Linden, L. (2005). Could have been a contender? The effects of limiting the number of candidates in Indian state parliamentary elections." Unpublished manuscript.

Rae, D. (1971). *The political consequences of electoral laws.* New Haven: Yale University Press.

Rehman, I. A. (2006). Polls and madrassah 'sanads'. The Dawn Newspaper - Opinion, June 8.

Stratmann, T. (2005). Ballot access restrictions and candidate entry in elections. *European Journal of Political Economy, 21,* 59-71.

Tullock, G. (1965). Entry barriers in politics. *American Economic Review, 55(1),* 458-466.

Tables

Table 1: Summary statistics for 1997 and 2002 election

Variable	Pol comp relation	Obs	Mean	SD	Min	Max
1997 data						
Disqualified: Binary 0/1 variable		198	All 0		0	0
Vote share	-	206	0.55	0.13	0.16	0.86
No. of candidates	+	206	8.55	8.52	2	107
Vote margin	-	206	0.26	0.16	0.00	0.75
Herfindahl-Hirschman index	+	206	0.57	0.11	0.26	0.91
Population		203	405946.3	69622.07	163933	635494
2002 original data						
Vote share	-	262	0.49	0.11	0.26	0.91
No. of candidates	+	262	7.54	3.87	2	26
Vote margin	-	262	0.17	0.14	0.00	0.88
Herfindahl-Hirschman index	+	262	0.62	0.10	0.17	0.82
Population		268	489094	72893.23	179078	717649
2002 data matched to 1997 constituencies						
Disqualified: Binary 0/1 variable		198	139 0s, 59 1s		0	1
Vote share	-	198	0.49	0.10	0.30	0.91
No. of candidates	+	198	7.41	3.73	2	23
Vote margin	-	198	0.17	0.13	0.01	0.88
Herfindahl-Hirschman index	+	198	0.62	0.09	0.17	0.80
Population		198	497211.7	61596.64	225441	692228

Vote share = winning candidate's votes/total votes polled in the constituency.

Number of candidates is the number of candidates who ran for election in the constituency.

Vote margin = (winner's votes - runner-up's votes)/total votes polled in the constituency.

Herfindahl-Hirschman index = $1 - \sum VS_i^2$, where VS_i = vote share of candidate i.

Table 2: Summary statistics for constant constituency variables

Variable	Obs	Mean	SD	Min	Max
Literacy	195	43.24	13.76	14.01	70.4
Small party candidate	198	0.19	0.39	0	1
Party size	198	136.45	69.06	1	178
Total small party candidates	198	5.11	4.96	0	35
Initial vote margin	206	0.26	0.16	0.00	0.75

Literacy is the percentage of constituency aged 10 and older which is literate, using 1998 census district level data, which is matched to the constituency level.

Small party is an indicator for whether the incumbent in 1997 belongs to a small party: defined as one which fields fewer than 50 candidates.

Party size is the number of candidates put up for election by the incumbent's party in 1997.

Total small party candidates is the number of candidates in each constituency who belong to a small party in 1997.

Initial vote margin is defined as equal to the vote margin in the 1997 election for both election years.

Table 3: Basic Panel Regression

	Vote share	Number of candidates	Vote margin	Herfindahl-Hirschman index
Pol comp relation	-	+	-	+
Disqualified	0.041	-1.101	0.049	-0.036
	(0.038)	(1.062)	(0.048)	(0.033)
Year	-0.095***	-0.409	-0.119***	0.074***
	(0.027)	(0.656)	(0.033)	(0.023)
Fixed effects	constituency	constituency	constituency	constituency
R-squared	0.98	0.95	0.84	0.99
Obs	396	396	396	396

OLS regression with robust standard errors clustered by constituency in parentheses.

Vote share = winning candidate's votes/total votes polled in the constituency.

Number of candidates is the number of candidates who ran for election in the constituency.

Vote margin = (winner's votes - runner-up's votes)/total votes polled in the constituency.

Herfindahl-Hirschman index $= 1 - \sum VS_i^2$, where VS_i = vote share of candidate i.

Disqualified equals 1 for a constituency in 2002, if the MNA elected from that constituency in 1997 did not have a Bachelor's degree or higher; it equals 0 for all constituencies in 1997.

Control included but not shown: constituency population (coefficient insignificant).

* denotes significance at 10%, ** at 5%, and *** at 1%.

Table 4: Literacy interaction

	Vote share	Number of candidates	Vote margin	Herfindahl-Hirschman index
Pol comp relation	-	+	-	+
Disqualified	0.309***	-4.652*	0.334**	-0.287***
	(0.102)	(2.817)	(0.149)	(0.086)
Literacy	0.007***	0.068***	0.003**	0.009***
	(0.001)	(0.022)	(0.002)	(0.001)
Literacy X Disqualified	-0.007***	0.097	-0.007**	0.006***
	(0.002)	(0.068)	(0.003)	(0.002)
Year	-0.105***	-0.034	-0.130***	0.081***
	(0.027)	(0.502)	(0.034)	(0.024)
Fixed effects	constituency	constituency	constituency	constituency
R-squared	0.98	0.96	0.85	0.99
Obs	378	378	378	378

OLS regression with robust standard errors clustered by constituency in parentheses.

