**Title**
Essays on Dynastic Politics and Political Participation



**Author**
Sonnet, Luke

**Publication Date**
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

# Essays on Dynastic Politics and Political Participation

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Political Science

by

Luke Sonnet

2020

ABSTRACT OF THE DISSERTATION

Essays on Dynastic Politics and
Political Participation

by

Luke Sonnet
Doctor of Philosophy in Political Science
University of California, Los Angeles, 2020
Professor Miriam A. Golden, Chair

This dissertation contains three essays on electoral politics, with focuses on political dynasties, political participation, and the political context of Pakistan. All three essays deal with the barriers that certain individuals face when participating in politics. The first essay explores how certain institutional changes, but not others, curtail the power of political families and the conditions under which electoral reforms create space for new candidates for legislative office. While the second essay makes a largely methodological contribution, the central result that political dynasties hold advantages even in close elections also has implications for who can enter and win elections. The third essay documents how social norms act as barriers to female political participation in Pakistan and considers the kinds of expectations that may increase female turnout when it lags far behind male turnout.

In the first essay, my co-authors, Ali Cheema and Farooq Naseer, and I estimate the effect of institutions on political dynasties. Previous evidence is mixed on whether institutional change disrupts elite power and policy prescriptions for how to improve the competitiveness of elections are unclear. We contribute to this literature by studying the effects of two

specific reforms enacted before the 2002 General Elections in Pakistan—reapportionment and an education minimum for legislative candidates. Introduced by General Musharraf's military regime, these reforms were specifically aimed at powerful political families. However, we provide evidence that only one was effective. Using a pseudo-regression discontinuity design relying on the formulaic assignment of seats to districts following reapportionment, we estimate that increasing the number of seats in the legislature dilutes the prevalence of political dynasties in elected positions. Political families were able to win just as many seats as before reapportionment, but were unable to capture many of the new seats, leaving room for politicians from other families to enter and win new seats. On the other hand, the education minimum was less effective at curtailing incumbent power. Using a differences-in-differences design, we estimate that areas most affected by the disqualifications caused by the education minimum were more likely to have connected family members entering and winning elections for the first time, while "outsiders" were even less likely to enter and win the newly vacated seats. We argue that elites may be more willing to respond to disqualifications than an expansion in the number of seats in a legislature due to diminishing marginal returns to the number of elected family members. This implies that increasing legislature sizes may be more effective at curtailing elite power than other institutional reforms that restrict who can run in an election.

The second essay makes a largely methodological contribution by providing evidence that members of political families are more likely to win very close elections. This imbalance in the dynastic status of winners and losers of close elections exists in datasets from Japan, the Philippines, Pakistan, and the United States. On average, the winner of a close elections is about 6 percentage points more likely to be a family member of a previously elected politician than a loser of a close election. This imbalance, while perhaps unsurprising given the well-documented advantage dynastic candidates have in general, has important implications for the validity of close elections regression discontinuity designs. Predicated on the assumption that those who barely win and those who barely lose elections are similar on many characteristics, close elections regression discontinuity designs aim to estimate the causal effect of winning an election. However, if there is a detectable imbalance in the characteristics of winners and losers of close elections, as with dynastic status here, then estimates using close elections regression discontinuity designs may be biased. In fact, I calculate that between 9 and 40 percent of top political science and economics papers that use these designs have their most robust estimate quartered by bias of this magnitude.

The third essay moves away from candidacy and the legislature and focuses on the issue of female political participation in Pakistan and the role that norms play in constraining or motivating behavior. Using data from 37 communities on the beliefs of men and women, their expectations about others in their community, and their social networks, this essay provides evidence that behavior and expectations about behavior are more strongly correlated with self-reported female political participation than normative beliefs and expectations. In other words, women self-report voting more often when they believe women in their social network vote and when women in their social network self-report voting. This relationship between self-reported voting and both empirical expectations—beliefs about what others actually do—and the actual behavior of others with a woman's own behavior is stronger than the relationship between self-reported voting and normative expectations—beliefs about what others think one ought to do. I argue that these results are explained by limited female mobility outside the home in Pakistan. Due to this restriction, female turnout takes the form of a coordination game and thus empirical expectations and behavior are more important than normative expectations in explaining when women will vote. Lastly, I provide evidence that expectations of both kinds serve to license political participation; women who think others support their right to vote will be more likely to turn out only if they themselves support a woman's right to vote. In a context like Pakistan, supportive social norms regarding suffrage only mobilize women when they themselves are supportive of female political participation.

The dissertation of Luke Sonnet is approved.

Chad J. Hazlett

Jeffrey B. Lewis

Daniel N. Posner

Miriam A. Golden, Committee Chair

University of California, Los Angeles

2020

*In memory of my mother*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

## ACKNOWLEDGMENTS

This dissertation is a product of the support, training, and mentorship of many faculty members. I want to thank my dissertation committee chair, Miriam Golden, for her support and instruction. From my first graduate seminar to mentorship as we manage field projects together, Miriam has provided both encouragement and amazing opportunities that made my research possible. I am grateful to my other committee members, Chad Hazlett, Jeff Lewis, and Daniel Posner, for their feedback throughout my training as papers and ideas came in piece by piece. I also thank Graeme Blair and Chad Hazlett for providing me opportunities to develop my ability to work in the field. Working on the VoteView project with Jeff Lewis and the DeclareDesign project with Graeme Blair and Alex Coppock are also highlights of my time as a graduate student. I also am fortunate to have worked with Saad Gulzar on several projects, who continues to be a source of friendship and scholarly energy.

I would also like to thank my collaborators. Chapter 1 is co-authored with Ali Cheema and Farooq Naseer, who provided an academic home for me in Lahore. Chapter 3 relies upon data and a survey instrument that come from a joint project co-led by Saad Gulzar and Muhammad Yasir Khan, along with myself. I also thank all of those at various conferences and workshops that have provided me with feedback, especially the members of the Comparative Reading Group at UCLA and the workshops at the European University Institute. Lastly, much of my work would not have been possible without the large number of research assistants who have contributed to data I use, designed survey instruments along side me, and spent many hours working with government officials. In particular, I would like to thank Mohammed Bukar and Mamoor Ali Khan for their friendship and dedicated work.

Support for research in this dissertation was provided by the UCLA Graduate Division, the LUMS-World Bank Pakistan at 100 pre-doctoral fellowship, the Abdul Latif Jameel Poverty Action Lab (Chapter 3), the International Growth Centre (Chapter 3), and the Stanford University Center for Global Development (Chapter 3). The human subjects research in Chapter 3 is covered by UCLA IRB#18-000783.

Without my fellow students, this dissertation would not have been possible. In particular I would like to thank Kye Barker, George Derpanopoulos, Imil Nurutdinov, Blair Read, Aaron Rudkin, and Joy Wilke for their friendship, support, advice on how to store coffee beans in a cupboard, and for providing an emotional home even as our work took us to different parts of the world. To the rest of my cohort, thank you for creating an excellent learning environment and for the late nights in the lounge pushing through problem sets.

Splitting your time between three locations while writing a dissertation requires a few constants to keep you centered. For always being there, I am grateful to my college friend group, without whom I would have a much shallower understanding of the world around me and a much deeper distrust of engineers and lawyers. My sisters, Aliya and Jasmine, provide me with unconditional love and a place to call home after so many years on the move together. I also thank my loving father for constant support and for giving me the nickname "doctor" from a young age, an unintential, but somewhat prophetic, nod to nominative determinism. Unfortunately for him, he never specified which kind of doctor. I wish I could thank my mother for basically everything, but a dedication will have to suffice. Any feelings of achievement and success that I could draw from completing this thesis are muted by her absence.

Finally, I am grateful for my best friend, Emily. In addition the fact that she pointed out the color of ggplot axis labels are grey by default, requiring me to regenerate every plot in this dissertation with black axis labels, Emily's continual supply of emotional support and love make any achievement of mine equally hers.

## VITA

| | |
|---|---|
| 2009–2012 | B.S. (Economics, Political Science, and History), University of Pittsburgh. |
| 2013– | Ph.D. student (Political Science), University of California, Los Angeles. |

## PUBLICATIONS

Boche, Adam, Jeffrey B. Lewis, Aaron Rudkin, and Luke Sonnet. "The new Voteview.com: Preserving and Continuing Keith Poole's Infrastructure for Scholars, Students and Observers of Congress", *Public Choice*, 176 (1-2): 17–32.

# CHAPTER 1

# Institutional Change and Dynastic Persistence in Pakistan: The Impact of Electoral Reapportionment and Education Minimums

**Co-authored with Ali Cheema (LUMS) and Farooq Naseer (LUMS)**

**Abstract:** The pervasiveness of families in electoral politics around the world highlights the ability of elites to self-perpetuate, even in the face of considerable institutional change. In this paper, we test how political dynasties in Pakistan responded to two institutional reforms introduced with the explicit goal of curtailing elite power. We demonstrate that electoral reapportionment and the introduction of education minimums—two key components of a reform package designed by General Musharraf in Pakistan following his 1999 coup—had different consequences for political families, a powerful elite group in Pakistan. Using a pseudo-regression discontinuity design leveraging district level reapportionment formulae, we estimate that an additional seat within a district decreases the probability a constituency is won by a dynastic candidate by 9 percentage points, diluting the power of the incumbent elite. On the other hand, using a differences-in-differences design, we estimate that the disqualification of incumbents via education minimums only causes churning within the elite; while incumbents were disqualified, their family members won at high rates in their stead and dynastic prevalence in politics was unchanged. We argue that these differential responses to institutional reforms may be the result of diminishing marginal returns to the number of family members in elected office. These results highlight the usefulness of legislature size

and reapportionment to curtailing the power of incumbent elites.

## 1.1 Introduction

The tendency of elected elites to entrench themselves and their close networks in power has been seen as a concern for democratic accountability since the time of classical elite theorists (Michels, 1911; Mosca, 1939; Pareto, 1991). The recent research on the formation and persistence of political dynasties (e.g. Dal Bó, Dal Bó and Snyder, 2009; Querubin et al., 2016) documents how the power held by elites can become closed and self-perpetuating, thereby creating the conditions for sustained elite capture even in the face of institutional change (Acemoglu and Robinson, 2006). The literature finds support for the observation by Mosca (1939) (as cited in (Dal Bó, Dal Bó and Snyder, 2009)) that "...every class displays the tendency to become hereditary..." and dynastic in a diverse set of electoral contexts.[1] An open question in the literature is what types of institutions and which institutional reforms have the potential to perpetuate or erode dynastic persistence.

The existing evidence on the effects of institutional design and reforms on elite power is mixed. Some find that reforms which are nominally designed to curtail the power of entrenched elites, such as term limits, seat reservations, and the enfranchisement of previously excluded social groups, can easily be navigated so that dynastic elites suffer little loss in actual power (Labonne, Parsa and Querubin, 2019; Auerbach and Ziegfeld, 2020; O'Connell, 2020; Acemoglu and Robinson, 2006). In contrast, others find that party-centered electoral systems may prevent the formation of political dynasties (Fiva and Smith, 2018). We contribute to this literature by studying the impact of two institutional reforms—electoral reapportionment and the introduction of education minimums for legislative candidates—on dynastic persistence in Pakistan. Using a novel dataset of family membership for all national and provincial candidates in Pakistan from 1970 to 2013, we show that the two reforms induced different responses in political dynasties and had differential impacts on

[1] Evidence on dynastic persistence has been documented in the United States (Dal Bó, Dal Bó and Snyder, 2009), India (Chandra, 2016; George and Ponattu, 2019), the Philippines (Querubin et al., 2016), Pakistan (Cheema, Javid and Naseer, 2013), and Japan (Smith, 2018).

dynastic persistence in subsequent years.

Reapportionment and education minimums were critical components of the set of constitutional and electoral reforms introduced by General Musharraf's military regime in 2002. The reforms were a part of the regime's effort to create an electoral authoritarian state based on partisan elections (Adeney, 2017).[2] They were aimed to shock the electoral dominance of powerful political families, which compromised a third of the National Assembly before the coup and dominated leadership positions in parties, and to lower the cost of entry for educated middle class professionals in electoral politics (Jalal, 2017; Waseem, 2006). These two components of the reforms had long-term effects: both reforms applied to the 2002 and 2008 elections held under military rule, and the 2002 electoral reapportionment also applied to the 2013 elections held five years after direct military rule had ended.

A set of prominent studies have highlighted the adverse representational outcomes that tend to result from an absence of regular readjustments in the size of the legislature in majoritarian electoral systems (Dahl and Tufte, 1973; Shugart and Taagepera, 1989; Lijphart and Aitkin, 1994; Allen and Stoll, 2014). The argument is based on a recognition that the increasing size of electoral constituencies, a consequence of the dissonance between the growth of the electorate and the legislature, is likely to heighten the entry barriers for contestants seeking to dislodge entrenched politicians. The literature argues that this is because elections are costlier and the costs of building competing coalitions against entrenched majorities are higher in larger constituencies (Allen and Stoll, 2014).

Similar considerations motivated the architects of the 2002 reforms to increase the size of the national and provincial legislatures[3] and reapportion and redelimit electoral constituen-

[2]The 2002 and 2008 elections were held under military rule and the power of the assemblies and their ministers was seriously curtailed during this period. Pakistan is a federal parliamentary government with a National Assembly and four provincial assemblies. While the reforms did not alter this structure, the 17th amendment to the constitution granted the President—General Musharraf himself—and provincial governors the power to remove the prime minister of the National Assembly and the Chief Ministers of each province, respectively.

[3]The 2002 Conduct Of General Elections Order increased the number of non-reserved general seats in

cies between administrative districts within provinces. The expectation was that the new, smaller electoral constituencies would induce greater competition for entrenched political families by enabling the entry of new contestants patronized by the military regime.[4] The framers thought that this would be a consequence of lower entry barriers because elections would be less expensive and the demand for new candidates would increase. Are political dynasties well situated to respond to this new demand? We would expect political dynasties to lose their advantage if a large portion of their success comes from the personal vote (Fowler and Hall, 2014) that is mobilized by well organized but highly localized voting networks. The adverse effect of reapportionment on dynastic persistence will be stronger if the marginal benefit of having additional family members in office tends to decline for political families.

We identify the effects of reapportionment and redelimitation using a pseudo-regression discontinuity design (RDD) that relies on the formulaic assignment of new seats across districts. Following the 1998 census, Pakistan redrew its constituencies and assigned the number of constituencies to each district following a population based formula. Districts that were just above a threshold for a half of a seat were assigned that additional seat, while districts just below did not get the additional seat. Using election results in the three elections following this delimitation, we estimate using a pseudo-RDD that assignment to an additional seat decreases the probability a dynastic candidate wins one of those seats by 8-10 percentage points. A district level analysis shows that, while the probability of an individual seat being won by a dynast goes down in districts with an additional seat, the total number of dynasts winning is unchanged. Taken together, these results suggest that political families are only occasionally challenging for new constituencies and rather expend effort to hold on to their existing mandates. Redelimitation and reapportionment thus dilutes the

the National Assembly from 207 to 272 and the total number of general seats in the provincial assemblies from 460 to 577 seats.

[4] For a contemporary op-ed in the leading English-language newspaper making this argument, see Waseem (2002). Also see European Union (2002) and Jalal (2017).

power individual dynasts have in the legislature by reducing the share of seats they hold.

Musharraf's electoral reforms also introduced a minimum Bachelor's degree requirement as a qualification to contest elections.[5] The objective was to create demand for new candidates by disqualifying strong incumbents and potential challengers from the pool of experienced candidates at the time (Afzal, 2014; Jalal, 2017; Waseem, 2002). However, unless they are differentially affected by the education minimum, political families may be better positioned to respond to a constricted pool of candidates resulting from this education minimum. If the core of a family's electoral constituency remains intact, the benefits of the family brand and localized networks can be transferred between generations and advantage dynastic candidates over outside challengers in the constricted pool of qualified candidates. Political families also have a strong incentive to put forward a new family member in order to retain at least one member of the family in office. Therefore, we would expect that members of political families may be particularly suited, then, to respond to quotas, minimums, and term limits (e.g. Labonne, Parsa and Querubin, 2019).

We follow Afzal (2014) and use a difference-in-differences design to estimate the effect of the education minimum imposed by General Musharraf before the 2002 elections. We estimate that as the share of national legislators removed from office by the education minimum increases, the probability a constituency is won by first-time dynastic candidates increases. The same is not true for non-dynastic first-time candidates. Furthermore, the share of constituencies won by dynastic candidates overall is unchanged. Therefore, incumbent elites were able to respond to this institutional restriction by replacing themselves with family members while outsiders were unable to seize seats vacated by powerful incumbents. The education minimum created churn within political families, but seems to have done little to curtail their overall power.

We argue that the differential responses of political families to reapportionment and

[5]Controversially, Islamic degrees were granted equivalence to a Bachelor's degrees, which may have advantaged religious parties (Rehman, 2006).

redelimitation and education minimums may rest in both the willingness and ability of political families to forward new candidates. We argue that that families face diminishing marginal returns to the number of family members in office, and will overcome barriers that remove an incumbent, but will not work as hard to get a second member of the family elected when presented with the opportunity. We also argue that institutions which target a traditional source of power for a group, such as how redelimitation targets the link between political families and local networks, can be effective at mitigating that groups response to institutional change. We provide observational evidence for one of these mechanisms, but leave a rigorous examination for future work.

In addition to contributing to the broader literature on institutional change and elite persistence highlighted above, this paper forwards an additional motivation for increasing legislature sizes: to curtail the power of entrenched elites. The evidence that reapportionment diluted the share of legislators coming from political families contributes to existing arguments for increasing legislature size in order to improve representativeness (New York Times Editorial Board, 2018) and the election of traditionally under-represented groups (Allen and Stoll, 2014; Kjaer and Elklit, 2014). Furthermore, these results imply that the contentious decision by the Pakistani government not to increase the number of seats following the 2017 census likely reinforced the power of existing elites.

The rest of the paper proceeds as follows. First, we describe the institutional context of politics in Pakistan, in particular we highlight the role played by political dynasties in this context. Second, we turn to the specific institutional changes that we study, their legal history, and potential consequences. Third, we present the research design and results for the effect of reapportionment and redelimitation on dynastic persistence. Fourth, we present the research design and results for the effect of education minimums. Finally, we explore some potential mechanisms for these findings before concluding.

## 1.2 Institutional context: hybrid regimes and dynastic elites in Pakistan

This section describes the institutional context and the structure of elite politics that defined the political environment for the Musharraf regime's 2002 electoral reforms.

### 1.2.1 Political institutions in Pakistan

Pakistan's Constitution enacted in 1973 provides for a parliamentary system with governance centered around elected legislatures and governments at the federal and provincial levels.[6] Power formally rests in the executive heads of these governments and their cabinets, who are elected by majorities in national and provincial assemblies. The majority of members in the national and provincial assemblies are directly elected from single-member electoral constituencies on a partisan basis under a plurality rule. They are referred to as Members of the National Assembly (MNAs) and Members of the Provincial Assembly (MPAs). Political parties are pivotal organizations in this institutional structure as they are empowered to endorse candidates to contest elections and appoint elected members to critical legislative and executive positions (Mufti, Shafqat and Siddiqui, 2020).

However, repeated cycles of military rule have shaped the environment in which political parties have evolved in Pakistan.[7] Pakistan inherited a weak party system at the time of independence along with what scholars of the country's political system have described as a tradition of "tutelary democracy" (Waseem, 1989; Wilder, 1999). The development

[6] Pakistan is a federation of four provinces: Balochistan, Punjab, Khyber Pakhtunkhwa, and Sindh. The federal government and each of the provinces had their own legislatures. The federal legislature also used to include constituencies situated in the Federally Administered Tribal Areas, but this region was merged with Khyber Pakhtunkhwa in 2018. The Islamabad Capital Territory also has constituencies in the federal legislature and is separate from the four provinces. Additionally, the semi-autonomous regions of Gilgit-Baltistan and Azad Jammu and Kashmir are governed separately.

[7] As of 2020, Pakistan has been under direct military rule for 36 years, about half of its existence.

of political parties was further retarded because of an unstable constitutional compact[8] and because adult franchise elections were not held on a national level until 1970 (Jalal, 1990, 1995; Waseem, 1989). It took a period of twenty years from independence for a popular federal party, the Pakistan People's Party (PPP), to emerge on the national scene (Jones, 2003; Wilder, 1999). The PPP, along with popular provincial and regional parties, spearheaded the transition to adult franchise democracy and civilian political supremacy that became the defining principles of a stable Constitution enacted twenty six years after independence.

Mainstream political parties continued to be challenged under periods of military rule that succeeded the 1973 Constitution. As these authoritarian regimes were politically and legally constrained to operate within the framework of an elected federal parliamentary structure enshrined in the Constitution, they had to establish hybrid regimes and share power with elected members of legislatures and governments. Therefore, weakening mainstream political parties and establishing a party loyal to the military regime became a sine qua non of these regimes.

These hybrid regimes were realized by introducing disqualification requirements, constituting legal and corruption cases against national and local party leaders, and institutionalizing electoral reforms designed to weaken the political leadership base of the mainstream political parties. Each military regime helped create a King's party of loyal politicians to ensure political stability. The range of electoral reforms used for this purpose varied; General Zia's regime in the 1980s introduced a non-partisan system for election to provincial and national assemblies, while General Musharraf's regime was denied this avenue by the Court.[9] Instead, General Musharraf resorted to increasing the size of the legislature and the

[8]Pakistan created three constitutions in its first 26 years of existence.

[9]In Benazir Bhutto v. Federation of Pakistan and one other (PLD 1989 SC 66), the Supreme Court ruled that the use of election symbols by parties gives substance to the fundamental right to free elections (Newberg, 1995). Furthermore in Mian Muhammad Nawaz Sharif v. President of Pakistan and others (PLD 1993 SC 473), the Supreme Court held that where a political party contests the election successfully, it has a fundamental right to form government provided it enjoys the requisite majority

Table 1.1: Dynastic incidence among first-time candidate across four countries

| | Japan | Philippines | United States | Pakistan |
|---|---|---|---|---|
| Election Years | 1947-2014 | 1946-2007 | 1788-1988 | 1985-2013 |
| N Candidates | 3123 | 2690 | 6257 | 4626 |
| % Dynastic | 14% | 29% | 8% | 20% |

Analysis and percentages drawn from a cross-national analysis in Sonnet (2020). Data on Japan are from (Smith, 2018), data on the Philippines are from (Querubin et al., 2016), and data on the United States are from (Dal Bó, Dal Bó and Snyder, 2009). We only consider first-time candidates (or first-time incumbents in the case of the United States) here. Note that a candidate is coded as dynastic if they had a family member precede them in office at a similar level. We discuss this further in Section 1.3 below.

use of electoral apportionment and education minimums for this purpose (Jalal, 2017).

### 1.2.2 Dynastic politics abroad and in Pakistan

The combination of repeated disruptions faced by political parties and the centrality of elected parliaments, in periods of both military and civilian rule, has resulted in the political space being filled by dynastic families and their electoral organizations. Table 1.1 shows that Pakistan has a high incidence of dynastic politics in comparison to other polities for which similar data is available. Note that for comparability to other datasets, we only consider the first election in which a candidate ran in Table 1.1 only first elections for candidates are considered.

In Pakistan, the leadership structure of dynastic organizations is built around political families. Cheema, Javid and Naseer (2013) find that the leadership structure of these organization is held together by ties of blood and leadership transitions typically involve sons, paternal nephews, and sons-in-law of the original dynasts. Khan (2020) finds that between 70-90 percent of women contesting elections for a general seat between 1988 and 2008 belonged to dynastic families.

Recent studies of micro politics in Pakistan show that the edifice of dynastic machines is built on highly localized vote blocks organized and led by local political entrepreneurs (Lyon, 2019; Martin, 2015; Mohmand, 2019; Nelson, 2009). These studies show that a variety of strategies are used by these machines to organize local vote blocks, which include using kinship and ethnic networks, patronage bonds, and ties of economic dependence. The literature also finds that at the local level voters reward local entrepreneurs with strong personalized connections with provincial and national politicians (Liaqat et al., 2019). Finally, except for a few studies on metropolitan cities in Pakistan (Liaqat, Cheema and Mohmand, 2020; Cheema, Liaqat and Mohmand, 2018), these studies find weak partisan bonds between politicians and voters.

This suggests that, as in many other contexts, dynasties form in Pakistan because family members often have significant advantages when running for office—they enjoy high name recognition that may help them at the polls (Feinstein, 2010; Rossi, 2017), their brand and resources help them gain access to attractive political parties (Geys and Smith, 2017; Chhibber, 2013; Chandra, 2016), and they may have been motivated to acquire political skills from a young age (Fox and Lawless, 2005).

Dynastic families are a rich source of political leadership in Pakistan. Many families develop strong partisan loyalties, however a large number regularly switch between parties and have provided the leadership base for the military-created "King's parties." We find that slightly more than one-third of families either run all candidates as with one party or as independents and around two-thirds have members running for two different parties. It is for this reason that military regimes have used the dual strategy of crafting the King's party around members of these dynastic political families, while at the same time trying to disrupt the dynastic basis of power through the reform of political institutions.

### 1.2.3 The Musharraf Coup and the 2002 electoral reforms

Like previous military coups, the Musharraf regime that took over power in 1999 also had to reinstate elected assemblies in order to gain parliamentary indemnity for all actions taken during the proclaimed emergency period (1999-2002).[10] The inability to engineer large scale defections of opposition politicians during the emergency period (Jalal, 2017) prompted the regime to enact electoral reforms in 2002 that weakened existing political elites. These reforms ensured that the 2002 elections were held under substantially different rules than previous elections. The majority of these were outlined by the Chief Executive's Order No. 7 and No. 24 of 2002. Together, these orders set the rules for subsequent elections in Pakistan that served the military regime's interest in breaking up entrenched elites who may be able to challenge their power. We first describe two key changes to electoral law, the number and location of constituencies as well as requirements for candidacy, before estimating the effects of these changes on the elites they targeted.

### Reapportionment and redelimitation

A cornerstone of the changes to the 2002 General Elections was the reapportionment of legislative seats both across and within provinces, as well as an increase in the number of seats overall. The National Assembly increased in size from 217 to 342 seats, although 60 of the new seats were reserved for women and 10 of the seats both before and after reform were reserved for non-Muslims.[11] The increase from 207 to 272 general seats—openly contested in first-past-the-post elections from single member constituencies—was accompanied by a reallocation of the percentages of seats to each province according to the 1998 census. In addition, the executive order that established the laws for the 2002 elections increased the sizes of each of the four provincial assemblies. The increase in the size of assemblies in 2002 is significant because it is a departure from other delimitations, which have either seen a modest increase in the size of the legislature or no increase at all.[12]

[10]The Musharraf regime issued a Proclamation of Emergency and insisted that the bloodless coup they initiated was not the beginning of a period of martial law.

[11]Both kinds of reserved seats are elected via a form of proportional representation.

[12]Although the 2017 census led to redelimitation before the 2018 elections, the civilian-led government chose not to increase the size of any assembly, even though the number of registered voters in Pakistan had

Seats for both assemblies were assigned at the district level; this meant that almost all districts would face an increase in the number of seats in the 2002 election. If a family wanted to capture a large share of seats within a district, they now would have to compete on multiple fronts. This method is crucial to our strategy to identify the causal effects of redelimitation on electoral outcomes. The number of seats assigned to a district was conditional on the share of seats that district was owed by its population relative to the population of its province (Election Commission of Pakistan, 2002, p. 37). We return to this allocation process below when discussing the research design.

Decreasing the number of constituents per seat, redrawing constituency borders, and increasing the overall number of seats was done in attempt to dilute the power of the entrenched elite and allow for the entry of new contestants (Waseem, 2002). By shrinking the size of constituencies, reapportionment could decrease the costs to running for office and allow for the entrance of candidates with fewer financial assets. Furthermore, redrawing constituencies can break the linkage between local organizations and powerful families. As discussed above, in Pakistan families rely heavily on these organizations and may draw much of their advantage from cultivating and investing in them. Redelimitation was thought to potentially break those linkages and may reduce their ability to maintain local strongholds. Indeed, in India, redelimitations have been shown to hurt incumbents who have their constituencies redrawn (Iyer and Reddy, 2013).

**Education minimums**

In the same executive order that ordered reapportionment and redelimitation, General Musharraf's military regime mandated that a Bachelor's degree—or equivalent—was mandatory for candidacy to the national and provincial assemblies. This is the first time that any such requirement was instituted in Pakistan (Election Commission of Pakistan, 2002, p. 5).

increased roughly 50 percent from 2002 to 2018.

Furthermore, candidates who had competed in previous elections affirmed their education status on candidate affidavit forms. This meant that those without Bachelor's degrees would find it difficult to claim they had all of a sudden completed this education requirement and instead were frequently disqualified.

The introduction of this bill was met with significant controversy; some criticized it as anti-democratic while others praised it, arguing that educated legislators would be less prone to corruption (Afzal, 2014; Jalal, 2017, p. 55). The strategic implications of the reform were not lost on contemporary observers either. The University Grants Commission of Pakistan decided that Islamic degrees met the standard, which prompted some observers to note how this guarded religious parties against the effects of the education requirements (Afzal, 2014, p. 55). Opposition and small parties were split on the matter; some argued that this barrier to entry made it more difficult to find candidates who could challenge for seats, while other small parties argued that the requirement removed many elite incumbents and created space for new candidates (Afzal, 2014, p. 55). In the end, 59 out of the 198 MNAs that Afzal (2014) was able to match were disqualified by this bill, a sizable number of incumbents to be ruled out of running for reelection.

## 1.3 Data

In order to study dynastic response to the two institutional changes above, we use a novel and original dataset of the family membership of legislative candidates in Pakistan from 1970 to 2013.[13] The data contain all candidates for election from the national and provincial assemblies for the 10 general elections held from 1970 to 2013 with two notable exceptions: (1) we only have data on winners from 1977 due to a lack of data on challengers and (2) we only have data on the top three placed candidates from 2013 owing to data collection

[13] First reported in Cheema, Javid and Naseer (2013) and cited as Cheema and Naseer (2013), we have since extended these data to include 2013.

constraints. We are left with a dataset of 30,213 candidates, for whom we code several important variables.

Our main outcomes of interest are built on the dynastic status of candidates in and winners of legislative elections from 2002 to 2013. Dynastic status is measured as ever having had a family member elected to office in an election wave that preceded the election wave being considered. Therefore, a father and a daughter who run together in an election for the first time would not be coded as dynastic candidates for our purposes here. However, if that daughter ran in 2002 and their father had won an election in 1997, then the daughter alone would be coded as a *dynastic candidate*. Therefore, dynasticism in our data captures the notion of being connected to a powerful political figure; dynastic candidates are following family members into the political arena. Note this means that the political dynast themselves, the first entrant into politics, is not coded as a *dynastic candidate*.

We admit both kinship and marital relations when considering whether two candidates are part of the same family. We measure political dynasties by tracing kinship and marital connections between candidates using journalists, key informants, and self-reported family data by candidates. This data collection was done in collaboration with the Monthly Herald of the Dawn group of newspapers, the largest English language newspaper group in Pakistan. Their independent data collection coincided with ours, and the two datasets were cross-verified, with Dawn district correspondents validating and updating family trees that were mismatched across datasets. We consider all blood relations at least as close as first cousins and grand uncles and aunts to be kin (shared great-grandparents). The vast majority of dynastic candidates are the nephew, brother, or son of a previously elected candidate. We also consider a few marital relations such as in-laws and the in-laws of immediate relatives; there are far fewer of these connections and the vast majority of dynastic candidates through marriage are sons-in-law or brothers-in-law. This is very similar to the definitions used in Japan (Smith, 2018) and the United States (Dal Bó, Dal Bó and Snyder, 2009). Data on dynasticism in the Philippines (e.g. Querubin et al., 2016) and India (George and Ponattu,

2019) rely instead on naming conventions or self-reported relationships to identify family ties.

For most of our analyses below, we consider the effect of a district level or super-district level policy on constituency level outcomes. Therefore, we collapse this candidate level data to the constituency level. Our main outcomes of interest are whether a *dynastic candidate wins* and whether a *first-time dynastic candidate wins*. In both instances, we take code a constituency as having a *dynastic candidate win* if the winner of the election is a *dynastic candidate*. This is a binary outcome measure. Similarly, we code *first-time dynastic candidate wins* as a binary variable that takes the value of one if the winner was a *dynastic candidate* in their first election and as zero otherwise. We return to some auxiliary outcome measures and district level measures in the results sections below when appropriate.

The exact dataset used in the analyses below, as well as how the institutional changes are measured, depends on the reform that we study. Therefore, which constituencies enter in to each analysis and how the treatments are clustered is covered in the appropriate results section below. As a result, the number of constituencies in each of our analyses may vary. Nonetheless, we never use data before 1988 as non-partisan elections and incomplete data make using the 1970, 1977 and 1985 elections impossible. The full constituency level dataset ranging from 1988 to 2013 consists of 5,201 constituency-elections. The 5,201 constituency-elections we have are 99.7 percent of the total possible of 5,215—we are missing data on one constituency in 2008 and 13 constituencies in 2013 due to cancelled elections.

The institutional reforms, our treatments of interest, are largely measurable at the district level. We discuss our assignment of constituencies to treatment conditions when we discuss how we identify the effects of both kinds of reforms below.

## 1.4 Effects of redelimitation

### 1.4.1 Research design

We leverage a natural experiment in Pakistan's redelimitation process to estimate the causal effect of receiving an additional seat within a district on outcomes in that district's constituencies. In this section, we describe how population cut-offs and seat shares are used to create as-if random variation in whether a district receives an additional seat following a redelimitation. In short, some districts have populations just large enough to entitle them to one more constituency than similar districts with a slightly smaller population.

#### 1.4.1.1 Discontinuity in seat assignment by district

Pakistan's legislative boundaries have been redrawn twice in the last 25 years. Following the 1998 census, the sizes of provincial and national legislatures were increased, with the new seats being distributed unevenly across provinces but according to population across districts within provinces. This redistricting followed a formula for almost all districts.[14] First, a quota for the number of citizens per legislator is determined by dividing the number of available seats by the population of the province. In the case of the provincial legislature, the full size of the legislature divided by the population of the province is the seat quota. For the national legislature, each province's seat allocation divided by the population of the province is the seat quota. Second, seats are assigned at the district level. Each district's population is divided by the seat quota to get a share of seats in the assembly; this seat share was then rounded to the nearest number to allocate the number of seats. In Figure 1.1, we present the official reported quotas for each province and legislative body (Election Commission of Pakistan, 2002, p. 37) and in Table 1.2 we present an example of how we use the census data and quotas to recreate the seat assignment process. Note there is a mistake in calculating

[14] For full details of this quota process, see (Election Commission of Pakistan, 2002, pp. 37–40).

the National Assembly quota for Punjab in the official document.

Figure 1.1: Official 2002 constituency quota calculations

**DETERMINATION OF QUOTA:**

7. As a first step, quota per seat in respect of each Province/Area/Agency in the National Assembly was determined as under:

| Province/Area/Agency | Population | No. of seats | Quota per seat |
|---|---|---|---|
| N-W.F.P. | 1,77,35,912 | 35 | 17735912-35=506740 |
| FATAs | 31,76,331 | 12 | 3176331-12=264694 |
| Federal Capital | 8,05,235 | 2 | 805235-2=402618 |
| Punjab | 7,36,21,290 | 148 | 73621290-148=496441 |
| Sindh | 3,04,39,893 | 61 | 30439893-61=499015 |
| Balochistan | 65,65,885 | 14 | 6565885-14=468992 |
| **Total:** | 13,23,44,546 | 272 | |

8. Quota per seat in respect of each Provincial Assembly was also determined as under:

| Name of the Assembly | Population | No. of seats | Quota per seat |
|---|---|---|---|
| Provincial Assembly, NWFP | 1,77,35,912 | 99 | 17735912-99=179151 |
| Provincial Assembly, Punjab | 7,36,21,290 | 297 | 73621290-297=247883 |
| Provincial Assembly, Sindh | 3,04,39,893 | 130 | 30439893-130=234153 |
| Provincial Assembly, Balochistan | 65,65,885 | 51 | 6565885-51=128743 |

Table 1.2: Example of seat share and seat calculations

| District | Assembly | 1998 Pop. | Quota | Seat Share | Seats |
|---|---|---|---|---|---|
| Attock | National | 1274935 | 497441 | 2.563 | 3 |
| Attock | Provincial | 1274935 | 247883 | 5.143 | 5 |
| Chakwal | National | 1083725 | 497441 | 2.179 | 2 |
| Chakwal | Provincial | 1083725 | 247883 | 4.372 | 4 |
| Rawalpindi | National | 3363911 | 497441 | 6.762 | 7 |
| Rawalpindi | Provincial | 3363911 | 247883 | 13.571 | 14 |

While only the seat quotas and not the exact seat shares are available in the 2002 Election Commission report, we can recreate the seat assignment process using the assigned number of seats in each legislature and the population data from the 1998 census that was used in the redelimitation. In the 1998 census, there were 100 districts in the four provinces of Pakistan plus the capital territory of Islamabad, treated as a unique district for the assignment of

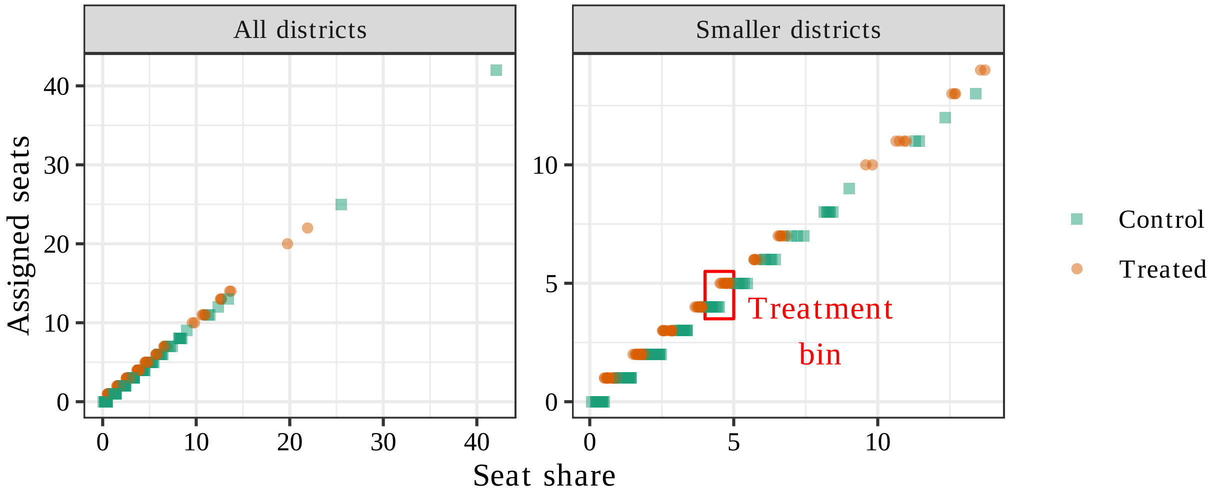
National Assembly legislators. Therefore there are 201 district-assemblies—two for each district within a province and one for Islamabad in the national legislature.

For each of these district-assemblies, the ECP calculated their assigned seat share—that is, the number of seats they ought to receive by dividing the population of the district by the seat quota for that province-assembly. We follow the same process here. Rounding these seat shares should give us the assigned number of seats for each district-assembly. This rounding creates a discontinuity at the each of the $x.5$ seat shares, where $x$ is some positive integer. A district with a 3.6 seat share gets 4 seats, and a district with a 3.4 seat share gets 3 seats. Therefore, we can consider a district $d$ "treated" if its $\text{SeatShare}_d$ modulo $1 \in (0.5, 1]$ and control if $\text{SeatShare}_d$ modulo $1 \in [0, 0.5)$. In Figure 1.2, we plot the assigned number of seats per district by the seat share of the district, with treated districts as orange circles and control districts as green squares. This figure reveals a series of discontinuities—at each 0.5 level on the y-axis, the number of assigned seats jumps by one. Therefore, the redelimitation resulted in a series of experiments where within certain bins of population size, districts either randomly get one more or one fewer seat depending on whether their "seat share" is above or below $x.5$. We call each of these experiments a treatment *bin*. In the right panel of Figure 1.2, we highlight one such treatment bin—districts with a seat share between 4 and 5 that were assigned either 4 or 5 seats depending on whether their seat share was greater or less than 4.5. In the main analysis we collapse each of these discontinuities into one discontinuity and take all districts above each cutoff as the treated districts.[15] Note that this figure has both national provincial constituencies within the same bin, while in the analysis they are treated as separate bins for ranges of national and provincial seat shares.

We argue that each of these treatment bins represents a natural experiment where districts fall on one side of the cut-off as-if randomly. Because the delimitation process was conducted after the census, we believe it unlikely that districts were able to sort on either

[15] In an alternative specification, we also estimate treatment effects at each discontinuity and average them to recover the average treatment effect. These results are reported in Appendix 1.A.

Figure 1.2: Treatment assignment at the district level




Each point is a constituency, both for the national and provincial assemblies. The right panel is the same as the left panel but zoomed in to the smaller districts.

side of the cut-off within each bin. While some districts would understand the value of inflating their population totals, inflating them so that they are just above a half share within one bin would require precise knowledge of the total population for the province and the total number of seats the legislature would contain. We also acknowledge that there were three instances where the seat share and treatment assignment link were explicitly broken; the National Assembly assignment for Okara and the provincial assembly assignments for Malakand and Nawabshah districts were off by one. They all received either one more or one less than they were due because their actual assignment would lead to too many or too few seats being assigned in the province. As these districts were closest to the cut-off, they were assigned the opposite treatment status in order to satisfy the total number of seats in the relative assembly. Note that in this paper we consider the intent-to-treat effect and assign these districts to their intended treatment status, rather than the treatment status they received.

It is also possible that each bin is simply too large and within each bin there are systematic differences between the treated and control constituencies. Indeed the population of districts is imbalanced across the cut-off—treated districts are more populous on average. We provide evidence of balance in the results below, where we use placebo tests to show that our key outcomes of interest are balanced across treatment and control in the pre-treatment period.

#### 1.4.1.2 Constituency-election dataset

We study outcomes generated at the constituency level. Almost all constituencies—817 of the 827 National and Provincial Assembly constituencies that are in our sample[16]—are contained within one district in the 2002 elections. The other 10 constituencies, 9 of which come from the sparsely populated Balochistan, are spread across multiple districts. This only happens when a district is small enough that their seat share for a particular assembly is less than 0.5—these districts should have zero seats in that assembly by the assignment mechanism, but of course leaving them without a representative is not acceptable. In these case, the population of this district is joined to a constituency of a neighboring constituency. This problem is explicitly addressed by Election Commission of Pakistan (2002), but they do not justify why they chose to merge certain constituencies other than to note that some districts were too small to justify their own constituency. We drop the 10 constituencies that cross district borders because we cannot map their outcomes to a particular district.

We also consider three election waves; the 2002 election immediately following redistricting and the 2008 and 2013 elections that used the same constituency boundaries. Therefore our chief dataset is a constituency-election dataset, wherein constituency outcomes are measured directly and treatment status is allocated at the district-assembly level that nests constituencies. This means that districts are often counted twice—once for their provincial constituencies and once for the national constituencies.[17]

[16]We exclude the 12 National Assembly constituencies that were at the time in the Federally Administered Tribal Areas and assigned by a different process

[17]Using this constituency level dataset, we can check whether our recreation of the redistricting process

In the analyses below, we exclude districts that have no other districts in the same treatment *bin*. These generally are districts that are so large that no other districts have similar populations and thus could plausibly serve as a counterfactual. This results in the omission of particularly large districts, such as Karachi, from our analysis. We also are forced to exclude a few constituencies that are in districts that are so small that the constituencies had to be formed in multiple districts. All but one of these constituencies is in the sparsely populated Balochistan, and are dropped from our analysis. The final dataset consists of all national and provincial constituencies from 2002, 2008, and 2013 that are not in the Federally Administered Tribal Areas, are wholly within one district as of 2002, and are within a district that had a district of similar size in the opposite treatment condition. This dataset contains 142 of the 201 possible district-assemblies, and, in 2002, 556 of the 827 possible constituencies. If every constituency had data for every election, we would have 1,668 constituency-elections, but a few constituencies elections had elections cancelled for various reasons. Our final dataset has 1,631 constituency-elections.

We rely on a pseudo-RDD strategy to estimate the effects of being just to the right of the cut-off—being as-if randomly assigned to getting one more seat—following the 1998 census. Our main model estimates the intent-to-treat effect of assignment to an additional seat within a district using

$$y_{cdpat} = \tau Z_{da} + \text{Bin}_{da} + \lambda_p \times \theta_t + \epsilon_{cdpat}, \tag{1.1}$$

where $y_{cdpat}$ is some outcome $y$ in constituency $c$, district $d$, province $p$, assembly type $a \in \{\text{Provincial}, \text{National}\}$, in election wave $t$; $\tau$ is the treatment effect of interest and $Z_{da}$ is

was accurate. The number of constituencies we match per district should equal the assigned number of seats for that district using just the population data. In 162 of the 201 district-assemblies, the assigned number of seats is equal to the number of constituencies we are able to match in the 2002 election data. In the other 39 districts, there is either no matched constituency—in the case where the district was too small—or the district has had one of its constituencies merged with one of those too-small districts and thus the count is not perfect.

a treatment indicator for whether the district $d$ is just above the $x.5$ threshold for the seat assignment for assembly $a$; $\text{Bin}_{da}$ is a treatment bin fixed effect for district $d$ in assembly type $a$—in other words all districts that share the nearest $x.5$ seat share cutoff and are part of the same "mini-experiment"; $\lambda_p$ and $\theta_t$ are province and election wave fixed effects, respectively, that we interact. We estimate these equations via OLS, clustering standard errors at the level of treatment assignment, the district-assembly. Therefore all constituencies within a district and in the same assembly across all three elections are within one cluster. We also consider an instrumental variables approach to estimate the effect of the number of constituencies on our outcomes of interest. In this case, we estimate the same equation as above except we take $Z_{da}$ to be the number of constituencies in a district $d$ and assembly $a$ and instrument for it using the aforementioned binary treatment indicator.

Note that this estimation strategy collapses the treatments at each discontinuity to one treatment indicator and relying on bin fixed effects to partial out heterogeneity across experiments. A similar approach instead estimates bin specific treatment effects and averages across treatment bins, following the logic of analyzing block randomized experiments detailed by (Lin et al., 2013). These results are qualitatively similar and reported in Appendix 1.A. We prefer the approach with bin fixed effects because they allow for the inclusion of province fixed effects as well.

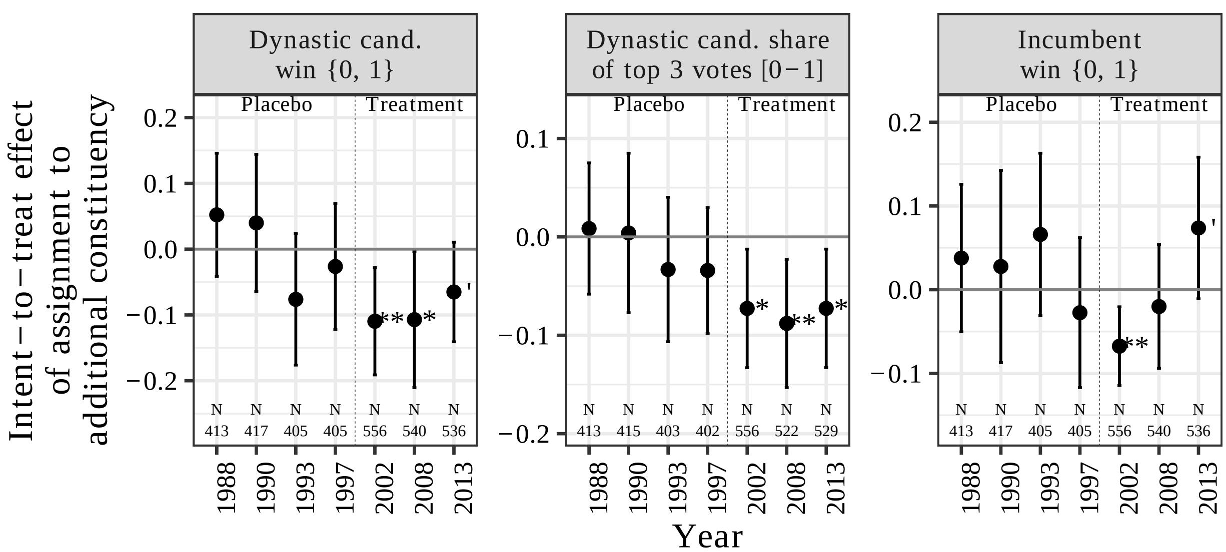
### 1.4.2 Results on dynastic persistence

In Figure 1.3, we present our estimated ITT effects by year. To the left of the dashed line, we estimate effects on years pre-redelimitation as a placebo test to support our identification strategy. For all three of the outcomes we consider here—whether a dynastic candidate won the constituency, the share of the top three vote shares[18], and whether an incumbent of any constituency won the election—there is no statistically significant difference between treated

[18]We do not have data on the dynastic status of candidates that placed 4th or lower in 2013 and thus only consider the top three positions

and control constituencies in the pre-redelimitation period, supporting our notion that the treatment assignment may be unrelated to the potential outcomes of constituencies. In the post-treatment period, the first two panels show a consistent and statistically significant decrease in both the probability a dynastic candidate wins a constituency as well as the share of the votes won by dynastic candidates when the constituency is in a district assigned to an additional seat. In the right-most panel, we show that the share of seats won by incumbents also decreases in districts with more seats, although this effect goes away as time passes. As a district grows in size, the number of incumbents available to win a seat in that district decreases, so this effect is somewhat mechanical. However, there is no clear lasting effect of this redelimitation on the power of incumbents to win elections. On the other hand, dynastic families continue to win at lower rates in following years, suggesting that adding seats to a district has a lasting effect on dynastic power within the district and the legislature.

Figure 1.3: Pseudo-RDD effect of one additional seat on dynastic prevalence by year




†, p-value < 0.1; *, p-value < 0.05; **, p-value < 0.01; ***, p-value < 0.001. . Standard errors are clustered at the district-assembly level. Each unit is a constituency in a particular assembly, and estimates are from separate OLS regressions fit within each election wave (year). The bars represent 95 percent confidence intervals. The four placebo estimates come from years that pre-date the redelimitation and are expected to be zero.

We report pooled ITT effects of assignment to an additional seat on different outcomes in

Panel A of Table 1.3 and the marginal effect of an extra seat using the IV estimation in Panel B. These estimates largely confirm what the election-specific results conveyed. In column (1), we estimate that when a constituency is within a district that is assigned an additional seat, it is 9.6 percentage points less likely to be won by a dynastic candidate. In column (2) we estimate that the share of the top three vote shares won by dynastic candidates is 0.078 lower for constituencies within a district that receives one additional seat. In column (3), the pooled effect of assignment to an additional seat is close to zero, reflecting how the initial dilution of incumbents winning seats goes away as time passes and new incumbents establish themselves. In Panel B, the IV analysis produces effects of the same magnitude—unsurprising given that the effect of assignment to an additional seat nearly perfectly predicts one additional seat per district.

#### Why are dynasts winning less often?

Is the decrease in the share of seats won by dynasts attributable to a decrease in their ability to win a redrawn constituency or to an unwillingness to have a family member challenge in a new one? Using district level outcomes, we show that the number of dynastic candidates winning across treated and control districts is the same and that there is no significant increase in the number of dynastic candidates competing. This indicates that dynastic candidates are not less likely to win when they run, but rather that they are are unable or unwilling to meet the demand for new candidates in new constituencies. In Table 1.4, we collapse our constituency level dataset to the district level but use the same estimation framework as above. We consider three district level outcomes: the share of constituencies won by dynastic candidates, a replication of the main result at the constituency level; the total number of dynastic candidates who win; and the total number of dynastic candidates.

In column (1) of Table 1.4, we replicate the result from the constituency level analysis and estimate that districts assigned to an additional seat have the share of seats won by dynasts decrease by 0.08. However, this decrease in the share of constituencies won by a dynastic

Table 1.3: Effect of assignment to one additional seat on dynastic prevalence

**Panel A: Intent-to-treat Effect**

| Outcome | Dynastic cand. win | Dynastic cand. share of top 3 votes | Incumbent win |
|---|---|---|---|
| | (1) | (2) | (3) |
| ITT | $-0.095^{**}$ | $-0.078^{**}$ | $-0.006$ |
| | (0.033) | (0.026) | (0.021) |
| Treatment Bin FEs | Yes | Yes | Yes |
| Province × Year FEs | Yes | Yes | Yes |
| N | 1632 | 1607 | 1632 |

**Panel B: IV 2SLS of Marginal Constituency**

| Outcome | Dynastic cand. win | Dynastic cand. share of top 3 votes | Incumbent win |
|---|---|---|---|
| | (1) | (2) | (3) |
| LATE | $-0.095^{*}$ | $-0.078^{*}$ | $-0.006$ |
| | (0.037) | (0.030) | (0.023) |
| Treatment Bin FEs | Yes | Yes | Yes |
| Province × Year FEs | Yes | Yes | Yes |
| N | 1632 | 1607 | 1632 |

$^{\dagger}$, p-value $<$ 0.1; $^{*}$, p-value $<$ 0.05; $^{**}$, p-value $<$ 0.01; $^{***}$, p-value $<$ 0.001. Standard errors are clustered at the district-assembly level. Each unit is a constituency in a particular assembly and election. There are 25 constituencies for which vote share data is missing and thus the number of observations drops for the vote share outcome. Panel A is estimated using OLS, and Panel B is estimated using 2SLS, with treatment status instrumenting for the number of actual constituencies in a district.

candidate is not the result of fewer dynastic politicians winning overall. Column (2) shows that, if anything, the number of dynastic candidates who win in treated district increases. Nonetheless, the fact that dynastic candidates are not capturing new seats is likely explained by the result in Column (3), which shows that treated districts have on average only 0.2 more dynastic candidates competing than control districts, and this estimate is statistically indistinguishable from zero. Therefore, it appears that dynasts are either unwilling or unable to find a family member who can win the additional constituency in their district.

Table 1.4: Effect of assignment to one additional seat on dynastic prevalence at the district level

| Outcome | Share cons. won by dynastic cands | N dynastic winners | N dynastic cands |
|---|---|---|---|
| | (1) | (2) | (3) |
| ITT | $-0.080^{\dagger}$ | 0.080 | 0.220 |
| | (0.041) | (0.177) | (0.398) |
| Treatment Bin FEs | Yes | Yes | Yes |
| Province $\times$ Year FEs | Yes | Yes | Yes |
| N | 428 | 428 | 428 |

$^{\dagger}$, p-value $< 0.1$; $^{*}$, p-value $< 0.05$; $^{**}$, p-value $< 0.01$; $^{***}$, p-value $< 0.001$. Standard errors are clustered at the district-assembly level. Each unit is a district in a particular assembly and election.

## 1.5 Effects of education minimum

In this section, we present the effects of education minimums on the kind of candidate that wins subsequent elections.

### 1.5.1 Research design

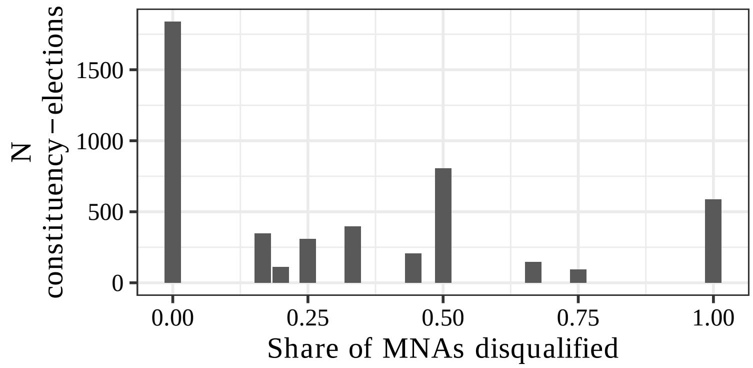
In order to study the effects of the education minimum requirement on the political elite, we study how disqualification of incumbents effects constituency level outcomes. To estimate the effect of disqualification, we employ a difference-in-differences design. To measure disqualification rates, we rely on Afzal (2014) for measurement of which MNAs were disqualified. No similar data exists for MPAs, as their educational status could not be ascertained for a large enough sample. Therefore, our main treatment is the disqualification of MNAs who did not have the requisite Bachelor's degree.

The disqualification of MNAs happens at the National Assembly constituency level; however, due to the redelimitation in 2002, it is difficult to match constituencies across the redelimitation in order to utilize a differences-in-differences design. In her paper, Afzal (2014) manages to match most National Assembly constituencies; however, in this paper we also consider effects on provincial assembly constituencies. Matching individual provincial assemblies to the national assemblies within which they lie is futile as few boundaries map one-to-one. However, we can again utilize the fact that most constituencies are contained within districts.

Therefore, we consider the share of MNAs disqualified within a district as our main treatment variable. However, in some areas—mostly the sparsely populated Balochistan—National Assembly constituencies contain multiple districts rather than the usual case where one district contains multiple constituencies. To accommodate these cases, we define the treatment level as the super-district. The "super-district" is the same as a district when all National Assembly constituencies are within its boundaries and it is the union of multiple districts when, in 1997, one National Assembly covered the entirety of all those districts. For example, in District Lahore, all constituencies in 1997 and 2002 are wholly within the districts boundaries, and thus the treatment is the share of 1997 MNAs in Lahore that were disqualified by the education requirement. However, the districts of Lower Dir and Upper Dir were within one MNA constituency in 1997; thus for all constituencies within Lower and Upper Dir, their treatment status is whether or not the MNA of the one 1997 constituency was disqualified. In this example, the districts of Lower Dir and Upper Dir form one "super-district." There are some cases where 1997 MNA constituencies crossed multiple districts and then new constituencies covered different pairs of districts; these cases are rare, confined to Balochistan, and dropped from our analysis.

In total, for all data from 1988 to 2013, we are able to assign 4,840 constituency-elections (93 percent) to a "super-district." Figure 1.4 presents the distribution of the share of MNAs disqualified in this sample of constituency-elections. Note that because the treatment is the same for all constituency-elections within a super-district, this figure will over-represent larger super-districts. In our analysis we consider results at both the constituency and super-district level.