Vote share = winning candidate's votes/total votes polled in the constituency.

Number of candidates is the number of candidates who ran for election in the constituency.

Vote margin = (winner's votes - runner-up's votes)/total votes polled in the constituency.

Herfindahl-Hirschman index = $1 - \sum VS_i^2$, where VS_i = vote share of candidate i.

Disqualified equals 1 for a constituency in 2002, if the MNA elected from that constituency in 1997 did not have a Bachelor's degree or higher; it equals 0 for all constituencies in 1997.

Literacy is the percentage of constituency aged 10 and older which is literate, using census district level data, which is then matched to the constituency level.

Controls included but not shown: constituency population (coefficient insignificant).

* denotes significance at 10%, ** at 5%, and *** at 1%.

Table 5: Small party interaction

	Vote share	Number of candidates	Vote margin	Herfindahl-Hirschman index
Pol comp relation	-	+	-	+
Disqualified	0.001	0.090	-0.007	-0.006
	(0.038)	(0.856)	(0.048)	(0.033)
Small party	0.556***	4.292*	0.260*	0.589***
	(0.126)	(2.582)	(0.144)	(0.111)
Small party X Disqualified	0.188**	-5.402*	0.254***	-0.138**
	(0.078)	(3.119)	(0.082)	(0.070)
Year	-0.095***	-0.402	-0.120***	0.074***
	(0.026)	(0.637)	(0.032)	(0.023)
Fixed effects	const	const	const	const
R-squared	0.98	0.95	0.85	0.99
Obs	396	396	396	396

OLS regression with robust standard errors clustered by constituency in parentheses.

Vote share = winning candidate's votes/total votes polled in the constituency.

Number of candidates is the number of candidates who ran for election in the constituency.

Vote margin = (winner's votes - runner-up's votes)/total votes polled in the constituency.

Herfindahl-Hirschman index $= 1 - \sum VS_i^2$, where VS_i = vote share of candidate i.

Disqualified equals 1 for a constituency in 2002, if the MNA elected from that constituency in 1997 did not have a Bachelor's degree or higher; it equals 0 for all constituencies in 1997.

Small party is an indicator for whether the incumbent in 1997 belongs to a small party: defined as one which fields fewer than 50 candidates across all constituencies.

Control included but not shown: constituency population (coefficient insignificant).

* denotes significance at 10%, ** at 5%, and *** at 1%.

Table 6

Checking the assumptions underlying the small party variable

	Incumbent party run	Incumbent party win	Small party win
Disqualified	0.0424	0.0827	0.1013
	(0.0785)	(0.1209)	(0.0791)
Small party X Disqualified	-0.5067**	-0.3462*	-0.7793***
	(0.2034)	(0.1773)	(0.1920)
Year	0.8489***	0.4173***	-0.0144
	(0.0433)	(0.0597)	(0.0523)
Fixed effects	const	const	const
R-squared	0.92	0.72	0.76
Obs	396	396	396

OLS regression with robust standard errors clustered by constituency in parentheses.

Incumbent party run is an indicator for whether a candidate from the incumbent party in 1997 ran again from the same constituency in 2002.

Incumbent party win is an indicator for whether a candidate from the incumbent party in 1997 won again from the same constituency in 2002.

Small party win is an indicator for a small party winning, defined separately for each election year.

Disqualified equals 1 for a constituency in 2002, if the MNA elected from that constituency in 1997 did not have a Bachelor's degree or higher; it equals 0 for all constituencies in 1997.

Small party is an indicator for whether the incumbent in 1997 belongs to a small party: defined as one which fields fewer than 50 candidates across all constituencies.

* denotes significance at 10%, ** at 5%, and *** at 1%.

Table 7: Initial vote margin interaction

	Vote share	Number of candidates	Vote margin	Herfindahl-Hirschman index
Pol comp relation	-	+	-	+
Disqualified	0.251***	-4.462*	0.334***	-0.192***
	(0.047)	(2.417)	(0.048)	(0.047)
Initial vote margin	0.787***	18.585***	0.587***	0.735***
	(0.134)	(2.869)	(0.151)	(0.119)
Initial vote margin X Disqualified	-0.911***	14.552**	-1.234***	0.677***
	(0.137)	(7.499)	(0.137)	(0.142)
Year	-0.095***	-0.412	-0.119***	0.074***
	(0.025)	(0.639)	(0.031)	(0.023)
Fixed effects	const	const	const	const
R-squared	0.98	0.95	0.88	0.99
Obs	396	396	396	396

OLS regression with robust standard errors clustered by constituency in parentheses.

Vote share = winning candidate's votes/total votes polled in the constituency.

Number of candidates is the number of candidates who ran for election in the constituency.

Vote margin = (winner's votes - runner-up's votes)/total votes polled in the constituency.

Herfindahl-Hirschman index $= 1 - \sum VS_i^2$, where VS_i = vote share of candidate i.

Disqualified equals 1 for a constituency in 2002, if the MNA elected from that constituency in 1997 did not have a Bachelor's degree or higher; it equals 0 for all constituencies in 1997.

Initial vote margin is defined as equal to the vote margin in the 1997 election for both election years.

Control included but not shown: constituency population (coefficient insignificant).

* denotes significance at 10%, ** at 5%, and *** at 1%.