Figure 1.4: Distribution of the share of MNAs disqualified by the education minimum at the district level




In the case of the constituency level data, we have a time-series cross-section rather than a panel as constituencies change across the 2002 redelimitation. At the super-district level, we have an unbalanced panel as some districts were created after 1988, when our data begins. At both levels, we rely on the assumption that there are no unobservable time-varying confounders between the share of MNAs disqualified in our outcomes and estimate the fixed effects model

$$y_{ipat} = \tau Z_d + \zeta_d + \lambda_p \times \theta_t + \epsilon_{ipat}, \tag{1.2}$$

by OLS, where $y_{ipat}$ is some outcome $y$ in unit $i$, either constituencies or districts depending on the dataset, province $p$, assembly type $a \in \{\text{Provincial}, \text{National}\}$, in election wave $t$; $\tau$ is the treatment effect of interest and $Z_d$ is the share of MNAs disqualified in district $d$ before the 2002 elections, and is set to 0 in the period before 2002; $\lambda_p$ and $\theta_t$ are province and election wave fixed effects, respectively, that we interact. We estimate these equations via OLS, clustering standard errors at the level of treatment assignment, the district. In what follows, we also estimate the same equation using leads and lags of $Z_d$ to test the validity of the research design and to understand how effects changed over time.

### 1.5.2 Results on disqualifications due to education minimums

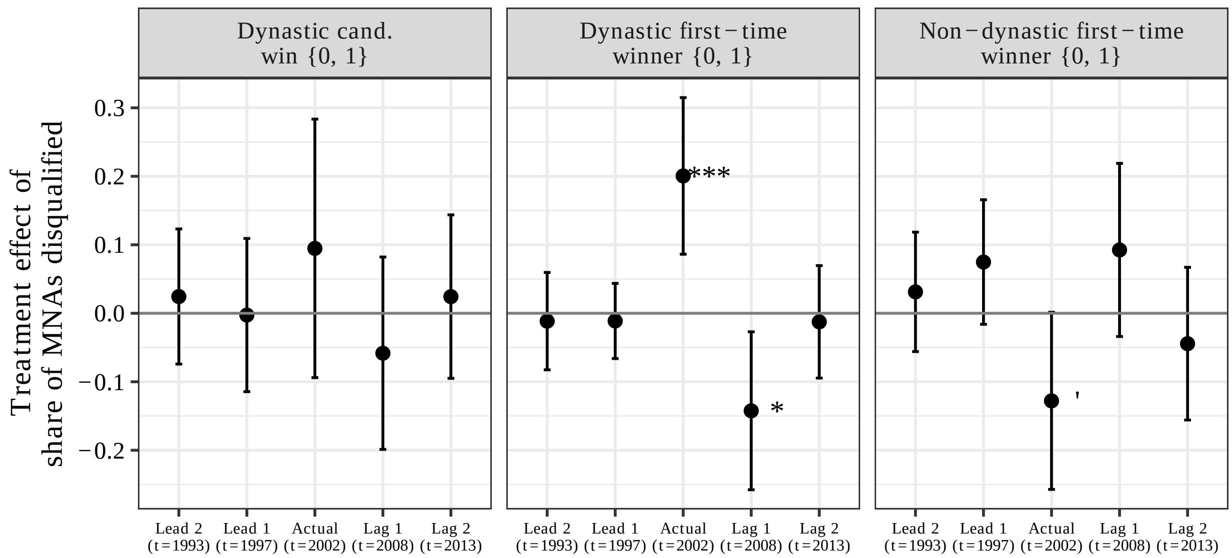
The difference-in-differences estimates at the constituency and super-district level are presented in Table 1.5. In Panel A presents the constituency level results and Panel B presents the super-district level effects. We estimate effects on three outcomes, measured by a binary indicator for the constituency level estimates and by the share of constituencies with the value of one for the super-district level estimates: (1) whether the constituency was won by a dynastic candidate; (2) whether a constituency was won by a first-time dynastic candidate; and (3) whether a constituency was won by a first-time non-dynastic candidate. These outcomes allow us to measure whether family politics are made more common place by candidate restrictions and what kind of candidate is taking the place of disqualified candidates. For each outcome, we estimate Equation 1.2 for the sample that just surrounds the introduction of the education minimum (the 1997 and 2002 elections) and a sample that includes all available data (1988 through 2013). Note that even though the education minimum was removed following the 2008 elections, there may still be lasting effects from the disqualification of incumbents.

As shown in columns (1) and (2) of Table 1.5, we estimate positive effects of disqualification of MNAs on dynastic prevalence, although the results are not statistically significant. While the introduction of education minimums successfully disqualified many incumbents, it appears not to have reduced the share of constituencies won by dynasts. If anything, dynasts may have been more prevalent after the introduction of the education minimum. In columns (2) and (3), we present the effect of the share of MNAs disqualified on the number of constituencies being won by first-time dynastic candidates. At both the constituency and district level, the effect of MNA disqualification on the prevalence of first-time dynastic winners is positive and, in three out of four specifications, statistically significant. Take for example the result in column (3) of Panel A; we estimate that constituencies in super-districts that had all MNAs disqualified experienced an 18.6 percentage point greater increase in the share of constituencies won by dynastic, first-time candidates than constituencies in super-districts where no MNA was disqualified. This sizable effect is replicated at the super-district level in Panel B. Here, the estimate in column (3) indicates that the share of constituencies won by first-time dynastic candidates increased by 14.9 percentage points more in super-districts where all MNAs were disqualified than super-districts where no MNAs were disqualified. This estimate is not statistically significant, largely owing to the small sample size. Column (4), including more election waves for each super-district, presents statistically significant results at both levels, although the effect size is smaller. Meanwhile, columns (5) and (6) estimate negative effects of disqualifying incumbents on the share of non-dynastic first-timers winning elections, although these results are largely statistically insignificant. To explain why including the 2008 and 2013 elections results in smaller effect estimates and to provide evidence for the validity of the differences-in-differences design, we now turn to a model with leads and lags.

Figure 1.5 presents results from estimating Equation 1.2 with two leads and two lags of $Z_d$ included to simultaneously test the plausibility of no time-varying confounders and demonstrate election-specific effects following the introduction of the education minimum. Each panel captures a single regression of one outcome on the five treatment indicators with fixed effects and standard errors are clustered at the super-district level. The left two estimates are the leads; for all outcomes they are statistically indistinguishable from zero and the point estimates are very close to zero for most of the leads. This provides some evidence for the validity of the differences-in-differences design; if there were strong time-varying confounders, then we might be able to estimate effects in the time periods preceding treatment.

Turning to the actual year of treatment, 2002, the estimates in the center panel indicate a strong positive effect of the share of MNAs disqualified on the probability a constituency is won by a dynastic candidate in their first election. Furthermore, as shown in the right panel, there is a negative effect on the share of constituencies being won by non-dynastic first-time candidates, although the estimate is only statistically significant at the 0.1 level.

Figure 1.5: Difference-in-differences estimates of effect of share of MNAs disqualified on dynastic churn with leads and lags




†, p-value < 0.1; *, p-value < 0.05; **, p-value < 0.01; ***, p-value < 0.001. . Standard errors are clustered at the super-district level. The unit of analysis is a constituency and each panel represents one regression for a particular outcome with indicators for two leads, two lags, and the actual year the education minimum was implemented. The bars represent 95 percent confidence intervals.

Table 1.5: Difference-in-differences estimates of effect of share of MNAs disqualified on dynastic churn

**Panel A: Constituency level effects**

| Outcome | Dynastic cand. win | | Dynastic first-time winner | | Non-dynastic first-time | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Share Disq. | 0.075 | 0.073 | 0.186** | 0.089** | −0.131† | −0.013 |
| | (0.099) | (0.063) | (0.060) | (0.034) | (0.068) | (0.039) |
| District FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Province × Year FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Subset | 1997-2002 | 1988-2013 | 1997-2002 | 1988-2013 | 1997-2002 | 1988-2013 |
| N | 1462 | 4994 | 1462 | 4994 | 1462 | 4994 |

**Panel B: Super-district level effects**

| Outcome | Share cons. won by dynastic cands | | Share cons. won by first-time dynastic cands | | Share cons. won by first-time non-dynastic cands | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Share Disq. | 0.030 | 0.038 | 0.149 | 0.072* | −0.124 | −0.031 |
| | (0.149) | (0.061) | (0.100) | (0.033) | (0.107) | (0.046) |
| District FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Province × Year FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Subset | 1997-2002 | 1988-2013 | 1997-2002 | 1988-2013 | 1997-2002 | 1988-2013 |
| N | 190 | 626 | 190 | 626 | 190 | 626 |

†, p-value $< 0.1$; *, p-value $< 0.05$; **, p-value $< 0.01$; ***, p-value $< 0.001$. Standard errors are clustered at the super-district level. Panel A is estimated on a time-series cross-section of constituencies and Panel B is estimated on an unbalanced panel. Columns (1), (3), and (5) use a dataset that ends in 2002, while columns (2), (4), and (6) extend until 2013.

This indicates that the education minimum was creating space for new candidates to succeed as long as they were connected to previously elected politicians; the share of unconnected ("outsider") candidates winning elections decreased in constituencies where incumbents were being disqualified. The education minimum was hurting incumbents, but incumbents were able to circumvent these problems by putting forth family members. It appears that the small parties who thought that the education minimum would make it harder for new candidates to emerge and challenge the elite (e.g. Afzal, 2014, p. 55) was substantiated.

Lastly, turning to the lags, the pattern reverses somewhat. The center and right panels show that areas with greater disqualification rates experienced a greater increase in first-time dynastic winners and a greater decrease in first-time non-dynastic winners in 2008, although only the former result is statistically significant. However, the left panel shows that there was little estimable change in dynastic prevalence throughout this whole period. These results indicate that the introduction of education minimums caused a wave of family members entering office in 2002 who likely remained in office in 2008—despite the greater decrease in new family members winning in treated areas, dynastic prevalence overall stayed the same. This implies that the introduction of the 2002 education minimum caused churning within the incumbent elite—the disqualified were able to replace themselves with their family members—and then those family members were able to stay in office. By 2013 the education minimum had been removed by a Supreme Court ruling. Nonetheless, we see little in the way of effects in 2013, likely because a new set of dynastic politicians had already solidified their position and there was little incentive for the previously disqualified to re-enter politics.

## 1.6 Mechanisms

Institutional reforms may affect political dynasties in many ways, affecting either the willingness or ability of political families to run and win elections. In this section, we suggest discuss two possible mechanisms for why we observe dynastic prevalence decreasing following reapportionment, but staying steady and responding actively to disqualification via education minimums. First, we argue that political families face diminishing marginal returns to the number of elected family members and thus will respond differently to different institutional reforms. These diminishing marginal returns suggest that political families will be more likely to replace a disqualified incumbent—as in the case with the education requirements—than branch out in a district with more seats to capture a second constituency. Second, we argue that the investments made by political families in local organizations may be an important part of their electoral success, but are particularly vulnerable to redelimitation and reapportionment.

### 1.6.1 Diminishing marginal returns to elected family members

There exists both a personal and a family level motivation for a family member to enter politics. At the individual level a family may see a career in politics as more profitable or meaningful due to their natural advantages as a candidate or intrinsic motivation from exposure to politics at a young age. At the family level, recognition that political connections are often very valuable may lead to candidate entry. Politically connected firms frequently profit from having political connections in Pakistan (Khwaja and Mian, 2005) and elsewhere (e.g. Fisman, 2001; Szakonyi, 2018). In addition to firms benefiting, individuals in political families also benefit—members of political families achieve better outcomes in the private sector labor market (Manacorda and Gagliarducci, 2020), are more present in the bureaucracy (Fafchamps and Labonne, 2017), and are better paid even without obvious signs of nepotism (Folke, Persson and Rickne, 2017). This literature largely studies whether connections to at least one politician yields benefits for families and firms; the marginal effect of having a second connected politician on private returns is likely much lower than having just one family member in office.

As a consequence, political families may be more interested in maintaining at least one member of the family in politics than in securing a second position within a legislature.

Diminishing marginal returns to a second elected family member are also likely outweighed by the relative benefit of that family member positioning themselves in the business, the bureaucracy, or another sector. For example, a member of a historically powerful political family in Pakistan was disqualified shortly after being elected in the 2013 general elections for having lied on their candidate affidavit forms. In the subsequent bye-election, the disqualified legislator's brother left his overseas business interests to return to Pakistan and run—and win—in the same constituency. In the brother's words, he simply had no choice but to return because the family would always need to maintain a political presence in their district (Sonnet, 2017). Before his brother was disqualified, he was not interested in running for office and was instead in charge of managing the family's economic interests.

Indeed, on the whole in Pakistan, families largely only run one candidate at a time. From 1985 to 2008, years for which we have the family relations of all candidates for the legislature in Pakistan, only 23.0 percent of families ever ran two or more candidates in the same election and only 9.4 percent of the time did more than one member of a family win an election. Running multiple candidates is likely an effort to ensure at least one candidate wins rather than getting an effort to secure multiple seats. This strategy of a family hedging their bets is supported by the fact that over 55 percent of families that run more than one candidate, do so with candidates running in different parties. Furthermore, 72.1 percent of families that run more than one candidate have candidates running for both national and provincial assembly positions. Families that do run multiple candidates, although rare, appear more interested in getting one member in multiple assemblies than two members or more in one assembly.

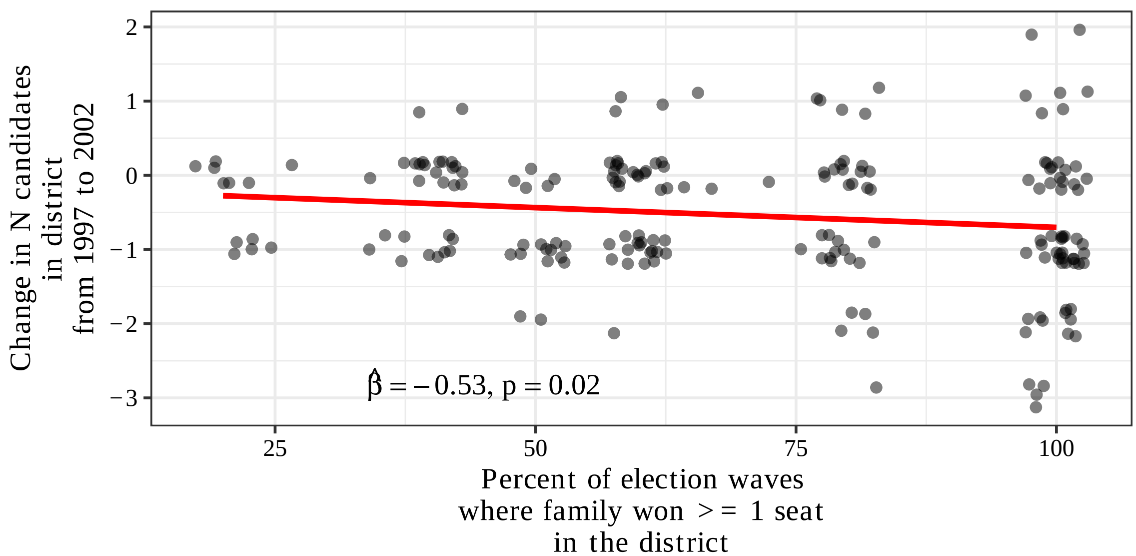
The large marginal returns to an elected family member, and the potential that these marginal returns are much lower for the second elected family member, may explain why political families respond to disqualification via the education requirement. Disqualification of the sole elected family member threatens the access of the family to a legislature. On the other hand, while the addition of new constituencies in nearby areas may dilute their power within the legislature, the main prize of maintaining an elected family member is not threatened.

### 1.6.2 Redelimitation and local organizations

Even if a family were interested in running a candidate in a new constituency, redelimitation may actually hurt their ability to do so if it disrupts one of their key advantages: local organization. Most explanations for the causes of the dynastic advantage point to name recognition and the strength of local organizations as key contributors (Feinstein, 2010; Rossi, 2017; Smith, 2018, e.g.). Redelimitation, through redrawing constituency boundaries, may disproportionately affect dynasts by disrupting the relationship between dynasts and investments in local organizations. If this were true, then reapportionment and redelimitation should be more damaging to the competitiveness of political families than other candidates and should hurt their performance post-redelimitation.

We provide some evidence for this relationship by testing whether families with longer histories in one district are more adversely affected by the 2002 reapportionment. Figure 1.6 plots the change in number of candidates a family runs in a district from 1997 to 2002 by the history of that family in the district, subsetting to all family-districts where the family ran two different candidates at some point before 2002 and won an election in 1997. In other words, this contains all incumbent families before the 2002 elections. The horizontal axis is the percent of election waves that a family won at least one constituency in the district and the vertical axis is the change in the number of candidates that a family runs in the district. All of the points are families in districts, jittered as the values on each axis are discrete. The downward slope of an associated linear model reveals that the stronger an incumbent family's history in a district, the fewer additional candidates they run in 2002 compared to 1997. Families with shorter histories are less adversely affected by the reapportionment and redelimitation. We leave more direct tests of this hypothesis for future work, where more direct measurement of redelimitation and local organization can shed more light on the relationship between the two.

Figure 1.6: Change in number of candidates a family runs in 2002 by history in district




Each point is a family-district pair for all incumbent families before the 2002 elections. An incumbent family is a family that has run at least 2 candidates from 1985 to 1997 and won a constituency in the district in 1997. The horizontal axis is the percent of election waves that a family won at least one constituency in the district and the vertical axis is the change in the number of candidates that a family runs in the district. The beta coefficient is from a regression of the change in candidates on the history of the family in the district, with standard errors clustered at the family level.

## 1.7 Conclusion

In this paper, we demonstrate that two different electoral reforms, reapportionment and the introduction of a education minimum for candidacy, had different consequences for the prevalence of dynastic politicians in Pakistani legislatures. Reapportionment, through increasing the number of constituencies within a district and redrawing constituency boundaries, diluted the share of constituencies won by dynastic candidates. Rather than seizing new constituencies and consolidating power when new seats were created, political families ran the same number of candidates, won a similar number of seats, but saw their share of each legislature decrease. On the other hand, when incumbents are disqualified for not meeting education minimums, political families respond to ensure their presence in the legislature endures. District where incumbents were disqualified at higher rates saw a greater share of constituencies won by first-time dynastic candidates.

Why are political dynasties susceptible to reapportionment and redelimitation? We argue that political families face strong incentives to ensure at least one family member is elected, but have fewer reasons to win a second seat when the number of legislators is increased. There are other possible explanations, such as redelimitation hurting the power of local networks to advantage political dynasties or whether political dynasties are somehow disadvantaged in the actual decision about how and where to redraw boundaries. However, we leave explorations of the micro-foundations of elite response to institutional reform for future work.

While Musharraf's explicit efforts to curtail the power of dynastic families in the legislature may have come up short in some ways, these results indicate that the "iron law of oligarchy" can be bent with the right set of tools. The differential response by dynasts to these two reforms has clear implications for how institutional design can curtail the power of the incumbent elite. As with term limits (Labonne, Parsa and Querubin, 2019) and seat reservations (Auerbach and Ziegfeld, 2020), education minimums are easily navigable by powerful incumbents. In the case of Pakistan, disqualified incumbents are able to get family members elected in their stead, and the share of the legislature captured by a particular kind of elite, political families, is unchanged. However, increasing the size of the legislature may have lasting consequences on the prevalence of elites in government.

# Appendix

## 1.A Robustness of reapportionment effects

This section contains alternative specifications and robustness checks for the estimated treatment effect of reapportionment on dynastic persistence.

Table 1.A.1 presents the treatment effect of one additional constituency within a district on dynastic prevalence using the following model

$$y_{cdpat} = \tau Z_{da} \times \text{Bin}^{c}_{da} + \epsilon_{cdpat}, \tag{1.A.1}$$

where $y_{cdpat}$ is some outcome $y$ in constituency $c$, district $d$, province $p$, assembly type $a \in \{\text{Provincial}, \text{National}\}$, in election wave $t$; $\tau$ is the treatment effect of interest and $Z_{da}$ is a treatment indicator for whether the district $d$ is just above the $x.5$ threshold for the seat assignment for assembly $a$; $\text{Bin}^{c}_{da}$ is the *centered* treatment bin fixed effect for district $d$ in assembly type $a$, interacted with the treatment indicator. Standard errors are clustered at the district-assembly level once more.

By centering the treatment bin indicators and interacting them with the treatment indicator itself, this estimator follows (Lin et al., 2013) and essentially averages over bin-specific treatment effects, weighting by the size of the bins. As a result, this estimator is more tightly linked to the design of the natural experiment as it relies on averaging the individual "experiments"; however, it precludes the use of province times year fixed effects. In any case, the results here are of similar magnitude and have similar uncertainty to those reported in Table 1.3.

Table 1.A.1: Effect of assignment to one additional seat on dynastic prevalence, averaging across bins

| Outcome | Dynastic cand. win | Dynastic cand. share of top 3 votes | Incumbent win |
|---|---|---|---|
| | (1) | (2) | (3) |
| ITT | $-0.076^{*}$ | $-0.065^{*}$ | 0.003 |
| | (0.029) | (0.025) | (0.020) |
| N | 1632 | 1607 | 1632 |

$^{\dagger}$, p-value $<$ 0.1; $^{*}$, p-value $<$ 0.05; $^{**}$, p-value $<$ 0.01; $^{***}$, p-value $<$ 0.001. Standard errors are clustered at the district-assembly level. Each unit is a constituency in a particular assembly and election. There are 25 constituencies for which vote share data is missing and thus the number of observations drops for the vote share outcome. ITT reported is $\tau$ in Equation 1.A.1.

# CHAPTER 2

# The Dynastic Advantage in Close Elections in Four Countries: Implications for Regression Discontinuity Designs

**Abstract:** Close elections regression discontinuity (RD) designs are widely used to study the causal effect of winning an election. Most researchers employing close elections RD designs assume that the potential outcomes of candidates are smooth functions of their vote margins when the vote margin is zero; in other words, they assume candidates who barely win elections are comparable to candidates who barely lose. The validity of this assumption in the context of close elections has been studied in several papers, largely focusing on whether incumbents are more likely to barely win elections, thus violating this assumption. In this paper, I demonstrate that a more difficult to observe covariate, dynastic status—i.e., whether one has a family member precede them in office—is imbalanced across close winners and close losers. Evidence for this imbalance exists in Japan, Pakistan, the Philippines, and the United States. On average, winners of close elections are roughly 6 percentage points more likely to be dynastic candidates and evidence of imbalance exists in each country. Sensitivity analyses reveal that 9 to 40 percent of close election RD papers in top political science and economics journals would have their most robust estimate quartered by this imbalance, if dynastic status correlates with their outcome of interest as much as it correlates who wins close elections.

## 2.1 Introduction

Regression discontinuity (RD) designs are widely used in political science to study the causal effects of winning an election. Generally, this involves studying the effects of winning on outcomes for the winning candidate or party (Lee, Moretti and Butler, 2004; Dal Bó, Dal Bó and Snyder, 2009, e.g.), or the effects of a candidate with a particular set of characteristics winning an election (Hall, 2015; Brollo and Troiano, 2016, e.g.). While RDs appear in many other contexts in the social sciences, close elections have received a considerable amount of attention. In the RD framework, close elections provide an opportunity to control for unobservable characteristics of candidates and to identify the causal effect of winning an election. There are two different sets of assumptions invoked by researchers. Some argue that winners and losers are decided as-if random in very close elections and thus the effects of winning are identified within this bandwidth. Others invoke the continuity assumption, whereby the potential outcomes of candidates are assumed to be continuous across the cutoff—in other words, there is no discontinuity in the potential outcomes when the margin of victory is zero. If the assumptions invoked are satisfied and the estimation strategy is appropriate—another source of contention in the literature but not the main focus here—one can use close elections RD designs to estimate the causal effect of barely winning an election.

Naturally, the use of these quasi-experiments to study causal effects has created a debate over their validity. For example, Caughey and Sekhon (2011) demonstrate that incumbency and several other correlated variables are imbalanced in close elections, with incumbents far more likely to win close elections. In response, Eggers et al. (2015) demonstrate that this imbalance is present only in one sample, the 20th century United States House of Representatives, and that in other countries and elections this imbalance does not exist nor does it exist for a host of other important pre-treatment covariates. Further evidence from (De la Cuesta and Imai, 2016) suggests that this imbalance is also highly sensitive to the specification employed in the balance tests. In this paper, I contribute to this conversation by studying a potentially important and difficult to measure covariate of electoral candidates: their membership in political families. I demonstrate that dynastic candidates—those who have been preceded by a family member at the same or similar level of politics—are more likely to win very close elections when pooling data across four different countries and decades of elections. Furthermore, this result comes from data wherein incumbency is held constant and therefore cannot explain the imbalance.

The importance of dynastic politics and the persistence of power within families is well documented in the political economy literature. These effects seem especially strong in candidate-centered environments (Dal Bó, Dal Bó and Snyder, 2009; Querubin et al., 2016), as some evidence shows this effect may not extend to more party-centric environments (Fiva and Smith, 2018; Van Coppenolle, 2017). There are a variety of reasons why having family members precede a candidate in politics makes that candidate more likely to win office. Many potential pathways for this perpetuation of power have been explored, including name recognition (Rossi, 2017; Dal Bó, Dal Bó and Snyder, 2009), financial advantages (Feinstein, 2010; Dal Bó, Dal Bó and Snyder, 2009), and connections to relevant party actors who control nominations (Smith, 2018; Chhibber, 2013; Chandra, 2016).

This paper demonstrates that this advantage is observable even in close elections. Evidence for an imbalance in the dynastic status of close winners and close losers exists in Japan, Pakistan, the Philippines, and the United States and employing a variety of specifications. On average, point estimates indicate that a candidate winning one of their early-career elections is roughly 3 to 11 percentage points more likely to be a dynastic candidate than a candidate who loses. This result presents what is essentially a failed placebo test across multiple contexts. In doing so, it makes a contribution to the literature on the validity of cross elections RD designs broadly. If the RD close elections design were valid, then winning a close election should have no bearing on the pre-election covariates of candidates; in other words, pre-election covariates should be balanced. Previous work has pointed to the possibility that incumbency status is imbalanced across close winners and close losers (Caughey and Sekhon, 2011) and this paper adds a second, related covariate, that is more difficult to collect for most researchers: the dynastic status of all candidates. By estimating positive and statistically significant estimates in the pooled sample and in every country-specific sample, I argue that researchers who employ close elections RD designs should carefully consider whether family connections could cause imbalances in their sample.

To quantify the magnitude of this imbalance and consider how analyses that omit dynasticism may be biased, I conduct a meta-analysis of the sensitivity of published results to the dynastic imbalance estimated in this paper. I collect estimates from close elections RD designs published in three top political science and five top economics journals. Between 9 and 40 percent of the most robust estimates in these papers would be quartered if the dynastic imbalance in close elections holds in their sample and if dynasticism is as important to their outcome of interest as it is to winning close elections. It is important to note that this does not mean these papers are necessarily biased; only that if dynasticism matters in their context for who wins close elections in their contexts, it may only take a small relationship between dynasticism and the outcome of interest to bias their results. For example, there are a series of studies focusing on the financial returns to office. If members of political families are wealthier and more likely to win close elections, than these studies will overstate the effect of holding office on wealth.

In the next section, I explain how we measure dynasticism and demonstrate the basic patterns of dynastic advantage in all elections and in close elections. I then review the main RD assumptions and introduce the placebo tests I use, before using this framework to document the imbalance in dynastic status across winners and losers. Finally, I turn to the sensitivity of published results to this imbalance before concluding.

## 2.2 Measuring dynasticism

I use four available candidate-level datasets that have the margin of victory for candidates and enough data to determine whether these candidates were preceded in office by a family member. Because collecting data on family connections of politicians is costly, there are only a few datasets that fit this criteria. I use data from Japan from 1947-2014 (Smith and Reed, 2018), Pakistan from 1985-2013 (Cheema, Javid and Naseer, 2013), the Philippines from 1946-2007 (Querubin et al., 2016), and the United States from 1788-1988 (Dal Bó, Dal Bó and Snyder, 2009). Table 2.1 summarizes the data used in the analysis below, as well as the kinds of elections for which data are available.

In the United States, only data on the first re-election attempt by a candidate exists (Dal Bó, Dal Bó and Snyder (2009) collected data on incumbents, not challengers), while in the other three countries data on a candidate's first election attempt is available. The focus on first elections (or re-elections) ensures that none of the units of analysis are incumbents (or all of them are single term incumbents, in the case of the United States), following arguments (Dal Bó, Dal Bó and Snyder, 2009; Querubin et al., 2016) and evidence (Caughey and Sekhon, 2011) that incumbency can predict winners of close elections.[1] I also only consider winning candidates and the best placed loser in each constituency in order to limit the possibility of including third and lower ranked candidates who have little chance of becoming treated as they are ranked below at least one better placed loser. The final dataset, then, is a sample of *early career* candidates from a wide variety of countries and time periods.

There is ample heterogeneity in this sample. The three Asian countries vary in their level of development over the time period captured, as well as having distinct religious and cultural backgrounds. Furthermore, this dataset includes nations with varying levels of dynastic presence in elections, from a low of eight percent in the United States to a high of

[1] Of course, party incumbency may still be playing a role here, although see evidence that a large part of the incumbency advantage is personal rather than partisan (Fowler and Hall, 2014)

Table 2.1: Data summary

| | Japan | Pakistan | Philippines | USA |
|---|---|---|---|---|
| Election Years | 1947-2014 | 1985-2013 | 1946-2007 | 1788-1988 |
| Election Attempt | 1st election | 1st election | 1st election | 1st reelection |
| Election Level | Nat'l lower house | Nat'l + provincial lower houses | Governor + nat'l lower house | Both nat'l houses |
| N Elections | 25 | 8 | 22 | 101 |
| N Candidates | 3123 | 4626 | 2690 | 6257 |
| % Dynastic Candidates | 14% | 20% | 29% | 8% |
| % Winning Candidates | 60% | 46% | 40% | 79% |
| % Dynastic Candidates who Win | 70% | 62% | 52% | 84% |

The Japan data come from (Smith and Reed, 2018), the Pakistan data were originally reported as (Cheema and Naseer, 2013) and are also discussed in (Cheema, Naseer and Sonnet, 2020), the Philippines data come from (Querubin et al., 2016), and the United States data come from (Dal Bó, Dal Bó and Snyder, 2009).

29 percent in the Philippines. Lastly, the elections considered for Pakistan, the Philippines, and the United States are first past the post, single-member district elections, while the elections considered in Japan are multi-member district elections.

There are two key variables in this analysis: *dynastic*, measuring whether or not the candidate is preceded in office by a family member; and the *margin* of victory for a candidate, measuring is how much a candidate beat the best-placed loser by if they won, or how much a candidate lost to the worst-placed winner if they lost.

In order to code whether a candidate is dynastic or not, I first must define family as well as what it means for a relative to have won office. In general, I follow the original author's rules for coding this variable, and present quotes from their original papers or codebooks defining the rule in Table 2.2. A variety of relationships by blood (parents, siblings, uncles and aunts, first cousins) and close marital relationships (spouses, brothers- and sisters-in-law) are admitted as relatives, although the data from the Philippines relies on a precise naming convention combined with geographical restrictions to identify family. For coding whether a relative preceded a candidate in office, most definitions require a relative to have won either the same office or an office at a similar level as the candidate in a previous election. This is largely a result of data constraints. For example, a candidate for the United States House whose father served in the Senate would be coded as dynastic, while the same candidate would not be coded as dynastic if their father was a mayor, or member of the state legislature. For the cases of Pakistan, the Philippines, and the United States, the relative must have won at the same level as the candidate, while in Japan a few extra executive positions are coded as family experience.

Table 2.2: Definition of Family

| Country | Rule for *Dynastic* = 1 |
|---|---|
| Japan | *Dynastic* = 1 if "the candidate is related by blood or marriage (e.g., child, grandchild, sibling, spouse, son-in-law, or other such close relative) to a politician who had previously served in the national legislature (either chamber), or executive (cabinet), regardless of district" (Smith and Reed, 2018, codebook p. 15). |
| Pakistan | *Dynastic* = 1 if the candidate is related by blood or marriage (child, grandchild, sibling, aunt, uncle, nephew, niece, spouse, son-in-law, daughter-in-law) to a politician who had previously served in either the national lower house or provincial legislature. |
| Philippines | *Dynastic* = 1 if a relative, "identified by finding a match of the middle name, last name, or husbands last name within the same province", was previously an incumbent governors or [...] congressmen within the same province" (Querubin et al., 2016, p. 158).[a] |
| United States | *Dynastic* = 1 if a relative by blood or marriage (e.g. parent, grandparent, great-grandparent, sibling, uncle, cousin, spouse) previously served in either chamber of the the United States Congress. Most of the data come from parents, uncles, siblings, and cousins, although see Table A1 of Dal Bó, Dal Bó and Snyder (2009) for more detail. |

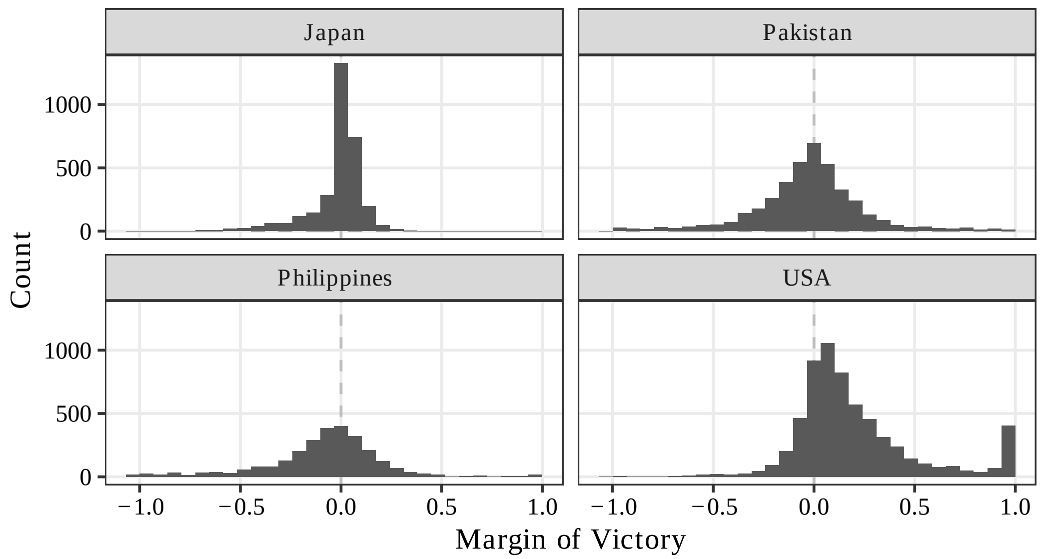
*a*. For more about matching families in Philippines, and the robustness and verification of this measure, see pages 158 to 160 in Querubin et al. (2016). Essentially, the author relies on a traditional naming convention to identify the ties between individuals.

These definitions introduce some heterogeneity in the definition of dynastic candidates, but on the whole they capture similar phenomena and rely on the judgment of researchers with strong area knowledge. Furthermore, this heterogeneity is mainly of consequence for interpreting differences in the results across countries, not of the overall pattern in the data. Only if the variation in who gets "mistakenly" coded as a dynastic candidate by country varies systematically with the probability they win very close elections would this matter for estimates of imbalance.

When coding the margin of victory, the definition for the three single-member district cases (Pakistan, the Philippines, and the US) is identical. For candidates who won their constituency, the margin of victory is their vote share minus the vote share of the first runner-up. For candidates who lost the election, the margin of victory is their vote share minus the vote share of the winner. In Japan, the electoral system is more complicated. I consider candidates competing in the multi-member districts and omit candidates who only compete on party lists. I code the margin of victory for those who win one of the seats in the multi-member district as the candidate's vote share minus the vote share of the best-placed candidate who did not win a seat. Similarly, for those who do not win seats, their margin of victory is their vote share minus the vote share of the worst-placed candidate who won a seat. Therefore, across all countries, the margin of victory for losers captures the distance to the lowest vote share needed to tie a seat getter, and for winners it captures the distance to candidate who they just beat to the last seat available in the constituency. Figure 2.1 displays a histogram of the margins of victory by country.

The United States sample has a greater share of winners (margins of victory greater than zero) and a peak at 1, owing to the sample being made up of first time reelection attempts, rather than first time candidates. Furthermore, the greater concentration of mass around 0 in Japan is a result of the multi-member districts where the winning cutoff tends to be lower than in the single-member district countries. In general, most candidates in this sample have winning margins under 10 percentage point, but there are also candidates winning in land slides across countries.

Figure 2.1: Distribution of margins of victory by country

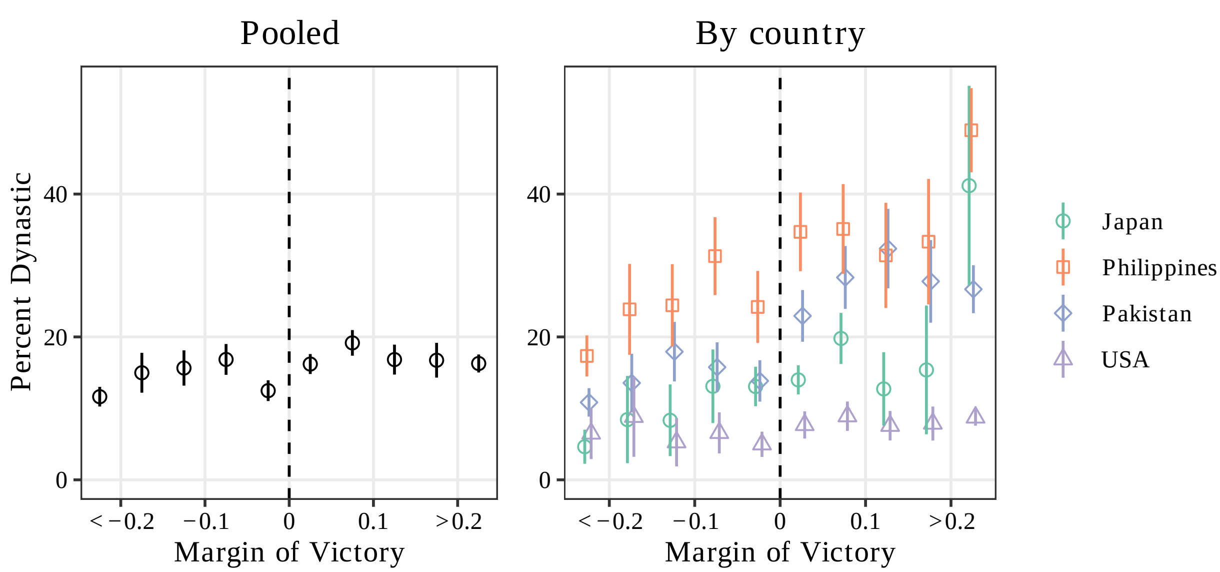



Each panel is a histogram of individual candidates and the margin of victory is the difference between themselves and the worst-placed winner if they are a loser, or themselves and the best-placed loser if they are a winner.

## 2.3 The dynastic advantage in close elections

Do dynastic candidates have an advantage in very close elections? I first explore some basic relationships before testing this relationship using the regression discontinuity (RD) framework. I present the raw relationship between the vote share for early career candidates and the probability they are dynastic candidates in Figure 2.2. On the horizontal axis, candidates are grouped by their margin of victory into 10 percentage point bins. On the vertical axis, I plot the percent of candidates in the corresponding margin of victory bin that are dynastic. In the left panel are the pooled results, while in the right panel the data are displayed by country. In general, as the margin of victory increases and one moves right along the horizontal axis, a greater share of candidates are dynastic. This demonstrates a positive relationship between a candidate's margin of victory and having a family member precede them in office. Furthermore, there is initial evidence of a discontinuity in this graph. Using the pooled data to compare those who lost their election by fewer than 10 percentage points and those who won their election by fewer than 10 (candidates just around the cutoff) there appears to be a jump in the number of dynastic candidates. Among candidates who just lost an election by fewer than 10 percentage points, 13.4 percent were dynastic candidates, while those who just won are dynastic candidates 16.8 percent of the time. This pattern also appears in the disaggregated data, although most starkly for Pakistan. In Pakistan, represented by the light blue diamonds in the right panel of Figure 2.2, candidates who lost by less than 10 percentage points were dynastic 12.7 percent of the time but those who won by less than 10 percentage points were dynastic candidates 22.7 percent of the time.

Figure 2.2: Percent of Dynastic Candidates by Margin of Victory




This figure displays the average share of candidates that are dynastic by margin of victory bins. Each bin is five percentage points wide, so the first bin to the left of the cutoff contains all candidates who lost by zero to five percentage points.

### 2.3.1 Close election regression discontinuity designs

To test this relationship more rigorously, I turn to placebo tests commonly used to validate RD designs. RD designs are used to study the effect of treatments where units become treated as soon as they cross some threshold in a running variable. In political science, RD designs are often employed to study the effect of barely winning an election (e.g. Lee, Moretti and Butler, 2004). Who wins an election is argued to be as-if random if the margin of victory is close enough. Given this assumption, researchers estimate the causal effect of barely winning an election on a variety of downstream outcomes without worrying that there are systematic differences between candidates who barely won elections and candidates who barely lost elections. A similar and weaker assumption is also often invoked: the validity of the RD design hinges on the continuity of the potential outcomes at the cutoff value. In other words, a unit's potential outcomes under treatment and potential outcomes under control must not be discontinuous at the RD threshold to infer that any post-treatment differences are attributable to treatment status. RD designs target the following estimand, the local average treatment effect (LATE) at the cutoff $c$ in running variable $Z$,

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)|Z_i = c], \tag{2.1}$$

where $Y_i(1)$ is the treatment potential outcome for unit $i$ and $Y_i(0)$ is the control potential outcome for unit $i$. This is equivalent to

$$\tau = \lim_{Z_i \downarrow c} \mathbb{E}[Y_i(1)|Z_i = c] - \lim_{Z_i \uparrow c} \mathbb{E}[Y_i(0)|Z_i = c], \tag{2.2}$$

where the necessity of the continuity assumption of the potential outcomes becomes more obvious; each term on the right hand side of Equation 2.2 would be impossible to unbiasedly estimate if the values of $Y_i(1)$ and $Y_i(0)$ that are observable jump right as they become unobservable. For example, if $Y_i(1)$ is lower for units just below the cutoff, then $\lim_{Z_i \downarrow c} \mathbb{E}[Y_i(1)|Z_i = c]$ cannot be unbiasedly estimated using just the treated units.

There are two chief potential violations of the validity of regression discontinuity designs: (1) there is a genuine discontinuity at the cutoff and the continuity assumption is invalid; (2) the continuity assumption may be valid, but there is not have enough data or the correct functional form to unbiasedly estimate the terms on the right hand side of Equation 2.2. The continuity assumption may be invalid if individuals are able to sort themselves to be just above a cutoff in order to get into the treatment group. In the case of elections, it could be the case that certain candidates are able to utilize election fraud and can change the results of the election such that they barely win when otherwise they would have lost. However, there also may be a lack of data, or an incorrectly specified estimator, that leads to bias in estimated treatment effects, even if the continuity assumption holds. In the case of close elections, if there are not enough elections that are close to the cutoff, then the functional form assumptions made on either side of the cutoff are increasingly important. It may be that random events (or more precisely, events unrelated to the potential outcomes of interest) may influence a few votes, but if the number of votes that are influenced is very small, then the set of elections for which the RD close elections design is valid is vanishingly small as one would have to know how to extrapolate from non-close elections to close elections. Unfortunately, balance tests like those shown in this paper cannot necessarily distinguish between violations of the continuity assumption and a lack of data in the bandwidth close enough to the cutoff that the treatment effect of interest is appropriately estimable. However, in what follows, I show that there is an imbalance for a wide variety of cutoffs and specifications, all of which are common in the literature.

#### 2.3.1.1 Validating close election designs

There are two tests that are commonly used to validate regression discontinuity designs: placebo balance tests and density tests. Placebo balance tests seek to falsify the continuity assumption by checking whether a variety of observable, pre-treatment characteristics are continuous at the cutoff. The argument goes as follows: if the RD design and estimation reveal "treatment effects" in variables that a treatment could not possibly affect, then the potential outcomes may similarly be discontinuous. These implausible "treatment effects" are analogous to imbalances in treatment and control samples. For example, if candidates who barely won elections had higher pre-election salaries, then the placebo test would estimate the impossible "treatment effect" of winning on pre-election salaries. If candidates with higher salaries have different average potential outcomes, then the potential outcomes will also be discontinuous at the cutoff. Thus, it is common to test for the validity of RD designs using a host of placebo tests for balance on variables that temporally precede the treatment and are therefore plausibly unaffected by it.

In addition to providing auxiliary evidence for the identifying assumption of RD designs, these placebo tests are also used to question the validity of RD designs. In their paper on close elections in the United States House, Caughey and Sekhon (2011) show that constituencies barely won by Democrats are more likely to have a Democratic incumbent. In other words, Democrats win close elections more often when they are the incumbent. This is a failed placebo test where a Democrat winning an election implausibly "causes" the Democratic party to have won the previous election in that constituency. They use this result to caution against using election results for regression discontinuities, at least in the context they consider. In response, Eggers et al. (2015) run similar placebo tests and demonstrate that this imbalance is detectable only in the U.S. during the post-war period. In other countries and time periods, or when one controls for party incumbency, there is no evidence that candidates from the incumbent party are more likely to win close elections.

First introduced by McCrary (2008), density tests check whether units are more tightly clustered on one side of the cutoff than the other. If units are sorting to one side of the cutoff, then it is plausible that units are manipulating the running variable and those that are able to do have different potential outcomes than those who are unable to sort. As a result, the assumption of continuity of potential outcomes is implausible if units are sorting around the cutoff. In the case of close elections RD designs, winners and losers will be balanced across the cutoff by design in a two-candidate contest. Thus, this test is only useful for close elections RD designs if one consider imbalances within some group, or if the treatment you care about is what type of candidate wins a close election. If there is evidence that some subgroup is sorting to one side or another then there similarly is evidence that the running variable is being manipulated and the potential outcomes may be discontinuous.

### 2.3.2 Failed placebo test of the effect of winning on dynastic status

In this paper, I rely chiefly on placebo tests to document the imbalance in dynastic status among close winners and losers. I prefer these to density tests as we are able to use all of the data in one estimation framework and be more transparent about how the magnitude of the imbalance as the RD estimation strategy changes. Nonetheless, I also conduct density tests within subgroup and report the results in Appendix 2.A. The results are qualitatively similar, although weaker overall.[2]

There are many estimators for causal effects within the RD framework, all of which imply slightly different functional forms for the potential outcomes. Estimates are often sensitive to the choice of estimator due to the lack of data near the cutoff and it is generally accepted that one should report a wide variety of estimation strategies. A common way to estimate the LATE in an RD framework is to estimate functions on either side of the cutoff and then measure the difference between these functions evaluated at the cutoff. This is often achieved by regressing the outcome on an indicator for being on the treated side of the threshold interacted with some function of the running variable, allowing for different functional forms on either side of the cutoff. The coefficient on the indicator for treatment status is an estimate of the LATE if the cutoff is set to zero on the running variable.

In this paper, I estimate the following equation,

$$\text{Dynastic}_i = \tau \text{Win}_i + f(\text{Margin}_i) + \lambda_t + \zeta_r + \epsilon_i, \tag{2.3}$$

where $\text{Dynastic}_i$ is a binary variable for whether the candidate is dynastic, $\text{Win}_i$ is whether

[2] The weaker results may be the result of having to do the analysis by sub-group. The imbalance of interest is both dynastic candidates winning more often and non-dynastic candidates losing more often. Having to do the density check by subgroup hides the fact that the difference between discontinuities in density across the threshold is really what matters.

the candidate wins a seat in their election, $\lambda_t$ is a decade fixed effect, $\zeta_r$ is a region fixed effect, and $\epsilon_i$ is a unit-level error. Here $f(\text{Margin}_i)$ is the function of the candidates margin of victory $\text{Margin}_i$ that can take many forms. In the simplest form, $f$ is a constant and the problem solves to a difference-in-means. In more complex forms it can take on local polynomials that are estimated separately on either side of the cutoff. The first set of results take $f$ to be a polynomial on either side of the cutoff and are estimated within pre-specified bandwidths, while later results use local polynomials and data-driven bandwidth selection.

The first set of results set $f$ to be a polynomial on either side of the cutoff like so:

$$\text{Dynastic}_i = \tau\text{Win}_i + \sum_{p=0}^{P} \alpha_p \text{Margin}_i^p + \beta_p \text{Win}_i \times \text{Margin}_i^p + \lambda_t + \zeta_r + \epsilon_i, \tag{2.4}$$

where $P$ is the order of the polynomial fit on either side of the cutoff. I estimate this model for values of $P \in \{0, 1, 2, 3\}$; when $P = 0$, the terms on $\alpha_0$ and $\beta_0$ are dropped and the problem reduces to a difference-in-means with fixed effects. For all models I cluster standard errors at the district level, defined as per the original authors in their work and as the actual districts in Pakistan. This model is commonly employed to estimate both LATEs and placebo tests in a RD framework.

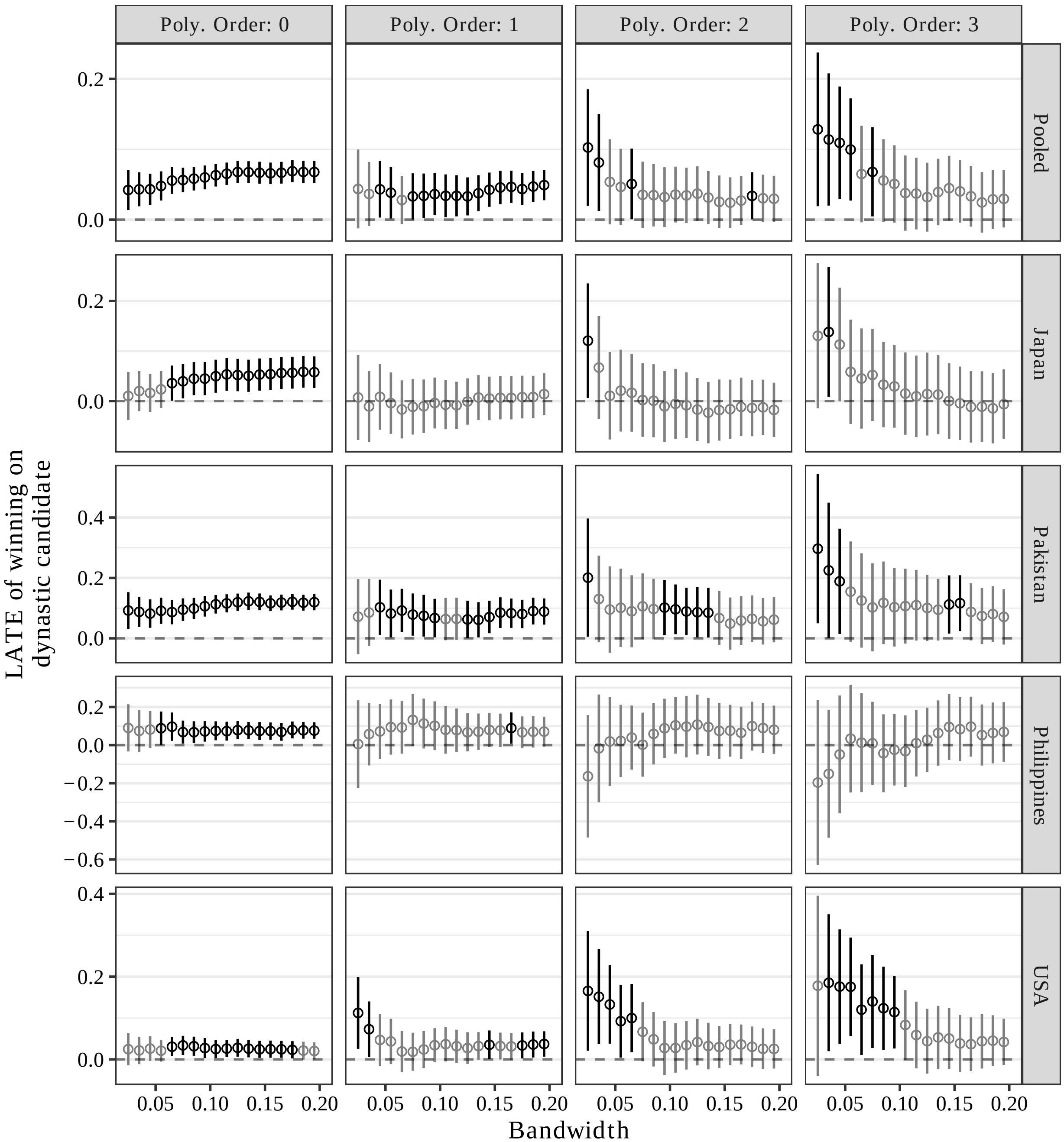
In Figure 2.3, I present the LATE of winning on dynastic status using the above specification for a variety of sub-samples of the data. On the vertical axis is the placebo LATE of winning a close election on being a dynastic candidate. On the horizontal axis are a variety of bandwidths used to trim the data. It is common practice to trim the data to a window where the assumptions of the RD model are more likely to hold, given that it is unlikely that candidates that won with large margins are likely to have potential outcomes that follow a similar functional form to candidates who are close winners and losers. The panels are organized into rows that correspond to different samples and into columns that correspond to different values of $P$, the order of the polynomial on the margin of victory running variable. The dots represent the estimated LATE, the line-ranges represent 95 percent confidence intervals. LATEs for which the p-value is below 0.05 are colored in black while others are in grey.

For example, the top-left panel presents estimates of the placebo test in the pooled dataset, with a polynomial of order 0 (i.e. a difference-in-means with the included fixed effects). The left most dot corresponds to the point-estimate when the bandwidth is at its smallest, 0.025. Therefore, the estimated treatment "effect" of barely winning an election on being a dynastic candidate is roughly 0.09 only candidates with $|Margin| < 0.025$ in the pooled dataset are considered. Moving right along the horizontal axis in the top-left panel, candidates who won and lost by larger and larger margins enter the estimation. The estimates on the pooled dataset (row one) are uniformly positive, and are statistically significant at the 0.05 level for a wide variety of specifications and bandwidths. The point estimates range from 0.024 and 0.126; in other words, I estimate that close winners in early career elections are 2.4 percentage points to 12.6 percentage points more likely to be dynastic candidates, depending on the specification used.

In both the United States and Pakistan, the majority of models exhibit imbalance between winners and losers in dynastic status, although there is less precision than in the pooled data. In Japan and the Philippines, the results are far more dependent on the specification used. There is no clear evidence of imbalance in either case, except for in the simple difference-in-means model (column 1, polynomial order = 0).

Presented in the above figure is a wide variety of specifications, a choosing the correct model from among them is quite difficult. The selection of the appropriate bandwidth is a difficult matter, as it is unclear ex ante how close an election has to be for the potential outcomes function to be smooth at the cutoff. Furthermore, estimating polynomials on either side of the cutoff can introduce overfitting and model dependence that creates large variation in the LATE at the cutoff. Fortunately, there are several tools to fine-tune the bandwidth selection, the estimation of smooth, local functions on either side of the cutoff, and produce robust confidence intervals with appropriate coverage (Calonico, Cattaneo and

Figure 2.3: Local average treatment effects of winning on being a dynastic candidate, by polynomial order and bandwidth




This figure displays a large number of LATEs of winning on being a dynastic candidate. Each panel contains a series of LATEs, where the value on the vertical axis is the magnitude of the LATE and the horizontal axis is the bandwidth used to estimate that LATE. The lines represent the 95-percent confidence intervals. The row panels represent the geographic sample used, and the column panels represent the order of the polynomial used in the regression discontinuity estimation. All models are estimated using OLS, include region and decade fixed effects, and have clustered standard errors at the district level.

Titiunik, 2014). I employ these tools in the `rdrobust` package in `R` in order to select the optimal bandwidth, estimate the potential outcomes functions on either side of the threshold using flexible, local polynomials, and correct confidence intervals. Note, there are still a wide variety of options one can select with robust RD estimators. The order of local polynomials, the algorithm for selecting the bandwidth, the order of the polynomials used in the bias correction, and more all go in to the choice of estimator. Here, I present results from a relatively simple model, where linear local polynomials are used and the bandwidth is selected as per Calonico, Cattaneo and Titiunik (2014).

In Table 2.3, I present the placebo LATE estimates using the data-drive bandwidth selection and local polynomial method described above. While the bandwidth selection is data-driven, it still may make sense to restrict the data that the procedure can use, as it may be plausible that those who won by huge margins or lost by huge margins are simply so different from those near the cutoff that include them in the same model is nonsensical. To address this possibility, each panel presents the LATE for the country-wise datasets and the pooled dataset after trimming the data to a closer sub-sample. The first panel uses all of the data, the second panel only uses candidates with $|Margin| < 0.2$, and so on; note that I am still using a data-driven bandwidth selection procedure, but in panels B through D, the data are first trimmed. The first row presents the LATE, with the standard error in parentheses below the LATE. I also present the data-driven estimated bandwidth and the sample size that falls within that bandwidth.

Using this data-driven and flexible procedure, a similar picture emerges. The placebo LATEs in the pooled data range from 3.6 percentage points to 11.3 percentage points, with all three of the four estimates statistically significant at the 0.05 level and the other significant at the 0.1 level. Again, at the country level, the strongest evidence for imbalance comes from Pakistan and the United States, although in the smaller samples samples the estimate for Japan is also large and statistically significant. The estimates in this table and the figure above both demonstrate ample evidence for imbalance in close winners and close losers in

Table 2.3: Local average treatment effects of winning on being a dynastic candidate using local polynomials

| | Japan | Pakistan | Philippines | USA | Pooled |
|---|---|---|---|---|---|
| **Panel A: Full data** | | | | | |
| LATE | 0.024 | 0.076$^{\dagger}$ | 0.088 | 0.036$^{\dagger}$ | 0.041$^{**}$ |
| | (0.033) | (0.040) | (0.058) | (0.022) | (0.015) |
| Est. BW | 0.061 | 0.171 | 0.156 | 0.138 | 0.160 |
| Eff. N | 1874 | 2749 | 1495 | 2996 | 10091 |
| **Panel B: Trimmed data ($\|MV\| < 0.2$)** | | | | | |
| LATE | 0.035 | 0.121$^{*}$ | -0.043 | 0.123$^{**}$ | 0.043$^{\dagger}$ |
| | (0.035) | (0.058) | (0.086) | (0.046) | (0.022) |
| Est. BW | 0.049 | 0.052 | 0.044 | 0.033 | 0.062 |
| Eff. N | 1679 | 1097 | 500 | 868 | 5544 |
| **Panel C: Trimmed data ($\|MV\| < 0.1$)** | | | | | |
| LATE | 0.122$^{*}$ | 0.157$^{*}$ | -0.098 | 0.143$^{**}$ | 0.087$^{*}$ |
| | (0.055) | (0.072) | (0.121) | (0.051) | (0.034) |
| Est. BW | 0.017 | 0.025 | 0.026 | 0.027 | 0.023 |
| Eff. N | 784 | 574 | 306 | 718 | 2341 |
| **Panel D: Trimmed data ($\|MV\| < 0.05$)** | | | | | |
| LATE | 0.154$^{**}$ | 0.236$^{**}$ | -0.169 | 0.184$^{**}$ | 0.126$^{**}$ |
| | (0.059) | (0.087) | (0.141) | (0.064) | (0.041) |
| Est. BW | 0.012 | 0.016 | 0.018 | 0.018 | 0.015 |
| Eff. N | 577 | 348 | 199 | 466 | 1537 |

$^{\dagger}$, p-value $<$ 0.1; $^{*}$, p-value $<$ 0.05; $^{**}$, p-value $<$ 0.01; $^{***}$, p-value $<$ 0.001. The reported estimated bandwidth was selected a data-driven bandwidth selection method. Each panel represents a different amount of data made available to the data-drive bandwidth selection procedure. The top panel admits all of the data, while subsequent panels trim the data before the bandwidth estimation is carried out. All LATEs estimated using local linear regression on either side of the cutoff, with robust bias-corrected standard errors and p-values, accounting for clustering at the district level. The pooled models include country fixed effects. All estimation conducted using the `rdrobust` package in `R` (Calonico, Cattaneo and Titiunik, 2014).

the dynastic status across several different contexts.

### 2.3.3 Explaining the dynastic advantage in close elections

What could be advantaging dynastic candidates in close elections? Many of the proposed mechanisms for dynastic persistence in the literature (e.g. name recognition, political connections, ambition, etc.) will advantage candidates whether they are in competitive or non-competitive elections and thus may not explain advantages in close elections. As discussed above, the imbalance could be caused by one of two problems: (1) there is a genuine discontinuity at the cutoff, and dynasts are able to barely win elections; or (2) there is not enough data to correctly estimate the discontinuity at the limit and the imbalance is a consequence of including data from elections too far from the cutoff with an inappropriate functional form.

The first explanation of a genuine discontinuity requires dynasts to engage in fraud or to precisely allocate resources. Using subsample effects, I can examine the conditions under which the dynastic imbalance is greatest to potentially discover when this sorting is likely. Appendix 2.B documents that the imbalance in the United States is largest in the time period that coincides with the imbalance in incumbency status estimated by Caughey and Sekhon (2011). The imbalance also appears to be larger for dynasts in the incumbent party, as demonstrated in Appendix 2.B. Therefore, it appears that the mechanisms enabling dynasts to win close elections likely coincides with the mechanisms allowing incumbents to win close elections. As such, the arguments made by Caughey and Sekhon (2011) may apply here; chiefly, that fraud is unlikely to explain sorting due to a lack of evidence from a subsample of elections they closely investigated, and that other, non-fraudulent benefits associated with being from the same party as the state or provincial government may be most important.

The second explanation that there is not enough data to appropriately estimate LATEs, then, appears more likely. However, the problem here is somewhat more difficult to articulate. Consider the case where the researcher has enough data. If there is no fraud or perfect observability of the vote count by campaigns, and if so much as one person is randomly discouraged or encouraged to vote, then there exist some elections for which estimating the LATE is simple. The researcher only needs to collect elections where the margin of victory is less than number of people randomly encouraged or discouraged from voting and treat it as a local experiment. However, knowing how many people would be randomly encouraged or discouraged to vote is not possible. Furthermore, this number may be so small that no, or very few, elections exist within that bandwidth. Fortunately, the assumption of continuous potential outcomes is weaker than the assumption that there is a random experiment within some bandwidth (De la Cuesta and Imai, 2016). Nonetheless, in order to accurately estimate a LATE, the continuity assumption must hold *and* the functional forms on either side of the cutoff must be correctly specified. As data further and further away from the cutoff is included in the estimation, the likelihood of that data reliably informing the effect at the cutoff decreases. Much more data near the cutoff would be needed. It is tempting to say, then, that to confirm whether the functional form is misspecified or there is a genuine imbalance, one simply needs to estimate an imbalance in a very small bandwidth. Indeed, in this paper, an imbalance exists even in very tight bandwidths. Nonetheless, what a "small" bandwidth means is undefinable. Very few elections exist where the margin of victory was within a few hundred votes and therefore reliance on functional form assumptions begins affecting estimates of close-elections RDD LATEs very quickly.

Nonetheless, neither the absence of fraud nor insufficient data at the cutoff are reasons to dismiss the imbalance demonstrated in this paper. Because this imbalance is estimable for a wide variety of estimation strategies and political contexts, and these strategies and contexts mirror those in which close elections RD designs are used by other social scientists, demonstrating the imbalance is sufficient to question the validity of some of those designs. The next section considers how this imbalance may affect published estimates relying on close elections RD designs.

## 2.4 Implications for close elections regression discontinuities

Are existing close elections RD results biased by this imbalance in dynasts winning close elections, and by how much? Before turning to a meta-analysis of extant close elections RD results, I use a paper by Eggers and Hainmueller (2009) as running example to explain how this imbalance relates to the central assumption of RD designs and the conditions under which the imbalance will bias estimates of treatment effects. I did not select this paper because I believe it is particularly prone to bias, but rather because of its thoughtful exposition and straightforward causal model. Eggers and Hainmueller (2009) study whether British candidates who barely win elections to become Members of Parliament are wealthier when they die. They invoke the usual assumption of continuous potential outcomes at the cutoff to justify their regression discontinuity estimator.[3] Therefore, they assume that the wealth of a candidate at their death had they won or had they not won are both continuous across the cutoff.

However, if wealth at death is a function of dynastic status near the cutoff, and if dynastic candidates are more likely to be on one side of the cutoff, then their estimates will be biased by not conditioning on dynastic status. Consider the case in which dynasts win close elections at higher rates and are also able to accumulate more wealth over their life time regardless of whether or not they won an election. In this situation, dynasts have higher treatment and control potential outcomes—they can earn more wealth over their lifetime than other individuals regardless of what their margin of victory is in this election. If they are also able to win close elections at higher rates, the treatment potential outcome—i.e. the wealth one would have if they won—would be discontinuous at the cutoff, because dynasts who have higher control and treatment potential outcomes are more likely to be on the winning side of the cutoff. Therefore, in this case, both the treatment and control potential outcomes will be discontinuous and their estimate is biased.

[3]Note they also employ a matching estimator and invoke a traditional selection on observables assumption in addition to their RD estimation.

Therefore, dynastic status will have to be correlated with both the potential outcomes under consideration as well as the treatment assignment mechanism—barely winning close elections. It is worth noting that given a small enough window and the presence of at least some random component to elections, potential outcomes will be smooth at the cutoff (Lee, Moretti and Butler, 2004). This is equivalent to stating that dynastic status is not associated with the treatment assignment mechanism, because for small enough margins of victory an election is truly as-if random. Unfortunately, as discussed above, this window may be so small as to be impractical for most data as demonstrated in this paper. For example, if the random component to elections only covered a few hundred votes, we would need ample elections within that window to guarantee continuous potential outcomes.

### 2.4.1 Sensitivity analysis

If these two conditions are met—dynasts have an advantage in close elections and dynasts have different potential outcomes for the outcome of interest near the cutoff—then close elections RD estimates will be biased. To understand the magnitude of the bias, I use the framework for sensitivity analysis proposed by Cinelli and Hazlett (2020). In this section, let $Z$ be the dynastic status of a candidate, $D$ be whether that candidate won an election, and $Y$ be some outcome of interest, such as wealth at death in the running example. In order for the effect of $D$ on $Y$ to be biased by the presence of dynasticism ($Z$), there must be a relationship between $Z$ and $D$—this is the dynastic advantage I document above—and between $Z$ and $Y$—this is the relationship between dynasticism and wealth at death. Understanding how much the dynastic advantage biases close elections effect estimates relies on stipulating the magnitude of these two relationships.

Given a postulated strength of the relationships between (1) dynasticism and winning close elections and (2) dynasticism and the outcome of interest, the bias in published RD estimates can be computed. I quantify these relationships by their partial $R^2$: the relationship between dynastic advantage and winning close elections, conditional on covariates $X$,

Table 2.4: Estimating relationship between dynasticism and winning close elections

| | Outcome: won election | |
|---|---|---|
| Dynastic cand. | 0.243** | 0.115*** |
| | (0.081) | (0.016) |
| Bandwidth | 0.005 | 0.1 |
| N | 503 | 7686 |
| $R^2_{D\sim Z}$ | 0.025 | 0.007 |

†, p-value < 0.1; *, p-value < 0.05; **, p-value < 0.01; ***, p-value < 0.001. The model used here is $\text{Win}_i = \tau \text{Dynastic}_i + \lambda_t + \zeta_r + \epsilon_i$, where all variables are described in Equation 2.3. Standard errors are classic standard errors as those are needed to construct $R^2_{D\sim Z|X}$, used in sensitivity analyses. Cluster robust standard errors are not much larger and still yield a significant coefficient dynastic candidate.

is represented as $R^2_{D\sim Z|X}$ and the relationship between dynasticism and the outcome of interest, wealth at death, is represented as $R^2_{Y\sim Z|X}$, again conditional on covariates $X$. Given the strength of these partial $R^2$ values, I then consider how much of the published effect of $D$ on $Y$ ($\hat{\tau}$) is bias due to dynasticism and present debiased estimates ($\hat{\tau}_*$).

I estimate $R^2_{D\sim Z|X}$ from the pooled data above and consider a range of hypothetical $R^2_{Y\sim Z|X}$ values. To estimate $R^2_{D\sim Z|X}$, I essentially flip the failed placebo tests above and regress winning an election on dynastic status (with province and year fixed effects, similar to the above models) within a certain bandwidth. I do so for two bandwidths: 0.5 and 10 percentage points, to show two different effects For these three values, the $R^2_{D\sim Z}$ in the pooled data is 0.025, 0.007 respectively. These results are reported in Table 2.4.

The relationship between an outcome of interest and dynasticism, $R^2_{Y\sim Z}$, cannot be measured for almost all of the published papers cited here, as dynasticism is often not a covariate available in published datasets. Instead, I postulate a range of potential values and compute the sensitivity of published results given a range of potential values. Taking $R^2_{D\sim Z|X}$, $R^2_{Y\sim Z|X}$, and the published estimates, standard errors, and degrees of freedom, I follow Cinelli and Hazlett (2020) and estimate the bias of the original estimate as

$$|\widehat{\text{bias}}| = \widehat{se}(\hat{\tau})\sqrt{\frac{R^2_{Y\sim Z|X}R^2_{D\sim Z|X}}{1-R^2_{D\sim Z|X}}(df)}, \tag{2.5}$$

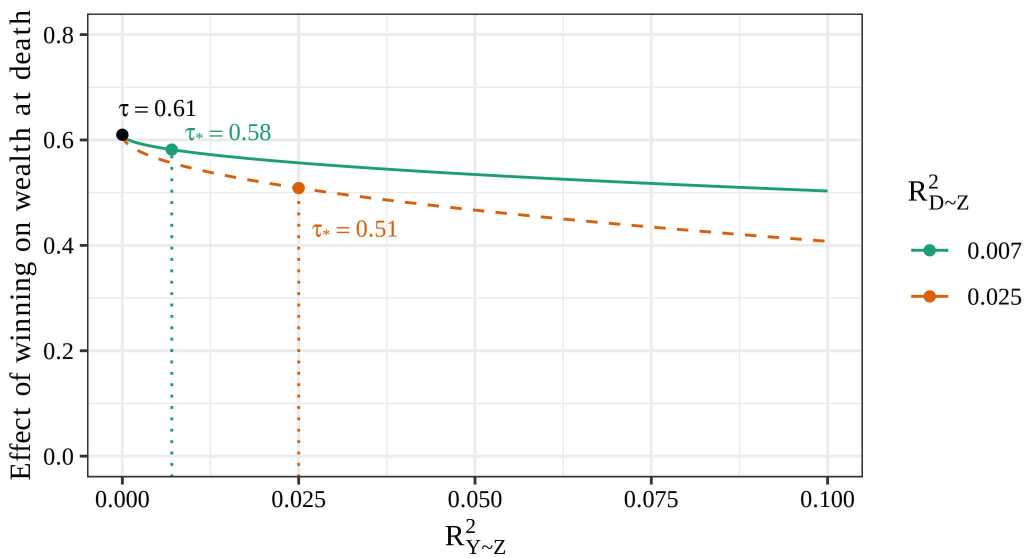
where $\widehat{se}(\hat{\tau})$ is the standard error of the original estimate and $df$ is the degrees of freedom.

In the running example, Eggers and Hainmueller (2009) estimate that when a Conservative candidate barely wins their election to become an MP, they die with log(assets) that are 0.61 greater ($\sim$ 83 percent greater) than Conservative candidates who barely lost elections. If Conservative *dynastic* candidates are both more likely to win close elections and more likely to accumulate more assets by the time they die, than these estimates will be upwardly biased. In Figure 2.4 I plot what the debiased estimates would be for a range of potential $R^2_{D\sim Z|X}$ and $R^2_{Y\sim Z|X}$ values. Note, these are hypothetical values; if there is no relationship between dynasticism and assets at death ($R^2_{Y\sim Z|X} = 0$) or there is no relationship between dynasticism and winning close elections in this context ($R^2_{Y\sim Z|X} = 0$), then there is no bias in their initial estimate.

In Figure 2.4, the horizontal axis is the range of different possible $R^2_{Y\sim Z|X}$ values—the hypothetical relationship between dynasticism and wealth at death. The vertical axis is the debiased treatment effect of a conservative barely winning a close election on their log assets at death. Each horizontal line represents a different estimated value of $R^2_{D\sim Z|X}$ from the pooled data—again, the different estimates come from slightly different specifications. To the right, as $R^2_{Y\sim Z|X}$ increases, and towards the lower lines, as $R^2_{D\sim Z|X}$ increases, the estimated treatment effect of a Conservative barely winning on wealth at death decreases substantially due to the bias from the dynasticism confounder. Because $R^2_{Y\sim Z|X}$ is unknowable without measuring dynasticism in this sample—and doing so for all published estimates would be very difficult—we cannot know which value on the horizontal axis corresponds to the actual debiased treatment effect.

In order to understand the sensitivity of a wide variety of published results, consider a

Figure 2.4: Debiasing the estimated effect of winning a close election on wealth from Eggers and Hainmueller (2009)




This figure displays how the published treatment effect $\hat{\tau}$ would vary for a variety of assumptions about $R^2_{D\sim Z|X}$ and $R^2_{Y\sim Z|X}$. The two lines represent two different values of $R^2_{D\sim Z|X}$, how much of the variation in close election winners $D$ is explained by dynasticism $Z$. The horizontal axis represents a range of potential hypothetical values of $R^2_{Y\sim Z|X}$, how much of the outcome $Y$ dynasticism explains. The vertical lines correspond to values of $R^2_{Y\sim Z|X}$ set equal to value of $R^2_{D\sim Z|X}$ for the line they intersect, with points and debiased estimates at their intersection.

few summaries of the above sensitivity plot. Each of the dotted vertical lines in Figure 2.4 and the intersecting point correspond to the debiased estimate when I hypothesize that the relationship between dynasticism and assets is just as strong as the relationship between dynasticism and winning close elections—in other words, when $R^2_{Y\sim Z|X}$ is set to the same value as $R^2_{D\sim Z|X}$. If these effects are equally strong in the sample at hand, then the debiased estimates are where the two vertical lines that intersect the debiased estimate lines. These intersections are marked by points on the two lines, corresponding to the two estimates of $R^2_{D\sim Z|X}$. In this example, when $R^2_{D\sim Z|X} = R^2_{Y\sim Z|X} = 0.025$, the debiased treatment effect is 0.51, 83 percent of the original estimate. In the next section, I follow a similar procedure to estimate the number of papers whose estimates would be quartered, halved, or have their signs reversed under the right conditions.

### 2.4.2 Meta-analysis of close elections RD estimates

In order to calculate the potential sensitivity of published results to this bias, I turn to a brief survey of the literature on close elections RD designs. First, I collect all papers in three top ranked political science journals (APSR, AJPS, and JOP) and three top ranked economics journals (AER, JPE, QJE) that I could find that use close elections regression discontinuity designs using JSTOR, Google Scholar, and manual citation tracing. I summarize these articles in Table 2.5.

Table 2.5: Articles using close elections regression discontinuity designs

| Journal | N | First Paper | Topic |
|---|---|---|---|
| QJE | 2 | 2004 | Effect of party/candidate type (2) |
| APSR | 8 | 2009 | Effect on other elections (2), Financial returns to office (2), Career advancement (1), Dynasticism (1), Federal alignment & service delivery (1), Incumbency advantage (1) |
| JOP | 11 | 2010 | Effect of party/candidate type (3), Incumbency advantage (3), Effect on other elections (2), Federal alignment & service delivery (2), Financial returns to office (1) |
| AER | 2 | 2011 | Effect of party/candidate type (2) |
| AJPS | 2 | 2011 | Effect of party/candidate type (1), Effect on other elections (1) |
| JPE | 1 | 2014 | Financial returns to office (1) |

Papers that use close elections RD designs in the three political science and top five economics journals. The topics were hand coded, and the full citations and list can be found in Appendix 2.C.

These papers are varied in their topic of study, treatment of interest, exact estimation strategy, and outcomes considered. However, they all rely on an individual or party barely winning or losing an election to study the effect of either (1) the candidate winning versus losing or (2) the type of candidate winning as compared to another type of candidate winning. The first kind of study often takes the candidate (or their party) as the unit of analysis and studies outcomes for that party or candidate. The three most common versions of this study are incumbency advantage studies, studies of how election results spill over onto other, related elections, and studies of the financial returns to office. In studies of the incumbency advantage, researchers estimate the effect of candidate or party winning on the probability that candidate or party wins a later election. In studies of cross-election spillover, researchers study the effect of a candidate winning, for example, a local election on the party's performance in national level elections. In studies of the financial returns to office, researchers estimate the effect of winning an election on the wealth of the winner or the economic performance of affiliated firms. All of these studies specifically study the effect of barely winning versus barely losing on outcomes directly tied to the individual winning or losing.

In the second set of studies, researchers use close elections to study the effects of certain *types* of candidates winning elections, often on outcomes in their constituency. This can range from what happens in a constituency when a Democrat barely beats a Republican in the United States, for example, to whether a member of the party in federal government barely winning an election increases transfers and service delivery from the federal government to the constituency. Rather than focusing on effects on the individual winning or losing, these designs instead use regression discontinuities to hold the features of constituencies constant, reasoning that places that are barely won by candidates of type $A$ are similar to places where candidates of type $A$ barely lost. In this case, the treatment $D$ is actually a candidate of type $A$ beating a candidate of type $B$, rather than whether a candidate of type $A$ wins. Thus, these results will only be biased if dynasticism and type $A$ are correlated in candidates. If they are, then the bias comes from attributing outcomes to what happens when candidate of type $A$ wins when really part of the effect is due to what happens when a dynastic candidate wins.

In Table 2.6, I report the percent of papers that would have their most robust estimate at least halved or switch signs if $R^2_{Y\sim Z|X} = R^2_{D\sim Z|X}$, for both values of $R^2_{D\sim Z|X}$ that I estimate. First, I take the most robust estimate for each paper—this is the estimate with the max $\hat{t_\tau}/\sqrt{df}$ where $\hat{t_\tau}$ is the t-statistic of the original estimate. Then, I estimate the debiased estimate for the three values of $R^2_{D\sim Z|X}$, where $R^2_{Y\sim Z|X}$ is set to the same value. Then I calculate the percent of those estimates that are quartered, halved, or reverse signed. The results are reported in the second, third, and fourth columns of Table 2.6, respectively. Note that I always assume the bias to work against the published estimate; it is also possible given the assumed relationship between dynasticism and the outcome of interest that these estimates could be made even larger.

Table 2.6: Sensitivity of RD estimates to potential dynastic imbalance

| $R^2_{D\sim Z}$ | % quartered | % halved | % change sign |
|---|---:|---:|---:|
| 0.007 | 9.1 | 9.1 | 4.5 |
| 0.025 | 40.9 | 13.6 | 9.1 |

The first column is the estimated strength of relationship between dynasticism and winning a close election as estimate from the pooled data. Each row therefore contains the percent of papers that would have their most robust estimate quartered, halved, or change sign if the relationship between dynasticism and their outcome of interest was equally strong and was signed against their original treatment effect.

Even for the weakest estimated relationship between dynasticism and winning close elections, hypothesizing that dynasticism has a similarly strong effect on the paper's outcome of interest will quarter the most robust estimate in 9.1 percent of papers. Relying on the tightly trimmed dataset to estimate $R^2_{D\sim Z|X} = 0.025$, over 40 percent of papers are calculated to have their estimate quartered. There are several caveats. For some of the papers, it is unlikely that there is a systematic relationship between dynasticism and the outcome of interest. For others, perhaps family politics are not widespread enough for the magnitude of the imbalance on dynastic status in close winners to hold. Nonetheless, due to the fragility of many regression discontinuity estimates, and due to the strong relationship between dynastic status and winning close elections estimated from this pooled data, the potential for debiased estimates to be significantly different from their published values is high.

## 2.5 Discussion

In this paper, I document a large and (in many specifications) statistically significant imbalance in the probability that a close winner of an election is dynastic. In other words, having a family member precede a candidate in office makes it more likely that candidate wins close elections. This effect is robust in Pakistan and the United States and may exist in Japan and Philippines although the evidence there is more mixed. When pooling the data, there is a clear average imbalance in the dynastic status of close winners and losers. I also document that this imbalance, when paired with an assumption that there is an equally strong relationship between dynastic status and an outcome of interest, halves the most robust published effect estimated using a close elections RD design in 9 to 41 percent of papers.

The magnitude of the potential sensitivity of published estimates is more important to elucidate the magnitude of the imbalance than it is to undermine the extant close elections RD literature. The assumption that dynastic status matters as much for each of these outcomes as it does for winning elections may be unlikely, and in some contexts dynastic politics are marginal. Furthermore, many of the papers included in the meta-analysis show balance on a whole host of important characteristics that may largely capture the effect of dynasticism even if it were to exist.

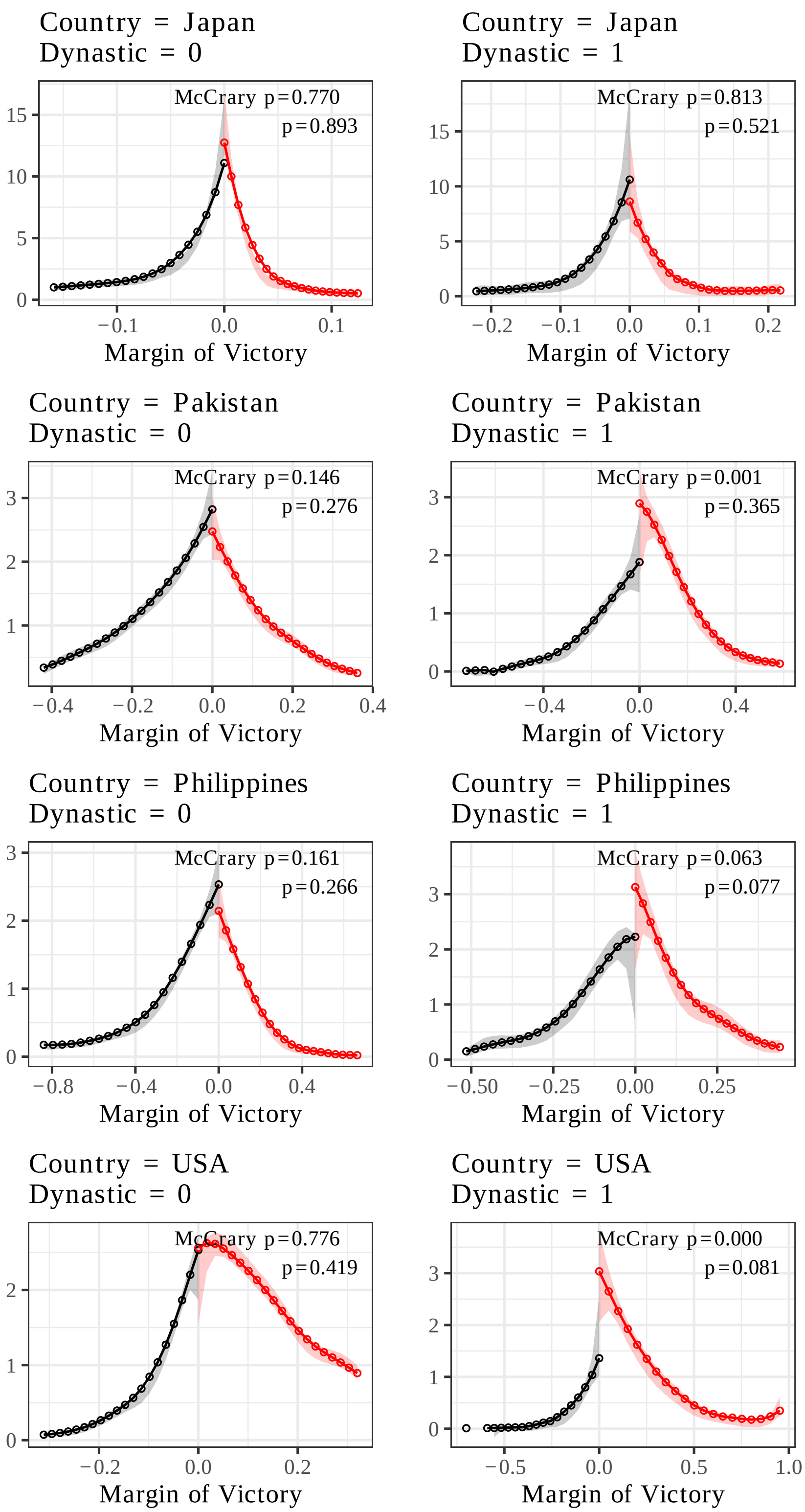
Nonetheless, dynasticism is hard to observe and expensive for researchers to collect. It captures a wide variety of latent skills, assets, and advantages that a politician may have access to that can influence not only their ability to win a close election but also many other outcomes for themselves and their party. Furthermore, this imbalance is not confined to one case as the imbalance in incumbents winning close elections is to the United States House in the post-war period. This imbalance exists across continents and across electoral systems. It is likely that in many contexts, there simply is not enough data on close enough elections to safe guard a design from all potential imbalances of this sort, and it is important for researchers to do their due diligence to track down as many potential auxiliary variables as possible to ensure they have balance. Unfortunately, collecting dynastic status of candidates may be prohibitively costly in some circumstances, in which case I advise authors to consider whether they believe dynasticism may play a role in the elections they study.

# Appendix

## 2.A Density tests for sorting of dynastic candidates

In this section I present the result of McCrary (2008) density tests and of local polynomial density tests as proposed by Cattaneo, Jansson and Ma (2019). These tests check whether certain individuals are sorted around some cut-off and are a complement to the failed placebo tests presented in the body of the paper. If there is sorting within dynasts, then there is evidence that dynasts are able to get just above the threshold and may have an advantage in close elections. The results are in Figure 2.A.1, where estimated local densities are plotted, with the corresponding p-value and the McCrary p-value printed on each panel. The McCrary test identifies sorting in Pakistan and the USA at the 0.05 significance level and the Philippines at the 0.1 significance level. The local polynomial estimator is more cautious and only estimates sorting in the Philippines and the United States at the 0.1 significance level.

Figure 2.A.1: Density tests for sorting along margin of victory by dynasts and non-dynasts by country



## 2.B Dynastic imbalance in different subsamples

This section presents the same placebo tests by sub-group. In Table 2.B.1 we demonstrate that the dynastic imbalance is most common in 20th century United States and Table 2.B.2 shows that this effect is also greatest when dynastic candidate is a member of the incumbent party.

Table 2.B.1: Robust LATEs of winning on being a dynastic candidate in the United States, subset by time period

| | Panel A: Full data | | |
|---|---|---|---|
| | 1788-1880 | 1882-1944 | 1946-1988 |
| LATE | -0.010 | 0.037 | 0.099$^{\dagger}$ |
| | (0.037) | (0.030) | (0.054) |
| Est. BW | 0.128 | 0.223 | 0.122 |
| Eff. N | 1072 | 1877 | 402 |
| | Panel B: Trimmed data ($\lvert MV\rvert < 0.2$) | | |
| | 1788-1880 | 1882-1944 | 1946-1988 |
| LATE | 0.071 | 0.091 | 0.125* |
| | (0.068) | (0.069) | (0.063) |
| Est. BW | 0.039 | 0.061 | 0.070 |
| Eff. N | 415 | 749 | 247 |

†, p-value < 0.1; *, p-value < 0.05; **, p-value < 0.01; ***, p-value < 0.001. The reported estimated bandwidth was selected a data-driven bandwidth selection method. Each panel represents a different amount of data made available to the data-drive bandwidth selection procedure. The top panel admits all of the data, while subsequent panels trim the data before the bandwidth estimation is carried out. All LATEs estimated using local linear regression on either side of the cutoff, with robust bias-corrected standard errors and p-values, accounting for clustering at the district level. All estimation conducted using the `rdrobust` package in `R` (Calonico, Cattaneo and Titiunik, 2014).

Table 2.B.2: Robust LATEs of winning on being from the incumbent party in Pakistan and the United States, subset by running with the incumbent party

| | Panel A: Full data | |
|---|---|---|
| | Not incumbent party | Incumbent party |
| LATE | 0.049 | 0.102*** |
| | (0.031) | (0.029) |
| Est. BW | 0.176 | 0.166 |
| Eff. N | 3427 | 2214 |

| | Panel A: Trimmed data ($\lvert MV\rvert < 0.2$) | |
|---|---|---|
| | Not incumbent party | Incumbent party |
| LATE | 0.100* | 0.147** |
| | (0.048) | (0.055) |
| Est. BW | 0.053 | 0.052 |
| Eff. N | 1398 | 924 |

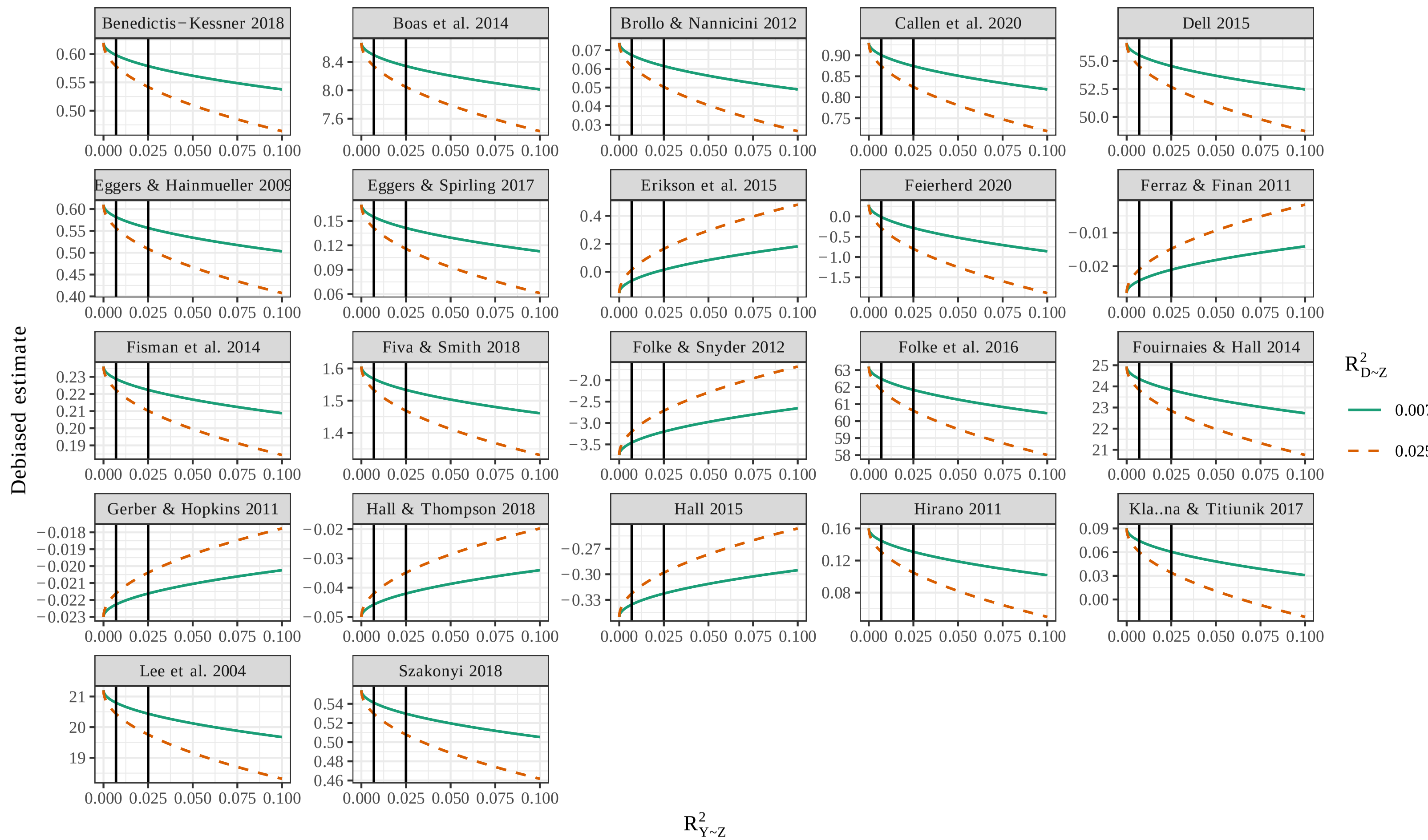
†, p-value < 0.1; *, p-value < 0.05; **, p-value < 0.01; ***, p-value < 0.001. The reported estimated bandwidth was selected a data-driven bandwidth selection method. Each panel represents a different amount of data made available to the data-drive bandwidth selection procedure. The top panel admits all of the data, while subsequent panels trim the data before the bandwidth estimation is carried out. All LATEs estimated using local linear regression on either side of the cutoff, with robust bias-corrected standard errors and p-values, accounting for clustering at the district level. All estimation conducted using the `rdrobust` package in `R` (Calonico, Cattaneo and Titiunik, 2014).

## 2.C Sensitivity analysis: articles and additional results

Table 2.C.1: List of close elections regression discontinuity articles

| Authors | Journal | Topic |
|---|---|---|
| Lee, Moretti and Butler (2004) | QJE | Effect of party/candidate type |
| Eggers and Hainmueller (2009) | APSR | Financial returns to office |
| Ferreira and Gyourko (2009) | QJE | Effect of party/candidate type |
| Schickler, Pearson and Feinstein (2010) | JOP | Effect of party/candidate type |
| Ferraz and Finan (2011) | AER | Effect of party/candidate type |
| Gerber and Hopkins (2011) | AJPS | Effect of party/candidate type |
| Hirano (2011) | JOP | Federal alignment & service delivery |
| Brollo and Nannicini (2012) | APSR | Federal alignment & service delivery |
| Folke and Snyder (2012) | AJPS | Effect on other elections |
| Boas, Hidalgo and Richardson (2014) | JOP | Financial returns to office |
| Fisman, Schulz and Vig (2014) | JPE | Financial returns to office |
| Fouirnaies and Hall (2014) | JOP | Incumbency advantage |
| Dell (2015) | AER | Effect of party/candidate type |
| Erikson, Folke and Snyder Jr (2015) | JOP | Effect on other elections |
| Hall (2015) | APSR | Effect on other elections |
| Carnes and Lupu (2016) | JOP | Effect of party/candidate type |
| Folke, Persson and Rickne (2016) | APSR | Career advancement |
| Caughey, Xu and Warshaw (2017) | JOP | Effect of party/candidate type |
| Eggers and Spirling (2017) | JOP | Incumbency advantage |
| Klašnja and Titiunik (2017) | APSR | Incumbency advantage |
| de Benedictis-Kessner (2018) | JOP | Incumbency advantage |
| Fiva and Smith (2018) | APSR | Dynasticism |
| Hall and Thompson (2018) | APSR | Effect on other elections |
| Szakonyi (2018) | APSR | Financial returns to office |
| Callen, Gulzar and Rezaee (2020) | JOP | Federal alignment & service delivery |
| Feierherd (2020) | JOP | Effect on other elections |

Figure 2.C.1: Sensitivity plots for most robust published estimate by paper




This figure displays how the published treatment effect $\tau_{original}$ would vary for a variety of assumptions about $R^2_{D\ Z|X}$ and $R^2_{Y\ Z|X}$. The two lines represent two different values of $R^2_{D\ Z|X}$, how much of the variation in close election winners $D$ is explained by dynasticism $Z$. The x-axis represents a range of potential hypothetical values of $R^2_{Y\ Z|X}$, how much of the outcome $Y$ dynasticism explains. The vertical lines correspond to values of $R^2_{Y\ Z|X}$ set equal to value of $R^2_{D\ Z|X}$ for the line they intersect, with points and debiased estimates at their intersection. Each panel is a separate paper represented by their most robust estimate.

# CHAPTER 3

# Doing As Your Neighbors Do, Not As They Say: Norms, Social Networks, and Female Political Participation in Pakistan

**Abstract:** Social norms play a large role in the decision to vote. In historically conservative contexts where female political participation lags behind enfranchisement, how men and women understand those norms may be an important constraint to female mobilization. In this paper, I study how the behavior and beliefs of a woman's social network, as well as her empirical expectations—beliefs about what others actually do—and normative expectations—beliefs about what others think one ought to do—constrain female political participation in Pakistan. Using social network data from 37 communities, I use a hierarchical social network auto-correlation model to demonstrate that women's self-reported turnout is more strongly correlated with empirical expectations and the behavior of one's peers than normative expectations. These results are in line with the view of voting as a coordination game, which is likely when female mobility outside the household is limited. I also demonstrate that supportive social norms about women's suffrage license female political participation for women who personally support female political participation, but are not sufficient to mobilize women who are uncertain about whether women ought to have the right to vote. Lastly, I consider policy implications for female mobilization efforts.

## 3.1 Introduction

Behavior, especially publicly visible behavior, is often influenced by social norms. When choosing how to act, social norms interact with an individual's privately held beliefs as well as their internal calculus about the private benefits and costs to a behavior. In the context of turning out to vote, the private benefits to voting are small enough that the social norms rewarding or punishing conformant behavior are likely to be large factors. Indeed, much of the recent literature on political participation has focused on the social contexts in which voters find themselves. Meta-analyses of traditional, survey-based studies of voter turnout often focus on a voters' peer group (e.g. Smets and Van Ham, 2013), while large-scale randomized controlled trials (e.g. Bond et al., 2012; Gerber, Green and Larimer, 2008) demonstrate how leveraging peer pressure and social networks can cause individuals to vote at higher rates.

Researchers often consider two kinds of social norms: those that summarize what people actually do and those that summarize what people think others ought to do. Bicchieri (2006) builds on this framework and argues that behavior is shaped by the *empirical expectations* about what others actually do and *normative expectations* about what others believe one ought to do. These expectations appear to be key drivers of behavior across contexts and behaviors (for a review, see Tankard and Paluck, 2016). If the goal is to change behavior through social norms, understanding whether perceptions of these norms are binding and how they are binding is an important step. The evidence on the efficacy of targeting either normative expectations—expectations about "prescriptive" or "injunctive" norms—or empirical expectations—expectations about "descriptive norms" is mixed.[1] In a historically conservative society where turnout among women lags far behind turnout of men, both descriptive norms and prescriptive norms may be constraining female political participation.[2]

[1] For a review of the relevant literature on mobilization, see Gerber and Green (2017) and for a review of the social psychology literature, see Miller and Prentice (2016).

[2] In this paper, I attempt to use "woman" where possible, as the data here is built on a sample of heads

In Pakistan, where the turnout rate among women is 9 percentage points lower than among men, recent survey evidence points to strong patriarchal norms as barriers to female political participation (Cheema et al., 2019). More specifically, female behavior movement outside of the household is often constrained in Pakistan; a woman choosing to vote often has to be accompanied by another woman. Indeed, in a survey in rural Sindh, a province of Pakistan, Giné and Mansuri (2018, p. 221) find that only 2 percent of women who voted went to a polling station alone and 87 percent of women went with another woman.[3] Therefore, while it may be helpful if a community believes women *ought* to have the right to vote, it is almost necessary that other women vote in order for a woman to be able to vote herself. Women appear to need to coordinate their voting behavior in order to go to the voting booth. As a result, both a woman's empirical expectations—what she expects other women will actually do—as well as what her peers actually do, should strongly predict whether she chooses to vote.

In this paper, I analyze social networks built on 37 communities in Pakistan to consider how different kinds of beliefs and expectations, as well as how the beliefs and behaviors of one's peers, correlate with female political participation. I consider three kinds of beliefs: (1) normative beliefs—personally held beliefs about how one ought to behave; (2) normative expectations—one's expectations about how *others* believe one should behave; and (3) empirical expectations—one's expectations about how others will behave.[4]

Using a hierarchical Bayesian network (auto-)correlation model, I test how these different kinds of beliefs and behaviors correlate with a woman's self-reported voting behavior. Using survey responses and constructed social networks, I jointly model how a woman's expectations about her peers as well as what her peers actually believe and say they do correlates

of household who identify as women. Nonetheless, I may use "female" at times when an adjective is more stylistically convenient.

[3] The remaining 11 percent went with a man.

[4] Note that normative expectations are predictions of "prescriptive norms" or "injunctive norms" and that empirical expectations are predictions of "descriptive norms".

with her behavior. Unsurprisingly, women who believe women ought to have the right to vote are more likely to self-report voting. However, a woman's *empirical expectations* about her peers and *behavior* of her female peers are more predictive of her behavior than her *normative expectations* and the actual *normative beliefs* of her peers. In other words, women in these communities do what they *think* other women do and what other women in their social network actually do, rather than what other women think they ought to do. This finding is in line with the view that voting for women in these contexts is a coordination game, and corroborates recent experimental evidence from the literature that empirical expectations may be more central to behavior than normative expectations (Bicchieri and Xiao, 2009).

While normative expectations fail to explain female political participation, the personal normative beliefs of women are highly predictive of their choice to turn out and also condition the relationship between expectations and self-reported behavior. A woman is most likely to report voting when she believes she should have the right to do so *and* they have expectations that other women vote or believe so as well. Expectations of others' beliefs or behavior are not sufficient to explain self-reported turnout on their own; even if a woman believes many other people support the right to vote, if she herself is unsure of the right to vote for women she will be less likely to self-report voting. In other words, expectations around female political participation can permit—or license—a woman to behave in line with her private beliefs, but may not be enough to induce her to change her private beliefs. This finding is consistent with the social psychology literature arguing that norms are easier to follow when one's beliefs are consistent with the norm (Tankard and Paluck, 2016, p. 198; Schroeder and Prentice, 1998).

These findings chiefly contribute to the literature on social norms and expectations as well as the literature on female political participation. If female political participation is lagging behind male political participation in historically conservative societies, then it may take the form of a coordination game. This paper demonstrates that in such a context women will be more responsive to the behavior of their peers—and expectations about this behavior—than about what one's peers believe ought to be the case. Nonetheless, this does not mean that normative expectations and normative beliefs are completely unimportant; this paper also demonstrates support for the licensing effect found in the social psychology literature.

These findings suggest a simple policy for increasing female political participation: target women with low empirical expectations and high levels of personal support for female political participation with information suggesting that other women self-report voting at higher levels than they anticipate. However, the number of women that underestimate turnout in their community is quite small in this sample. In fact, most people have empirical expectations about on par with self-reported turnout in their communities. A larger gap exists between normative expectations and the average level of support for female political participation in their community. In other words, both men and women underestimate the prescriptive norms held by members of their community. This gap exists for both men and women, and the gap appears even larger among men. In the absence of a gap between empirical expectations and practice, it may be implausible that treatments focused on providing information that voters already have will change their behavior. Instead, interventions should focus on the large gap that exists between average normative beliefs and normative expectations.

In the next section, I discuss some implications of the literature on social norms for the study of female political participation in Pakistan. I then turn to the survey and data used to test these implications, as well as how I construct the social networks used in the analysis. In Section 3.4, I introduce the network autocorrelation model and test what kinds of expectations and beliefs are strongly correlated with self-reported turnout within and across individuals. In Section 3.6 I turn to implications for the understanding of gendered social norms in patriarchal societies and for mobilization efforts to close the gender gap before concluding.

## 3.2 Social norms and political participation

Norms are often defined as either the way things are or the way most people think things ought to be. Actual practices and behaviors—how people behave—are characterized as *descriptive norms* and individuals hold *empirical expectations* that are their beliefs about how others behave. The collection of beliefs about what ought to be—how people should behave—are often called *prescriptive* or *injunctive norms* and individuals hold *normative expectations* about what others believe ought to be. This classification of norms into two separate camps is widespread and can be found in economics (Bicchieri, 2006), political science (Gerber and Green, 2017), and psychology (Tankard and Paluck, 2016). While these norms are often correlated, there are many situations in which people believe one ought not to engage in a behavior, but often still do because the behavior is believed to be widespread, such as corruption (Köbis et al., 2015, e.g.). The reverse is also possible, where one believes a behavior is widespread even if they themselves believe they ought not engage in the behavior, such as drug use among adolescents (Schroeder and Prentice, 1998).

Which kind of expectation is more pivotal to explaining behavior is unclear and likely context dependent. In a lab experiment, (Bicchieri and Xiao, 2009) demonstrate that empirical expectations are more important than normative expectations in predicting payments in a Dictator game, arguing that the costs to breaking normative expectations are low when the disapproved behavior is common. On the other hand, recent evidence from Saudi Arabia indicates that correcting misperceived normative expectations can increase female labor force participation (Bursztyn, González and Yanagizawa-Drott, 2018). However, because information about one norm often conveys information about another norm, even experimental settings can fail to disentangle the two (Blanton, Köblitz and McCaul, 2008) and the effectiveness of targeting one norm versus the other remains unclear (Miller and Prentice, 2016).

### 3.2.1 Female political participation as a coordination game

Because voting is a necessarily observable act that requires travelling outside of the home to a public location, there are likely complementarities to social neighbors choosing to vote that may outweigh normative considerations if such travel is costly. In our sample area, women frequently report requiring permission to leave their homes for social gatherings and often will not travel far from home unless travelling with someone else. Even in urban settings, such as Lahore, over 70 percent of men would advise women against taking non-segregated public transport (Sajjad et al., 2017). These norms spill over onto female political participation. In a rural sample in Sindh, Pakistan, Giné and Mansuri (2018, p. 221) find that only two percent of women who voted in 2008 traveled to the polling station alone. Over 87 percent of women who voted traveled to the polling station with another woman. Therefore, in a context where female mobility is restricted, voting appears to take the form of a coordination game wherein women will be able to vote if they can find another woman who is also willing and able to vote.

When a behavior takes the form of a coordination game, then the behavior of one's peers as well as one's expectations about how others behave should be more predictive of an individual's behavior. Normative expectations and the normative beliefs of one's peers are not sufficient to overcome the coordination problem—knowing your peers support a woman's right to vote is unlikely to mobilize a woman to vote if she requires another woman to accompany her to the polling station. Therefore, female political participation, when it takes the form of a coordination game, should be more responsive to empirical expectations and the behavior of one's peers than normative expectations about and the beliefs of one's peers. A secondary implication of voting as a coordination game for women in these contexts is that more socially connected women will be more likely to find someone with whom they can vote. As a result, women with more social ties should be more likely to report voting. Of course, there there are large number of reasons why women that have more social connections vote (see Prillaman, 2018); nonetheless, the positive relationship between social connections and voting still is an implication of voting as a coordination game.

### 3.2.2 Support for women's suffrage licenses behavior

Furthermore, widespread support for a woman's right to vote serves as permission to vote rather than as an indication that non-voters will be shamed.[5] Similarly, in a historically conservative context where female political participation is low, increased expectations about the number of women voting is more likely to be understood as increased permission to vote, rather than as a new norm that will lead to sanctions for women who do not vote. Consider a woman who believes women ought to have the right to vote. If she learns that other women are voting, she is more likely to vote than a woman who is uncertain about whether a woman ought to have the right to vote. As a result, she will be more responsive to the social norms when they come into line with her own preferences, a result that has been explored in the social psychology literature (see *licensing*, Tankard and Paluck, 2016, p. 198). The prediction is rather intuitive and applies to the inverse situation. Consider instead a woman who believes that women ought *not* to vote. If she learns that other women are voting and that others believe women should be allowed to vote, they may not respond to that information by voting as it disagrees with their privately held belief. If a woman learns that many other women are voting and/or that many other people believe that women should *have the right* to vote, she still may not vote if she herself holds a different belief. Therefore, I expect that women who believe women should have the right to vote will be more likely to vote if they have high normative and empirical expectations. I also expect that women who are uncertain whether women should have the right to vote will not be more likely to vote even when they believe other women vote and believe that support for women's suffrage is widespread.

[5] Although related, whether one ought to vote is different from whether one should have the right to vote, the focus of this study.

## 3.3 Data

The data used in this paper come from surveys conducted in rural District Peshawar, Khyber Pakhtunkhwa (KP), Pakistan. These surveys were conducted in March and April of 2019 and serve as a baseline for a follow-up experiment with Saad Gulzar and Muhammad Yasir Khan. District Peshawar is home to the capital of the province of KP and many surrounding towns and loosely clustered communities. District Peshawar is one of the most developed districts in KP, with a Human Development Index (HDI) of 0.756 in 2015, around the same level of countries such as Ukraine and Ecuador. This is well above the national HDI in 2015 of 0.560.[6] We do not work in settlements in the western part of the district, close to the Federally Administered Tribal Areas, where conservative values are more prevalent and backlash was more probable.

As a result, this sample is not representative of Pakistan as a whole. While the peri-urban and rural areas surveyed are certainly less developed and less wealthy than the provincial capital that lies in the center of District Peshawar, they are not far removed from a large urban center nor are they urban themselves. Furthermore, the exclusion of areas deemed to be too conservative or where door-to-door canvassing was not going to be sanctioned by local elites means the results here do not speak to strictly conservative parts of the country. These areas were chosen for two main reasons: (1) more rural areas tend to exhibit slightly larger gender gaps in Pakistan and thus are more likely areas for the aforementioned experiment to succeed, and (2) we collected information on social networks within settlements, explicitly excluding ties individuals may have outside of the settlement in which they live. In urban areas, defining a settlement in this way would be completely implausible as the distinction between "settlements" would be arbitrary.

The survey was a door-to-door near-census of households[7] in 37 settlements in District

[6]See the UNDP Pakistan report here: https://www.pk.undp.org/content/pakistan/en/home/library/human-development-reports/PKNHDR.html.

[7]A household is defined as a group of individuals, often bound by kinship ties, who share a kitchen and

Peshawar. In total, there are 4,892 households in these 37 settlements, an average of about 132 households per settlement. In each household, we attempted to survey both male and female heads of household. Men and women were surveyed separately by matched gender enumerators. We were only able to survey both female and male heads of household in 3,738 households, although we were able to reach 4,348 of the female heads of household. In the analyses below, I primarily use a dataset trimmed to households where we were able to reach the female head of household, although I will highlight when I use the dataset where we were able to reach both heads of household.

In addition to collecting some basic demographic, political knowledge, and family data, this survey chiefly focused on asking about social networks and beliefs and behavior with respect to female political participation. I will describe how we built social networks from this data before explaining how we measured beliefs and providing some descriptive information about this sample.

### 3.3.1 Measuring social networks

Social networks can be measured either via the direct observation of behavior or relationship statuses or through respondent reports of various kinds of social ties. I use survey responses about several different types of connections that households may have with one another in order to construct a social network. We asked both men and women about their social networks. However, in this study I will always refer to the network nominated by women unless otherwise stated. To build these social networks, we first conducted a census of all households in each settlement and built a list of possible households to which a respondent could be socially connected. Then, during the baseline survey, we ask respondents about who they engage with in one of four different ways. For each of these kinds of interaction, respondents nominate up to five other households within the same settlement from the

routinely share meals. A physical house or compound could have several households by this definition.

census list. Note that this restricts the social network to be within the settlement, as defined above. Table 3.1 contains the four questions used to construct social networks. For each type of network, respondents could nominate up to five other households with whom they were connected either by: (1) family, (2) socialization, (3) discussion of community issues, and (4) discussion of for whom to vote.[8] The analyses below use the union of all of these nominations as the full social network.

Table 3.1: Measuring different social networks

| Network type | Survey question |
|---|---|
| "Family" | "Please list up to 5 households besides this one where any of the following relatives live: siblings, parents, husband's parents, husband's siblings, children." |
| "Socialize" | "Please list up to 5 households with whom you spend the most time during a typical week." |
| "Community Issues" | "Please list up to 5 households with whom you discuss community issues, such as gas, electricity, roads, rubbish, and other problems." |
| "Discuss Voting" | "Please list up to 5 households with whom you and your family consulted when deciding who to vote for." |
| All network types | The union of the above four networks. |

I build directed social networks based on the nominations by women separately from the directed social networks I build based on nominations by men. In a directed network, ties between individuals flow from the nominating household to the household they nominated. This process allows us to identify which households are referential for each household, while not assuming that the relationship is equal in the other direction. The set of connections I consider is fully captured by the adjacency matrix $\mathbf{W}$. Each element $W_{ij} = 1$ if household $i$ mentioned household $j$ as the answer to any of the four questions. In all cases, households are linked to one another and not to individuals within households; the only difference across the

[8]Note, the limit of five households is not seriously problematic for our measurement as respondents only listed five households 1.02% of the time they were asked to list their connections.

female and male-nominated networks is whether the respondent nominating other households is the female or male head of household.

Figure 3.1 is a representation of one of the settlement's female-nominated networks. Each node, or circle, represents a household, and the size of the node for a household corresponds to the *indegree* of that household. The *indegree* is the number of times that a particular household is nominated by women in other households in any one of the four kinds of networks. More precisely, the *indegree* centrality for unit $j$ is the sum of the $j$th column of the adjacency matrix $\mathbf{W}$. As such, it serves as a measure of how many other households nominated an individual household, and thus view it as central to their social lives.[9] For example, if household A is mentioned as a family member of household B and as a household that household C socializes with, then it has an indegree of two. Note that if a household is mentioned by another household in two kinds of networks (e.g. family and social) this still only contributes one to the indegree of the household.

Figure 3.1: Example of one social network

Each node in this network is a household, the size of a node is the households indegree—the number of households that nominated it as a tie, and a line between nodes indicates one household nominated the other.

[9]There are other measures of centrality, such as eigenvalue centrality, which I describe and use later on.

### 3.3.2 Measuring beliefs

In surveys of both men and women we administered a series of questions about social norms and beliefs.

#### 3.3.2.1 Personal normative beliefs

We ask three questions about the respondent's support for various forms of female political participation, each on a five point scale. These three questions were about *normative* beliefs—about what respondents believed ought to be the case. We asked how much respondents agreed with the following statement: "women should have the right to vote if they so choose." Respondents could "strongly disagree", "disagree", "neither strongly disagree nor agree", "agree", or "strongly agree". Second, we asked respondents if they believed on a 5-point scale that "women should be able to participate in gatherings to discuss politics", with answers ranging between "definitely no", "probably no", "maybe", "probably yes", and "definitely yes". Third, using the same scale, we asked if "women should be able to run for office." Respondents were asked to give their personal beliefs about these questions, and they were answered in the sequence presented here. I code this scale from zero to four in the following analyses. I use the sum of these beliefs in order to simplify later analyses and call this the *female belief index* or the *male belief index*, depending on respondent.

#### 3.3.2.2 Normative and empirical expectations

After asking about support for each of the three forms of female political participation, we ask respondents to consider 10 typical women and 10 typical men in their community and estimate how many of them would state support each of those three behaviors. This captures a respondent's *normative* expectations—the share of other people in the immediate community that are believed to believe women *ought* to be able to engage in these forms of political participation. Therefore, for each respondent I have three sets of normative expectations, each ranging from 0 to 10. To simplify some later analyses, I may sum these values to create an index of normative expectations so that the final *female normative expectation* and *male normative expectation* indices range from 0 to 30. We also ask respondents about their *empirical* expectations. We ask all respondents how many women, out of a typical 10, will self-report voting in the 2018 general elections. This variable ranges from 0 to 10 and I denote this the *female turnout prediction* or *male turnout prediction* in below tables and figures.

### 3.3.3 Self-reported turnout

I measure female political participation using self-reported turnout. We ask whether or not the respondent voted in the last general election in 2018—*self-reported turnout*—and code the response as one if they answer in the affirmative and zero otherwise. Survey respondents often over-report turnout (Karp and Brockington, 2005), which can distort studies of political mobilization (Vavreck et al., 2007). In the absence of validated turnout data, such as in Pakistan, self-reports are often the only measure of political participation that can be obtained.[10] In our survey, we asked individuals if they had voted in the most recent 2018 general election (an election about 8 months preceding the survey), and coded all answers that were affirmative as ones and all other answers as zeros.

In the 37 settlements where we work, the female heads of household that we interview self-report voting at an unlikely rate of 69.7 percent. Given national turnout was 51.6 percent of registered voters, it is unlikely that this self-reported turnout rate is accurate. While the respondents are heads of household, who may be more likely to turn out, it is unlikely that this overcomes the fact that they are women—who tend to turn out at far lower rates—and that the denominator in this sample is all female heads of household, not just registered voters.

[10] Recent efforts to measure turnout by observing inked fingers, a practice used to discourage double voting, may be promising if surveys can be fielded immediately after elections.

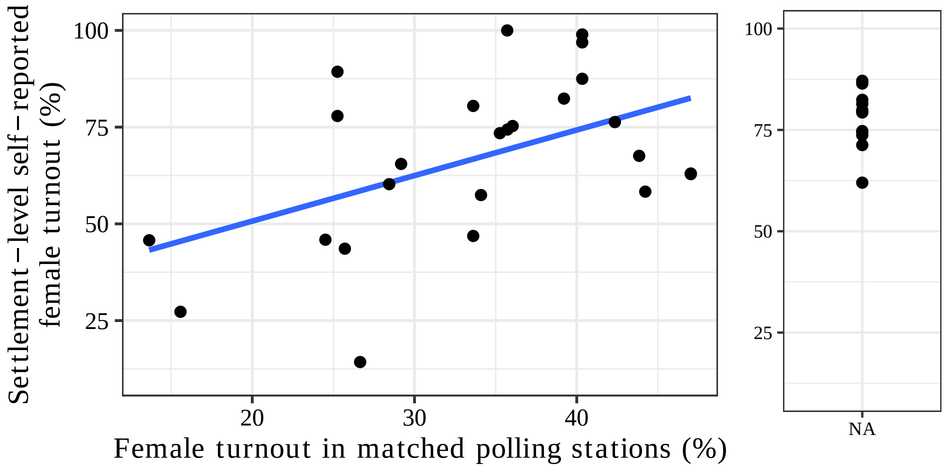
Nonetheless, does self-reported turnout capture something meaningful about past behavior? Or does it at least capture some combination of actual behavior and willingness to state a behavior that may, in some of the contexts we consider, not be socially desirable? I leave discussing the possible interpretations of this measure to the discussion of results below, and for now report that this self-reported measure is correlated with the aggregate-level female turnout in the settlements for which that data are available. In Figure 3.2, I compare the official reported female turnout as a percent of registered voters on the horizontal axis with the average of self-reported turnout in the corresponding settlement on the vertical axis. Note that there is not a one-to-one mapping of settlements to polling station areas—some settlements have multiple polling stations and some settlements will share a polling station with other settlements. Furthermore, 12 of the 37 settlements do not have available, gender-disaggregated voting totals—data exist for the other 25 because the polling stations are gender-disaggregated and thus turnout at the polling station level is the same as turnout by gender. The average self-reported turnout in these 12 settlements is represented on the right panel of Figure 3.2.

Overall a correlation of 0.49 exists between self-reported turnout among female heads of households and aggregate female turnout at the polling station level. The OLS fit returns a slope of roughly one, indicating that an increase in 10 percentage points in self-reports corresponds to an increase in 10 percentage points in actual turnout. This indicates that while there is serious over-reporting, on average overreporting is happening to a similar degree across settlements that have high and low overall female turnout.

### 3.3.4 Other data

To collect some data on socioeconomic status, we ask women to place their household into an *income bin*. This question had four categories, pertained to the monthly income of the household, and we code it here from one to four. We also ask whether or not the household owns a motorized vehicle and code them as one if they mention any kind of motorized vehicle

Figure 3.2: Self-reported female turnout and matched polling station female turnout




Each point represents a settlement. The vertical axis is the average self-reported turnout of female heads of household within that settlement and the horizontal axis is the average female turnout in polling stations that serve at least some of that settlement. The right panel represents settlements for which no gender-disaggregated turnout data is available.

and 0 otherwise. Together these two questions serve to measure the economic standing of the household. We also ask the woman if she knows whether a major political party, the PML-N, is in the provincial government or not. If they answer the question correctly, they are coded as a one and as a zero otherwise. I take this as a measure of *political knowledge*.

### 3.3.5 Centrality

In Figure 3.3, I present averages of the socioeconomic variables and personal normative beliefs by binned *indegree centrality* of the household as nominated by other women in the network. All responses in this table are those reported by the female heads of household. In social networks, the *centrality* of an individual is often a key predictor of how influential they are on the behavior of others in the community (Breza and Chandrasekhar, 2019) and whether information that enters a community is spread widely and new behaviors are adopted (Banerjee et al., 2013). These individuals are often considerably different along several dimensions, as their centrality in the network is endogenous to their household characteristics. With respect to politics, more central individuals are more likely to be involved in informal political exchanges and are more likely to win elections, independent of wealth and previous electoral success (Cruz, Labonne and Querubn, 2017).

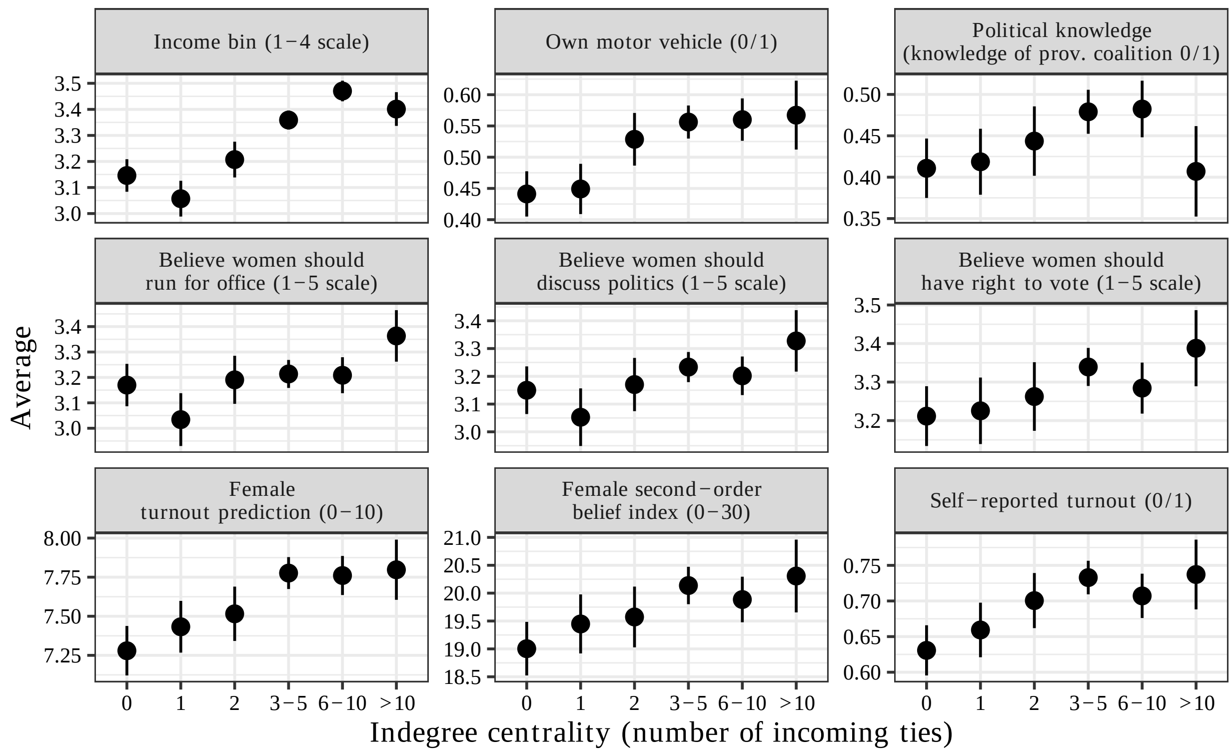
Furthermore, one of the implications of female turnout in conservative societies as a coordination game is that more connected women will both have more information about who will vote and be more likely to be connected to a woman who does vote. As a result, I expect that more connected women self-report voting more often.

In the first row of Figure 3.3, I present the socioeconomic and political knowledge variables. More central households—those with greater *indegrees*, i.e., mentions by other households—self-report higher monthly income, are more likely to own a motor vehicle, and have somewhat greater levels of political knowledge except for the most central households, a puzzling result. In row two, there is no clear relationship between centrality and high levels of personal support for female political participation, although the most connected households do have the highest level of support for all three forms of participation. The third row demonstrates that the normative and empirical expectations are more strongly correlated with indegree centrality, as is self-reported turnout. If a woman's indegree centrality was not predictive of her behavior, then it would be implausible that voting is a coordination game—it is much more difficult for a woman with no connected households to coordinate to go to the polls.

## 3.4 Results: direct and network correlates of self-reported turnout

I extend the classic linear network autocorrelation model (see Ord, 1975; Doreian, 1981) in a Bayesian framework to test whether empirical expectations or normative beliefs are more tightly associated with self-reported voting behavior, conditional on other potential covariates. I describe the model I employ before turning to the results.

Figure 3.3: Female centrality, demographics, political engagement, and political beliefs




Each panel summarizes a different covariate on the baseline survey. Each point is the average of that outcome for households with a particular indegree, indicated on the horizontal axis. The line ranges are 95 percent confidence intervals around the mean.

### 3.4.1 Building on network autocorrelation models

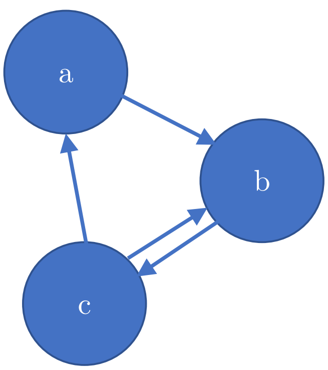
Due to the likelihood of network autocorrelation and the non-independence of error shocks over the social network, I model survey responses using a modified linear network autocorrelation model. The traditional model is,

$$\mathbf{y} \sim MVN(\rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta}, \mathbf{I}\sigma^2), \tag{3.1}$$

where $\mathbf{y}$ is the outcome vector of interest, $\mathbf{W}$ is a weight matrix that captures the social network structure (often a row-normalized adjacency matrix), $\rho$ is the estimated autocorrelation of the outcome, $\mathbf{X}$ is the usual $N$ by $P$ data matrix, $\boldsymbol{\beta}$ are the $P$ direct effects, and $\sigma$ is the constant variance of the disturbances. Usually, the covariate of interest is $\rho$, which measures whether or not there is any network effect and whether the auto-correlation is positive or negative.

An important choice in this model is the specification of $\mathbf{W}$—this choice essentially encodes how variables across households relate to one another and only the magnitude of this relationship is left for estimation. Take the network in Figure 3.4 for example. In this network, there are three households, (a), (b), and (c). In this network, the arrows point from the household that nominates to the household that is nominated. In this example, (a) said they were connected to household (b), while (c) said they were connected to both households.

Figure 3.4: Example of a directed network




(a) A simple directed network

$$\mathbf{W} = \begin{array}{c c} & \begin{array}{ccc} a & b & c \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \end{array}$$

(b) The corresponding (directed) adjacency matrix

$$\mathbf{W} = \begin{array}{c c} & \begin{array}{ccc} a & b & c \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.5 & 0.5 & 0 \end{pmatrix} \end{array}$$

(c) The corresponding (directed) row-normalized adjacency matrix

The $w_{ij}$th element of a directed adjacency matrix represents a tie *from* household $i$ to household $j$. In traditional linear network autocorrelation models there are many ways to specify the weight matrix rather than the simple adjacency matrix. One common choice is to row-normalize $\mathbf{W}$ so that each row sums to 1, as is shown in Figure 3.4c. The choice of $\mathbf{W}$ implies a particular form of network disturbance; if a weight matrix is not row-normalized, then individual $i$ with many incoming ties will have a greater value of $\mathbf{w}_i^\top \mathbf{y}$ than someone with only one tie. Thus, a row-normalized weight matrix, such as the one I employ here, essentially models network disturbances as an equal pressure coming from all an individual's ties, but that the sum of those pressures is equal across individuals no matter how many ties they have.[11] If one believed peers have additive rather than average affects on one's outcomes, then data should not be normalized. For further discussion of the choices a researcher can make when specifying $\mathbf{W}$ in a linear network autocorrelation model, see Leenders (2002).

This class of model allows for answers to the following kind of question: "what is the magnitude of the network-based correlations in the data, conditional on individual-level covariates" and vice versa. These models are analogous to spatial autocorrelation models, where the distance between units is geographic rather than social, but are specified in the same way.

I extend the classic linear network autocorrelation model in two ways. First, I aim to study whether there are cross-covariate network effects by including various network effect measures $\rho_k$ for some set of covariates $k$ in $1, \ldots, K$.[12] Second, I specify the hierarchical model for each settlement $s \in \{1, \ldots, S\}$ as

$$\mathbf{y}_s \sim MVN\left(\sum_{k=1}^{K} \rho_{ks} \mathbf{W}_s \mathbf{z}_{ks} + \mathbf{X}_s \boldsymbol{\beta}_s, \mathbf{I}\sigma^2\right), \tag{3.2}$$

$$\rho_{ks} \sim N(\gamma_k, \tau_k) \ \ \forall k \in \{1, \ldots, K\}, \tag{3.3}$$

$$\beta_{ps} \sim N(\theta_p, \zeta_p) \ \ \forall p \in \{1, \ldots, P\}. \tag{3.4}$$

In this model, $\mathbf{z}_{ks}$ is a column vector for household covariate $k$ within settlement $s$ that may be correlated through the social network—represented by $\mathbf{W}_s$—with $\mathbf{y}_s$ (including $\mathbf{y}_s$ itself). For example, in order to estimate whether the stated support for female political participation in socially adjacent households is correlated through the network with self-

[11]It is also possible to model this using a weight matrix that more heavily values contributions of more central individuals, say by element-wise multiplication of $\mathbf{W}$ by that row's eigenvector centrality.

[12]This is not dissimilar from taking weighted averages of neighbors covariates and including them in $\mathbf{X}$.

reported turnout, I set the stated support for female participation as one of the $k$ column vectors $\mathbf{z}_{ks}$ and self-reported turnout as $\mathbf{y}_s$. In this model $\rho_{ks}$ and $\beta_{ps}$ are settlement-level parameters drawn from sample-level parameters $\gamma_k$ and $\theta_k$. I set weak, non-informative priors on these sample parameters and the choice of weak prior is not important to the results presented below.

### 3.4.2 Correlates of self-reported turnout

In Table 3.2, I present the results of fitting this model where self-reported turnout in the 2018 elections is the outcome. The main purpose of this table is to understand how beliefs and behavior are correlated, with a specific eye on what kinds of beliefs and whose beliefs are most strongly correlated with behavior. Specifically, I focus on direct correlations of five kinds of variables with self-reported turnout: (1) socio-economic status and network position; (2) personal normative beliefs and normative and empirical expectations; (3) the belief set of the male head of household; and (4) the behavior of women and personal normative beliefs of men and women in the female-nominated social network.

In Table 3.2, I report $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$, the mean of the sample-level parameters from which the settlement-level parameters $\boldsymbol{\beta}$ and $\rho_k$ are drawn. In column (1) of Table 3.2 I include only the economic status of the household, network centrality of the household in the female-nominated network, here measured as the eigenvector centrality[13], and the political knowledge of the household. In this model, the 95 percent credible interval of the parameters on political knowledge and eigenvector centrality are strictly positive, with substantially large effects. Women in isolated households—those with no connections—report voting at 23.9 percentage point (pp) lower rates than the most connected households, conditional on the model; the range of the centrality measures is 0 to 1, although the majority of the units are between 0 and 0.4. Furthermore, women who know the composition of the provincial

[13]Eigenvector centrality essentially weights connections to other connected households more heavily than connections to less connected households.

government are 4.1 pp more likely to self-report voting. These are unsurprising results—women who are connected to other households are more likely to vote, as are women with higher levels of political knowledge. They are also consistent with voting as a coordination game. Women in isolated households will find it more difficult to find another individual with whom they can vote.

In columns (2) and (3) I report the main results on social norms and political participation. In both models, a woman's personal normative beliefs and empirical expectations are likely to have positive relationships with the outcome, while the credible regions for a woman's normative expectations contain zero. The *Fem. belief index* ranges from 0 to 12 as it is the sum of the three four-point personal belief questions; the standard deviation of this scale is 2.52. The parameter estimated here indicates women who are one standard deviation more likely to support female political participation are 8.6 percentage points more likely to self-report turning out. Similarly, the empirical expectations, *Fem. turnout prediction*, are on a scale from 0 to 10 and count the number of women, out of 10 women in the community, that they believe self-report voting. Therefore, each unit increase corresponds to a 10 pp increase in a woman's empirical expectations. The estimated parameter indicates that a woman who believes 10 percent more women around her self-report voting is 1.9 pp more likely to self-report voting. Meanwhile, the two sets of normative expectations—those about other women *Fem. normative expectations index (F)* and those about other men *Fem. normative expectations index (M)*—have 95 percent credible intervals that contain 0 and point in opposite directions. In other words, statistical uncertainty remains about the relationship between normative expectations and self-reported voting; the posterior distribution of these parameters contains a non-negligible mass on both positive and negative values.

In addition to increased statistical uncertainty about the relationship between normative expectations and self-reported voting, the standardized median posterior is also smaller than the median posterior on empirical expectations. In column (4), I present the median posterior values standardized by the standard deviations of the covariates. In standard deviation units,

Table 3.2: Correlates of self-reported turnout in a modified network auto-correlation model

| | Outcome = self-reported turnout | | | |
|---|---|---|---|---|
| | Covariates in original scaling | | | Standardized scaling |
| | (1) | (2) | (3) | (4) |
| **Direct effects $\gamma$** | | | | |
| Intercept | 0.566* | 0.112 | −0.017 | |
| | (0.462, 0.673) | (−0.034, 0.254) | (−0.217, 0.189) | |
| Income bin | 0.019 | 0.022 | 0.021 | 0.016 |
| | (−0.004, 0.043) | (−0.001, 0.045) | (−0.002, 0.044) | |
| Own motor vehicle | 0.030 | 0.013 | 0.013 | 0.006 |
| | (−0.002, 0.062) | (−0.017, 0.044) | (−0.016, 0.043) | |
| Network centrality | 0.239* | 0.214* | 0.203* | 0.034 |
| | (0.139, 0.341) | (0.124, 0.308) | (0.115, 0.294) | |
| Pol. knowledge | 0.041* | 0.021 | 0.020 | 0.010 |
| | (0.008, 0.073) | (−0.017, 0.059) | (−0.016, 0.055) | |
| Fem. belief index | | 0.034* | 0.034* | 0.085 |
| | | (0.023, 0.045) | (0.024, 0.044) | |
| Fem. turnout prediction | | 0.019* | 0.019* | 0.037 |
| | | (0.004, 0.034) | (0.004, 0.033) | |
| Fem. normative expectations index (F) | | −0.004 | −0.003 | −0.019 |
| | | (−0.012, 0.004) | (−0.009, 0.003) | |
| Fem. normative expectations index (M) | | 0.003 | 0.003 | 0.019 |
| | | (−0.003, 0.011) | (−0.003, 0.009) | |
| Male belief index | | | −0.001 | −0.002 |
| | | | (−0.008, 0.007) | |
| Male turnout prediction | | | 0.009 | 0.017 |
| | | | (−0.001, 0.019) | |
| Male normative expectations index (F) | | | −0.002 | −0.010 |
| | | | (−0.006, 0.003) | |
| Male normative expectations index (M) | | | 0.002 | 0.008 |
| | | | (−0.003, 0.007) | |
| **Network effects $\theta$** | | | | |
| Self-reported turnout | | | 0.185* | 0.047 |
| | | | (0.113, 0.258) | |
| Fem. belief index | | | −0.004 | −0.007 |
| | | | (−0.015, 0.006) | |
| Male belief index | | | −0.001 | −0.001 |
| | | | (−0.013, 0.011) | |
| N | 4348 | 4348 | 4348 | 4348 |
| S | 37 | 37 | 37 | 37 |

This table reports the median posterior values of $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ from Equation 3.3. Under "Direct effects" I report the median $\boldsymbol{\gamma}$ posterior values and under "Network effects" are the median $\boldsymbol{\theta}$ posterior values. Direct effects measure the relationship between a respondent's own covariates and their self-reported turnout while the network effects measure the relationship between their social peers' covariates and their responses. Underneath the median posterior values are 95 percent credible intervals in parentheses and asterisks next to the median posterior values indicate credible intervals that exclude zero. Columns (1) through (3) report median three separate models with all of the covariates in their original scaling while column (4) reports the same model as that in column (3) but with the median posterior values rescaled to be standardized and more easily comparable across covariates.

the posterior median of the parameter on both male and female normative expectations (-0.019 and 0.019) are less than one-third of the posterior median parameter on the personal normative beliefs (0.085) and are roughly half of the posterior median parameter on empirical expectations (0.037). The magnitude of the relationship between empirical expectations and self-reported turnout is twice as large as the relationship between normative expectations and self-reported turnout.

Thus, column (2) presents clear evidence that empirical expectations and personal normative beliefs are more strongly correlated with self-reported turnout than normative expectations in this sample. Women are willing to self-report turning out to vote when they believe women ought to have the right to political participation, conditional on their expectations. In other words, expectations about others in one's community are not sufficient to explain self-reported behavior; women who self-report voting are not only responding to social cues but also believe in their right to vote. Furthermore, there is no evidence that expectations about men are any more or less important than expectations about women.

In column (3), I add potential within-household and network effects. I add the belief set of the male head of household as well as the network effects of women in adjacent households self-reporting turnout, and the personal normative beliefs of adjacent households. Again, when considering direct effects, woman's personal normative beliefs and their empirical expectations remain the only strong predictors of self-reported turnout. The belief set of the male head of household does not have any clear relationship with self-reported turnout among women. Furthermore, the network effects demonstrate that women whose socially connected households have women who do not vote at all self-report turning out at a 18.5 pp lower rate than women whose socially connected neighbors all vote. This overstates the likely range as the standard deviation of the network-weighted average of socially adjacent women who vote is 0.25. So moving one standard deviation up in the share of network-adjacent women who vote corresponds to a 4.6 pp greater rate of self-reported turnout. There are no such network effects of the personal normative beliefs of the men and women in adjacent households.

Column (3) presents evidence that the behavior of one's neighbors, and the predictions about the behavior of one's neighbors, is more strongly correlated with self-reported turnout than the normative beliefs of one's neighbors or the beliefs about the beliefs of one's neighbors. Not only do posterior distributions on social neighbors' normative beliefs contain zero, the magnitude of the median posterior value is significantly smaller than the value on social neighbors' voting behavior, as reported in Column (4). Taken together, these last set of results point to the strength of empirical expectations and the self-reported behavior of social neighbors in explaining which women vote. Conditional on a host of personal covariates, women who vote are far more likely to be in a social network of other women who vote, rather than a social network of other women who support a woman's right to vote.

This is consistent with the theoretical expectation that voting as a coordination game invites a stronger relationship between empirical expectations and behavior. The network correlation of self-reported turnout is also consistent with a model of female political participation in conservative societies as a coordination game.

There are three main caveats. First, these are not causal estimates of the effect of beliefs on voting. These results instead capture what kinds of beliefs and self-reported behavior co-exist in networks. I make claims that these data show women who believe that other women vote, even if other women may not support a woman's right to vote, are the same women who self-report voting.

Second, the election women report voting in precedes their belief elicitation, and observing other women voting is likely to increase one's empirical expectations. Indeed, these results are indications that this may be true. Unfortunately, I cannot rule out the possibility that women used to believe no women were going to turn out, but then when they voted they realized how many other women were voting and thus updated their empirical expectations about the behavior of others. However, if I take this view of results, than we also have to believe that these same women are not updating their beliefs about how women *ought* to behave. In this case, the central finding that women who report turning out to vote are women who believe other women turn out to vote, independent of whether they believe others think they should vote, remains.

Third, there is strong correlation between a woman's personal normative beliefs and her normative expectations. There is also a strong correlation between a woman's personal normative beliefs and those of her social network. Therefore, normative expectations may still be playing a role in predicting behavior if the majority of the relationship is captured by the effect on personal normative beliefs. While studying such dynamics would require different data than that at hand, as a robustness check I omit a woman's personal normative beliefs from the models estimated in columns (2) and (3). Even with this omission, the credible intervals for the expectations and the network effects of the personal normative beliefs continue to contain zero and the magnitude is little affected. These results are presented in Appendix 3.A. There is little evidence that the relationship between normative expectations and behavior is suppressed by the relationship between a woman's private normative beliefs and turnout.

### 3.4.3 Mechanism: social norms license voting behavior

One potential explanation for a weak relationship between normative expectations about the right to vote and voting may be that these expectations only *permit* women to vote and are not sufficient to induce a woman who is uncertain about the behavior to vote. This secondary hypothesis that norms have a greater impact when they *license* the behavior of an individual can be tested using a simple interaction.

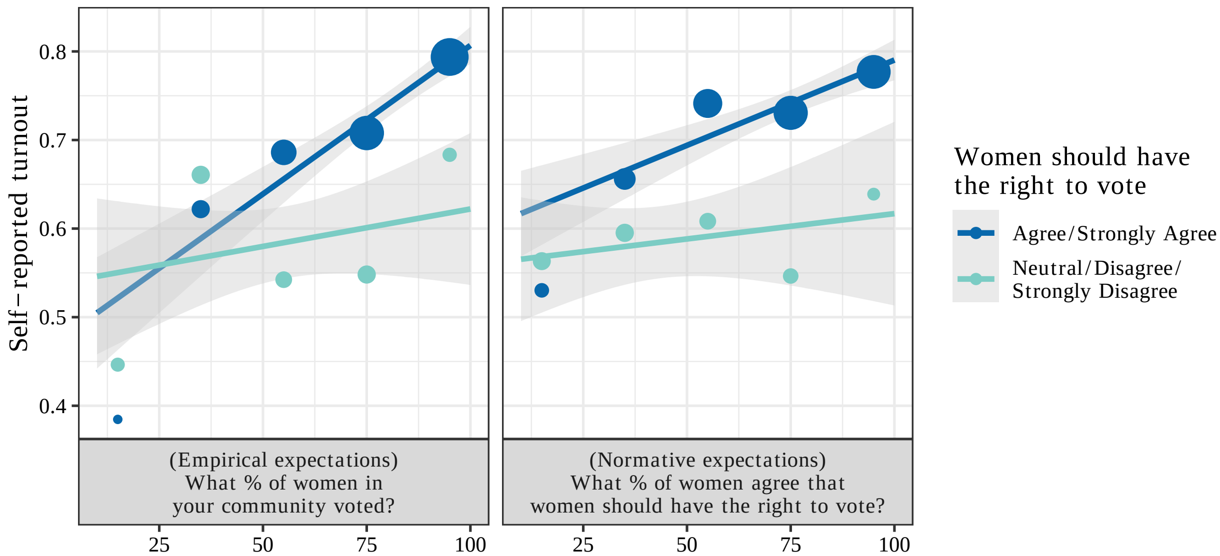
If a woman in this sample believes that women ought to have the right to vote, then learning that other women in their community vote or support the right to vote licenses them to engage in that behavior. Women who are unsure about women's suffrage are unlikely to be compelled to vote when they learn that others support a woman's right to vote. Although there is no change in expectations captured by this survey, the static predictions remain largely the same. Women who believe that women should have the right to vote will be more likely to vote when they also believe other women support their behavior in either their actions or their beliefs.

To test this hypothesis, I present simple binned averages of self-reported voting by both expectations and personal normative beliefs in Figure 3.5. The x-axes are empirical and normative expectations, respectively. Respondents selected between 1 and 10, and I have multiplied it by 10 here for simplicity. The vertical axis is self-reported turnout, averaged over bins of the horizontal axis, where every two values are binned together—e.g. 10 and 20 percent are averaged and placed into one bin, 30 and 40 into another, and so on. The binned averages are also calculated by personal normative beliefs, with the dark blue circles representing women who agree or strongly agree that women should have the right to vote while the light blue circles represent women who neither agree nor disagree, or disagree that women should have the right to vote. The size of the circle corresponds to the number of respondents who are in that expectation and personal belief group. The lines represent simple linear OLS estimates by group with 95 percent confidence intervals shaded in grey.

The slope of both lines is greater for women who agree that women should have the right to vote and is flatter for women who do not agree women should have the right to vote. This supports the licensing hypothesis, where women are more responsive to widespread approval of a norm when they themselves approve of the norm. Women who do not approve of the norm are less likely to vote than women who do, even if others engage in or approve of female political participation. The difference in slopes is substantively significant; women with very low expectations self-report voting around 55 to 60 percent of the time regardless of their personal beliefs, but women with very high expectations who support a woman's right to vote are 20 percent more likely to vote than women who who are uncertain or disagree with a woman's right to vote. Nonetheless, the difference in slopes is only significant at the 10 percent level for both kinds of expectations when not clustering standard errors; when standard errors are clustered by settlement the difference is not statistically significant.

In another context where political participation among women is much higher, it is possi-

Figure 3.5: Self-reported turnout and the interaction between personal beliefs and expectations




Each point is a set of female respondents grouped by their personal support for a woman's right to vote and their expectations, binned by values of 20. For example, all women who believe 0-20 percent of women ought to have the right to vote and who personally agree women should have the right to vote are in one point. The size of the point corresponds to the number of women in that bin. The vertical axis is the share of women who self-report voting in that group and the horizontal axis is their binned expectations. The lines represent lines of best fit with 95 percent confidence intervals computed using from robust OLS standard errors.

ble that high empirical expectations would have a different relationship with voting behavior. In that context, high empirical expectations might capture that a woman expects women to turn out and thus may be punished for not turning out herself. However, in this case where female political participation is traditionally quite low, high empirical expectations appear to signal permission rather than a requirement that one should vote.

## 3.5 Targeting future interventions

Taken together, the strong relationship between empirical expectations and the behavior of one's peers with voting, as well as the interaction between personal beliefs and expectations, implies a subsample of women who would be responsive to individualized mobilization efforts: women who themselves support a woman's right to vote and have low normative and empirical expectations of others. Unfortunately, beyond the caveat that these results are static and do not estimate the effects of *changes* in expectations that are causally identified, there are two main problems with these implications.

First, the group that this applies to is quite small. Although the majority of women state they support a woman's right to vote, personal normative beliefs that women should have the right to vote are correlated with expectations. There are few people who are strong supporters of a woman's right to vote who also have low expectations of others in their community. This can be seen in the small size of the leftmost dark blue dots in Figure 3.5 above.

Second, many people already accurately estimate how many others in their community will self-report voting. On the other hand, respondents' normative expectations are well under average normative beliefs of their community. In column 1 of Table 3.3, I report the average level of personal normative beliefs of women and men, as well as self-reported turnout. This percent who agree with the three statements is a simple recoding of the 0 to 4 scales presented above. Answers are coded as "agreement" if they choose either "probably yes" and "definitely yes" (or analogously, "agree" and "strongly agree" to the voting question) to the three questions introduced above. Panel A contains averages among female respondents, and Panel B contains averages among male respondents. In Panel A, I present information about women's beliefs. For example, 82.4 percent of women agree that women should be discussing politics. In Panel B, I present information about men's beliefs. For example, 75.6 percent of men agree that women should be discussing politics. In general, men express greater levels of support for women's actual participation in politics through voting or candidacy, but are less supportive of women discussing politics. In general, there is fairly high support for female political participation across all three questions, and across both men and women.

In columns 2 and 3 of Table 3.3, I report the normative expectations of women and men in Panels A and B and I report their empirical expectations in Panel C. In Panel A, I

Table 3.3: Underestimation of community support for female political participation

**Panel A: On Women's Beliefs**

| | **Women's personal beliefs** | **Normative expectations about women** | |
|---|---|---|---|
| Women should ... | % women agree | Avg. woman's prediction for "What % of women in your community agree?" | Avg. man's prediction for "What % of women in" your community agree?" |
| ...discuss politics | 82.4 | 63.9 | 49.6 |
| | (0.6) | (0.4) | (0.4) |
| ...have the right to vote | 84.5 | 69.0 | 62.6 |
| | (0.6) | (0.4) | (0.4) |
| ...be able to run for office | 83.2 | 64.9 | 54.4 |
| | (0.6) | (0.4) | (0.4) |

**Panel B: On Men's Beliefs**

| | **Men's personal beliefs** | **Normative expectations about men** | |
|---|---|---|---|
| Women should ... | % men agree | Avg. woman's prediction for "What % of men in your community agree?" | Avg. man's prediction for "What % of men in" your community agree?" |
| ...discuss politics | 75.6 | 58.3 | 40.1 |
| | (0.7) | (0.4) | (0.3) |
| ...have the right to vote | 93.9 | 64.0 | 58.5 |
| | (0.4) | (0.4) | (0.4) |
| ...be able to run for office | 85.0 | 57.9 | 46.8 |
| | (0.6) | (0.4) | (0.3) |

**Panel C: On Women's Behavior**

| | **Reported behavior** | **Empirical expectations** | |
|---|---|---|---|
| | % of women | Avg. woman's prediction | Avg. man's prediction |
| Self-reported turnout | 70.3 | 76.2 | 66.6 |
| | (0.7) | (0.3) | (0.3) |

Standard errors reported in parentheses. The personal beliefs are the percent of respondents who either "agree" or "strongly" agree with statements of support for the different kinds of female political participation. The expectations are averages of predicted percentages derived from the question: "How many other [men/women] in your settlement would agree with the statement ...".

present the expectations of women and men about other women. In Panel B, I present the expectations of women and men about other men. For example, on average, women believe that 63.9 percent of other women support women discussing politics, while, on average, men believe that only 49.6 percent of other men support women discussing politics. Both estimates are considerably lower than women's stated level of support for women discussing politics, 82.4 percent. This pattern repeats across all outcomes, and across men and women's beliefs. Both women and men underestimate the stated level of support for female political participation among both women and men. In general, men predict lower levels of support for female political participation than do women.

In Panel C, I consider the empirical expectations of men and women in this sample, and compare them to the self-reported level of voting. While 70.3 percent of women self-report voting, women estimate that 76.2 percent of other women will self-report voting, while men estimate that 66.6 percent of women will self-report voting. The large gap in perceived normative beliefs is not replicated here, although if one considers actual turnout rates among women, respondents are still overestimating the behavioral norm. However, again if the results are taken at face value, it appears there is stronger evidence for a clear gap between normative beliefs than between behavior and expectations about that behavior.

Of course, there are several caveats to this result. First, our measurement of personal beliefs is not incentivized, and it is possible that respondents are over-reporting their support for female political participation. However, in this case it is not entirely clear which direction respondents would be biased as there are some broad, societal pressures that may make the more conservative viewpoint "desirable." Unfortunately, however, this bias will not be distributed uniformly in our sample, and may correlate with some of our other variables of interest. Nonetheless, the answers can also be taken at face value and understood as a combination of what people believe and what they think they should say to an enumerator. In that situation, the questions still capture some sense of what a respondent believes *ought* to be the case, if not their true, personal beliefs. Furthermore, the gap between stated beliefs and what they believe about their community should be largely unaffected by these problems. Because we explicitly say "when asked about whether they agree with [behavior], how many out of 10 women in your community would say they do agree", we are asking respondents to predict the responses of others and thus the comparison of their answers with the beliefs about others are regarding the same understanding of what *ought* to be.

Given that there is little to no gap between self-reported turnout and expectations about self-reported turnout, an intervention designed around empirical expectations may have little impact. Unlike studies of drinking among adolescents, where adolescents greatly overestimate the incidence of drinking among their peers and thus are likely to update when told the actual incidence (Schroeder and Prentice, 1998), individuals here are unlikely to update when the new information fits their prior expectations so closely. However, the large gap between normative expectations and normative gaps allows for more updating if an intervention was to correct misperceived norms. There is evidence for this kind of intervention improving female labor force participation among women in Saudi Arabia; when men were told other men supported female labor force participation more than they anticipated, they were more likely to support women in their households searching for jobs (Bursztyn, González and Yanagizawa-Drott, 2018).

## 3.6 Discussion

In a historically conservative society, female nominated social networks and female beliefs and expectations play an important role in explaining their self-reported political participation. Expectations that other women will vote and the self-reported behavior of other women both are strong predictors of a woman reporting she voted. The relationship between normative expectations—what a woman thinks others think *ought* to be—and voting is weaker.

Interestingly, the role of men is muted in this analysis. Conditional on a woman's personal beliefs, the centrality of the household, and the behavior of a woman's peers, neither the beliefs or expectations of the male head of household nor the beliefs of men in socially adjacent households are strongly predictive of self-reported political behavior. Of course, this is not evidence that men are not important gatekeepers of female political behavior in this sample; rather it shows that women's beliefs, expectations, and their networks are sufficient to explain their behavior. The extent to which those are shaped by men is outside the scope of this paper, but may clarify the ways in which men constrain female political participation in Pakistan.

The strength of the relationship between empirical expectations and female political participation is consistent with viewing women voting in a conservative society is a coordination game. Given that voting is made culturally and logistically easier by other women also voting, having social neighbors who votes not only signals that voting is acceptable but lowers costs to voting. This may explain why such large spillover effects exist in simple door-to-door knowledge-based mobilization efforts in Pakistan (Giné and Mansuri, 2018). However, because women appear to have empirical expectations in line with the reported behavior of their peers, interventions focused on changing empirical expectations appear to have a limited audience. Due to the large documented gap in normative expectations and average normative beliefs, an intervention aimed at increasing normative expectations may actually be more efficacious despite a weaker relationship between normative expectations and behavior in this sample.

This paper also provides evidence that these social norms about the right to vote in conservative societies largely *license* voting behavior—in other words, they permit women who believe they should have the right to vote to vote. Other women who are unsure about their right to vote may not be responsive to permissive social norms. Furthermore, because the evidence presented here is observational, women with low normative and empirical expectations may also lack skills, financial resources, or ability to coordinate that could be necessary conditions for female political participation in conservative societies (see Prillaman, 2018). Future work should test whether correcting incorrect normative expectations is sufficient to increase female political participation or whether complimentary skill and resource interventions are also necessary.

# Appendix

## 3.A Robustness of network correlation model

This section contains additional estimates from the hierarchical Bayesian network autocorrelation model. In Table 3.A.1, I present estimates from the same models as in columns (2) and (3) in Table 3.2 although omitting the personal normative beliefs.

Table 3.A.1: Correlates of self-reported turnout in a modified network auto-correlation model: robustness to exclusion of personal normative beliefs

| Outcome = self-reported turnout | | |
|---|---|---|
| | (1) | (2) |
| **Direct effects** $\gamma$ | | |
| Intercept | 0.298* | 0.083 |
| | (0.151, 0.442) | (−0.123, 0.287) |
| Income bin | 0.021 | 0.019 |
| | (−0.003, 0.046) | (−0.004, 0.044) |
| Own motor vehicle | 0.020 | 0.019 |
| | (−0.011, 0.050) | (−0.012, 0.050) |
| Network centrality | 0.230* | 0.216* |
| | (0.140, 0.324) | (0.126, 0.309) |
| Pol. knowledge | 0.018 | 0.018 |
| | (−0.019, 0.057) | (−0.019, 0.055) |
| Fem. turnout prediction | 0.017* | 0.017* |
| | (0.002, 0.031) | (0.003, 0.030) |
| Fem. normative expectations index (F) | 0.003 | 0.004 |
| | (−0.004, 0.011) | (−0.002, 0.010) |
| Fem. normative expectations index (M) | 0.004 | 0.003 |
| | (−0.002, 0.011) | (−0.002, 0.009) |
| Male belief index | | 0.001 |
| | | (−0.007, 0.008) |
| Male turnout prediction | | 0.009 |
| | | (−0.001, 0.019) |
| Male normative expectations index (F) | | −0.002 |
| | | (−0.007, 0.003) |
| Male normative expectations index (M) | | 0.001 |
| | | (−0.004, 0.006) |
| **Network effects** $\theta$ | | |
| Self-reported turnout | | 0.198* |
| | | (0.122, 0.273) |
| Fem. belief index | | 0.002 |
| | | (−0.009, 0.013) |
| Male belief index | | 0.001 |
| | | (−0.011, 0.013) |
| N | 4348 | 4348 |
| S | 37 | 37 |

This table reports the median posterior values of $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ from Equation 3.3. Under "Direct effects" I report the median $\boldsymbol{\gamma}$ posterior values and under "Network effects" are the median $\boldsymbol{\theta}$ posterior values. Direct effects measure the relationship between a respondent's own covariates and their self-reported turnout while the network effects measure the relationship between their social peers' covariates and their responses. Underneath the median posterior values are 95 percent credible intervals in parentheses and asterisks next to the median posterior values indicate credible intervals that exclude zero. Columns (1) through (2) report median three separate models with all of the covariates in their original scaling.

Bibliography

Acemoglu, Daron and James A Robinson. 2006. "De facto political power and institutional persistence." *American Economic Review* 96(2):325–330.

Adeney, Katharine. 2017. "How to understand Pakistan's hybrid regime: the importance of a multidimensional continuum." *Democratization* 24(1):119–137.

Afzal, Madiha. 2014. "Do barriers to candidacy reduce political competition? Evidence from a bachelors degree requirement for legislators in Pakistan." *Public Choice* 161(1-2):51–72.

Allen, Geoff and Heather Stoll. 2014. "Representation as a numbers game: The link between legislative size and the representation of minorities and women." Working paper.

Auerbach, Adam Michael and Adam Ziegfeld. 2020. "How Do Electoral Quotas Influence Political Competition? Evidence from Municipal, State, and National Elections in India." *The Journal of Politics* 82(1):397–401.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo and Matthew O Jackson. 2013. "The diffusion of microfinance." *Science* 341(6144):1236498.

Bicchieri, Cristina. 2006. *The grammar of society: The nature and dynamics of social norms.* Cambridge University Press.

Bicchieri, Cristina and Erte Xiao. 2009. "Do the right thing: but only if others do so." *Journal of Behavioral Decision Making* 22(2):191–208.

Blanton, Hart, Amber Köblitz and Kevin D McCaul. 2008. "Misperceptions about norm misperceptions: Descriptive, injunctive, and affective social normingefforts to change health behaviors." *Social and Personality Psychology Compass* 2(3):1379–1399.

Boas, Taylor C, F Daniel Hidalgo and Neal P Richardson. 2014. "The spoils of victory: campaign donations and government contracts in Brazil." *The Journal of Politics* 76(2):415–429.

Bond, Robert M, Christopher J Fariss, Jason J Jones, Adam DI Kramer, Cameron Marlow, Jaime E Settle and James H Fowler. 2012. "A 61-million-person experiment in social influence and political mobilization." *Nature* 489(7415):295.

Breza, Emily and Arun G Chandrasekhar. 2019. "Social Networks, Reputation, and Commitment: Evidence From a Savings Monitors Experiment." *Econometrica* 87(1):175–216.

Brollo, Fernanda and Tommaso Nannicini. 2012. "Tying your enemy's hands in close races: the politics of federal transfers in Brazil." *American Political Science Review* 106(4):742–761.

Brollo, Fernanda and Ugo Troiano. 2016. "What happens when a woman wins an election? Evidence from close races in Brazil." *Journal of Development Economics* 122:28–45.

Bursztyn, Leonardo, Alessandra L González and David Yanagizawa-Drott. 2018. Misperceived social norms: Female labor force participation in Saudi Arabia. Technical report National Bureau of Economic Research.

Callen, Michael, Saad Gulzar and Arman Rezaee. 2020. "Can political alignment be costly?" *The Journal of Politics* 82(2):000–000.

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2014. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica* 82(6):2295–2326.

Carnes, Nicholas and Noam Lupu. 2016. "What good is a college degree? Education and leader quality reconsidered." *The Journal of Politics* 78(1):35–49.

Cattaneo, Matias D, Michael Jansson and Xinwei Ma. 2019. "Simple local polynomial density estimators." *Journal of the American Statistical Association* pp. 1–7.

Caughey, Devin and Jasjeet S Sekhon. 2011. "Elections and the regression discontinuity design: Lessons from close US house races, 1942–2008." *Political Analysis* 19(4):385–408.

Caughey, Devin, Yiqing Xu and Christopher Warshaw. 2017. "Incremental democracy: The policy effects of partisan control of state government." *The Journal of Politics* 79(4):1342–1358.

Chandra, Kanchan, ed. 2016. *Democratic Dynasties: State, Party and Family in Contemporary Indian Politics*. Cambridge University Press.

Cheema, Ali, Asad Liaqat and Shandana Khan Mohmand. 2018. "Political Connections and Vote Choice: Evidence from Pakistan." Working paper.

Cheema, Ali, Farooq Naseer and Luke Sonnet. 2020. "Institutional Change and Dynastic Persistence in Pakistan: The Impact of Electoral Reapportionment and Education Minimums." Working paper.

Cheema, Ali, Hassan Javid and Muhammad Farooq Naseer. 2013. "Dynastic politics in Punjab: Facts, myths and their implications." *Institute of Development and Economic Alternatives Working Paper* (01-13).

Cheema, Ali and Muhammad Farooq Naseer. 2013. "Dataset of Dynastic Politicians in Pakistan." Institute of Development and Economic Alternatives.

Cheema, Ali, Sarah Khan, Shandana Khan Mohmand, Anam Kuraishi and Asad Liaqat. 2019. "Pakistan's Participation Puzzle: A Look at the Voting Gender Gap.".
**URL:** *https://www.usip.org/blog/2019/07/pakistans-participation-puzzle-look-voting-gender-gap*

Chhibber, Pradeep. 2013. "Dynastic parties: Organization, finance and impact." *Party Politics* 19(2):277–295.

Cinelli, Carlos and Chad Hazlett. 2020. "Making sense of sensitivity: Extending omitted variable bias." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82(1):39–67.

Cruz, Cesi, Julien Labonne and Pablo Querubn. 2017. "Politician Family Networks and Electoral Outcomes: Evidence from the Philippines." *American Economic Review* 107(10):3006–37.
**URL:** *https://doi.org/10.1257/aer.20150343*

Dahl, Robert Alan and Edward R Tufte. 1973. *Size and democracy*. Vol. 2 Stanford University Press.

Dal Bó, Ernesto, Pedro Dal Bó and Jason Snyder. 2009. "Political dynasties." *The Review of Economic Studies* 76(1):115–142.

de Benedictis-Kessner, Justin. 2018. "Off-cycle and out of office: Election timing and the incumbency advantage." *The Journal of Politics* 80(1):119–132.

De la Cuesta, Brandon and Kosuke Imai. 2016. "Misunderstandings about the regression discontinuity design in the study of close elections." *Annual Review of Political Science* 19:375–396.

Dell, Melissa. 2015. "Trafficking networks and the Mexican drug war." *American Economic Review* 105(6):1738–79.

Doreian, Patrick. 1981. "Estimating linear models with spatially distributed data." *Sociological methodology* 12:359–388.

Eggers, Andrew C, Anthony Fowler, Jens Hainmueller, Andrew B Hall and James M Snyder Jr. 2015. "On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races." *American Journal of Political Science* 59(1):259–274.

Eggers, Andrew C and Arthur Spirling. 2017. "Incumbency effects and the strength of party preferences: Evidence from multiparty elections in the United Kingdom." *The Journal of Politics* 79(3):903–920.

Eggers, Andrew and Jens Hainmueller. 2009. "MPs for Sale? Returns to Office in Postwar British Politics." *American Political Science Review* 103(4):319–42.

Election Commission of Pakistan. 2002. General Elections 2002: Report. Technical report.
**URL:** *https://www.ecp.gov.pk/ge/ge2002vol1.pdf*

Erikson, Robert S, Olle Folke and James M Snyder Jr. 2015. "A gubernatorial helping hand? How governors affect presidential elections." *The Journal of Politics* 77(2):491–504.

European Union. 2002. Pakistan National and Provincial Assembly Election: Final Report. Technical report.

Fafchamps, Marcel and Julien Labonne. 2017. "Do politicians relatives get better jobs? evidence from municipal elections." *The Journal of Law, Economics, and Organization* 33(2):268–300.

Feierherd, Germán. 2020. "How mayors hurt their presidential ticket: Party brands and incumbency spillovers in Brazil." *The Journal of Politics* 82(1):195–210.

Feinstein, Brian D. 2010. "The dynasty advantage: Family ties in congressional elections." *Legislative Studies Quarterly* 35(4):571–598.

Ferraz, Claudio and Frederico Finan. 2011. "Electoral accountability and corruption: Evidence from the audits of local governments." *American Economic Review* 101(4):1274–1311.

Ferreira, Fernando and Joseph Gyourko. 2009. "Do political parties matter? Evidence from US cities." *The Quarterly journal of economics* 124(1):399–422.

Fisman, Raymond. 2001. "Estimating the value of political connections." *American economic review* 91(4):1095–1102.

Fisman, Raymond, Florian Schulz and Vikrant Vig. 2014. "The private returns to public office." *Journal of Political Economy* 122(4):806–862.

Fiva, Jon H and Daniel M Smith. 2018. "Political dynasties and the incumbency advantage in party-centered environments." *American Political Science Review* 112(3):706–712.

Folke, Olle and James M Snyder. 2012. "Gubernatorial midterm slumps." *American Journal of Political Science* 56(4):931–948.

Folke, Olle, Torsten Persson and Johanna Rickne. 2016. "The primary effect: Preference votes and political promotions." *American Political Science Review* 110(3):559–578.

Folke, Olle, Torsten Persson and Johanna Rickne. 2017. "Dynastic political rents? Economic benefits to relatives of top politicians." *The Economic Journal* 127(605).

Fouirnaies, Alexander and Andrew B Hall. 2014. "The financial incumbency advantage: Causes and consequences." *The Journal of Politics* 76(3):711–724.

Fowler, Anthony and Andrew Hall. 2014. "Disentangling the personal and partisan incumbency advantages: Evidence from close elections and term limits." *Quarterly Journal of Political Science* 9(4):501–531.

Fox, Richard L and Jennifer L Lawless. 2005. "To run or not to run for office: Explaining nascent political ambition." *American Journal of Political Science* 49(3):642–659.

George, Siddharth Eapen and Dominic Ponattu. 2019. "Like Father, Like Son? The Effect of Political Dynasties on Economic Development." Working paper.

Gerber, Alan S and Donald P Green. 2017. Field experiments on voter mobilization: An overview of a burgeoning literature. In *Handbook of economic field experiments*. Vol. 1 Elsevier pp. 395–438.

Gerber, Alan S., Donald P. Green and Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *The American Political Science Review* 102(1):33–48.

**URL:** *http://search.proquest.com.ezproxy.royalroads.ca/docview/214437890*

Gerber, Elisabeth R and Daniel J Hopkins. 2011. "When mayors matter: estimating the impact of mayoral partisanship on city policy." *American Journal of Political Science* 55(2):326–339.

Geys, Benny and Daniel M Smith. 2017. "Political Dynasties in Democracies: Causes, Consequences and Remaining Puzzles." *The Economic Journal* 127(605):F446–F454.

Giné, Xavier and Ghazala Mansuri. 2018. "Together we will: experimental evidence on female voting behavior in Pakistan." *American Economic Journal: Applied Economics* 10(1):207–35.

Hall, Andrew B. 2015. "What happens when extremists win primaries?" *American Political Science Review* 109(1):18–42.

Hall, Andrew B and Daniel M Thompson. 2018. "Who punishes extremist nominees? Candidate ideology and turning out the base in US elections." *American Political Science Review* 112(3):509–524.

Hirano, Shigeo. 2011. "Do Individual Representatives Influence Government Transfers? Evidence from Japan." *The Journal of Politics* 73(4):1081–1094.

Iyer, Lakshmi and Maya Reddy. 2013. Redrawing the Lines: Did Political Incumbents Influence Electoral Redistricting in the World's Largest Democracy? Working Paper, 14-051 Harvard Business School.

Jalal, Ayesha. 1990. *The State of Martial Rule: The Origins of Pakistan's Political Economy of Defence*. Cambridge University Press.

Jalal, Ayesha. 1995. *Democracy and authoritarianism in South Asia: A comparative and historical perspective*. Cambridge University Press.

Jalal, Ayesha. 2017. *The Struggle for Pakistan: A Muslim Homeland and Global Politics*. Cambridge: Harvard University Press.

Jones, Philip Edward. 2003. *The Pakistan People's Party: Rise to Power*. Oxford University Press.

Karp, Jeffrey A and David Brockington. 2005. "Social desirability and response validity: A comparative analysis of overreporting voter turnout in five countries." *The Journal of Politics* 67(3):825–840.

Khan, Sarah. 2020. Women in Electoral Politics: An Account of Exclusion. In *Pakistan's Political Parties: Surviving Between Dictatorship and Democracy*, ed. Mariam Mufti, Sahar Shafqat and Niloufer Siddiqui. Georgetown University Press.

Khwaja, Asim Ijaz and Atif Mian. 2005. "Do lenders favor politically connected firms? Rent provision in an emerging financial market." *The Quarterly Journal of Economics* 120(4):1371–1411.

Kjaer, Ulrik and Jørgen Elklit. 2014. "The impact of assembly size on representativeness." *The Journal of Legislative Studies* 20(2):156–173.

Klašnja, Marko and Rocío Titiunik. 2017. "The incumbency curse: Weak parties, term limits, and unfulfilled accountability." *American Political Science Review* 111(1):129–148.

Köbis, Nils C, Jan-Willem Van Prooijen, Francesca Righetti and Paul AM Van Lange. 2015. "Who doesnt?The impact of descriptive norms on corruption." *PloS one* 10(6).

Labonne, Julien, Sahar Parsa and Pablo Querubin. 2019. Political dynasties, term limits and female political empowerment: evidence from the Philippines. Technical report National Bureau of Economic Research.

Lee, David S, Enrico Moretti and Matthew J Butler. 2004. "Do voters affect or elect policies? Evidence from the US House." *The Quarterly Journal of Economics* 119(3):807–859.

Leenders, Roger Th AJ. 2002. "Modeling social influence through network autocorrelation: constructing the weight matrix." *Social networks* 24(1):21–47.

Liaqat, Asad, Ali Cheema and Shandana Khan Mohmand. 2020. Who Do Politicians Talk to? Political Contact in Urban Punjab. In *Pakistan's Political Parties: Surviving Between Dictatorship and Democracy*, ed. Mariam Mufti, Sahar Shafqat and Niloufer Siddiqui. Georgetown University Press.

Liaqat, Asad, Michael Callen, Ali Cheema, Adnan Khan, Farooq Naseer and Jacob N Shapiro. 2019. "Political Connections and Vote Choice: Evidence from Pakistan." Working paper.

Lijphart, Arend and Don Aitkin. 1994. *Electoral systems and party systems: A study of twenty-seven democracies, 1945-1990*. Oxford University Press.

Lin, Winston et al. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics* 7(1):295–318.

Lyon, Stephen M. 2019. *Political Kinship in Pakistan: Descent, Marriage, and Government Stability*. Lexington Books.

Manacorda, M and S Gagliarducci. 2020. "Politics in the Family Nepotism and the Hiring Decisions of Italian Firms." *American Economic Journal: Applied Economics* 12(2):67–95.

Martin, Nicolas. 2015. *Politics, Landlords and Islam in Pakistan*. Routledge.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of econometrics* 142(2):698–714.

Michels, Robert. 1911. *Political Parties*. The Free Press.

Miller, Dale T and Deborah A Prentice. 2016. "Changing norms to change behavior." *Annual review of psychology* 67:339–361.

Mohmand, Shandana Khan. 2019. *Crafty Oligarchs, Savvy Voters: Democracy under Inequality in Rural Pakistan*. South Asia in the Social Sciences Cambridge University Press.

Mosca, Gaetano. 1939. *The Ruling Class*. New York: McGraw-Hill.

Mufti, Mariam, Sahar Shafqat and Niloufer Siddiqui, eds. 2020. *Pakistan's Political Parties: Surviving Between Dictatorship and Democracy*. Georgetown University Press.

Nelson, Matthew J. 2009. *In the Shadow of Shari'ah: Islam, Islamic Law and Democracy in Pakistan*. Columbia University Press.

New York Times Editorial Board. 2018. "America Needs a Bigger House." *Dawn* .
**URL:** *https://www.nytimes.com/interactive/2018/11/09/opinion/expanded-house-representatives-size.html*

Newberg, Paula. 1995. *Judging the State: Courts and Constitutional Politics in Pakistan*. Cambridge University Press.

O'Connell, Stephen D. 2020. "Can quotas increase the supply of candidates for higher-level positions? Evidence from local government in India." *Review of Economics and Statistics* 102(1):65–78.

Ord, Keith. 1975. "Estimation methods for models of spatial interaction." *Journal of the American Statistical Association* 70(349):120–126.

Pareto, Vilfredo. 1991. *The rise and fall of the elites: an application of theoretical sociology*. Transaction Publishers.

Prillaman, Soledad A. 2018. "Strength in Numbers: How Women's Groups Close India's Political Gender Gap." Working paper.

Querubin, Pablo et al. 2016. "Family and Politics: Dynastic Persistence in the Philippines." *Quarterly Journal of Political Science* 11(2):151–181.

Rehman, I. A. 2006. "Polls & madressah 'sanads'." *Dawn* .
**URL:** *https://web.archive.org/web/20150319043920/https://www.dawn.com/news/1069081*

Rossi, Martín A. 2017. "Self-Perpetuation of Political Power." *The Economic Journal* 127:455–473.
**URL:** *https://doi.org/10.1111/ecoj.12443*

Sajjad, Fizzah, Kate Vyborny, Ghulam Abbas Anjum and Erica Field. 2017. "Overcoming Barriers to Women's Mobility: Improving Women's Access to Public Transport in Pakistan." *CDPR Policy Brief* (PB1705).

Schickler, Eric, Kathryn Pearson and Brian D Feinstein. 2010. "Congressional parties and civil rights politics from 1933 to 1972." *The Journal of Politics* 72(3):672–689.

Schroeder, Christine M and Deborah A Prentice. 1998. "Exposing Pluralistic Ignorance to Reduce Alcohol Use Among College Students 1." *Journal of Applied Social Psychology* 28(23):2150–2180.

Shugart, Matthew Soberg and Rein Taagepera. 1989. *Seats and votes*. New Haven.

Smets, Kaat and Carolien Van Ham. 2013. "The embarrassment of riches? A meta-analysis of individual-level research on voter turnout." *Electoral studies* 32(2):344–359.

Smith, Daniel M. 2018. *Dynasties and Democracy: The Inherited Incumbency Advantage in Japan*. Stanford University Press.

Smith, Daniel M. and Steven R. Reed. 2018. "The Reed-Smith Japanese House of Representatives Elections Dataset.".
**URL:** *https://doi.org/10.7910/DVN/QFEPXD*

Sonnet, Luke. 2017. Private interview with Akbar and Yusuf Ayub Khan.

Sonnet, Luke. 2020. "The Dynastic Advantage in Close Elections in Four Countries: Implications for Regression Discontinuity Designs." Working paper.

Szakonyi, David. 2018. "Businesspeople in elected office: Identifying private benefits from firm-level returns." *American Political Science Review* 112(2):322–338.

Tankard, Margaret E and Elizabeth Levy Paluck. 2016. "Norm perception as a vehicle for social change." *Social Issues and Policy Review* 10(1):181–211.

Van Coppenolle, Brenda. 2017. "Political dynasties in the UK House of Commons: the null effect of narrow electoral selection." *Legislative Studies Quarterly* 42(3):449–475.

Vavreck, Lynn et al. 2007. "The exaggerated effects of advertising on turnout: The dangers of self-reports." *Quarterly Journal of Political Science* 2(4):325–343.

Waseem, Mohammad. 1989. *Politics and the State in Pakistan.* Progressive Publishers.

Waseem, Mohammad. 2002. "Issue of electoral reforms." *Dawn* .
**URL:** *https://web.archive.org/web/20200526204649/https://www.dawn.com/news/1062720*

Waseem, Mohammad. 2006. *Democratization in Pakistan: A study of the 2002 elections.* Oxford University Press.

Wilder, Andrew. 1999. *The Pakistani voter: Electoral politics and voting behaviour in the Punjab.* Oxford University Press.