Gender and Natural Experiments in Developing Countries[*]

Irma Clots-Figueras
University of Kent

Lakshmi Iyer
University of Notre Dame

Chapter for *Handbook of Experimental Development Economics*, eds. Utteeyo Dasgupta and Pushkar Maitra, under contract with Edward Elgar.

1. Introduction

This chapter summarizes research on the effects of women's leadership and gender-based policies in developing countries. Women are disadvantaged on many fronts globally, and particularly so in developing countries. Women's wages are 77% of men's wages globally, and only 26% of the world's parliaments are comprised of women. In developing countries, 57% of women have completed secondary school compared to 64% of men, and only 44% of women participate in the labor force compared to 73% of men.[1] Women are also subject to gendered issues such as high rates of maternal mortality and exposure to physical and sexual violence. In 2017, there were 211 maternal deaths for every 100,000 live births and almost one-third of women worldwide were estimated to have experienced sexual violence.[2]

Given this backdrop, many countries have established a range of policies targeted towards women. These run a gamut from quotas in political office and other employment, equalizing opportunities for education across men and women, directing specific monetary transfers to women recipients and targeting microfinance schemes to women borrowers. In this chapter, we examine the effects of these types of gender-based policies on overall development outcomes and those of women specifically.

[*] Email: i.clots-figueras@kent.ac.uk; liyer@nd.edu. We thank Nicolas Chehade, Regina Mannino and Eduardo Pages Jimenez for research assistance, and an anonymous referee for helpful comments.
[1] Data from the UN's Gender Inequality Index, UN Women and the Inter-Parliamentary Union.
[2] Data from the UN Population Fund and the World Health Organization, on behalf of the United Nations Inter-Agency Working Group on Violence Against Women Estimation and Data (2021).

In keeping with the theme of the Handbook, we examine primarily the circumstances under which gender-based interventions can be analyzed in an experimental, or quasi-experimental, framework. These include cases where such policies are implemented as randomized controlled trials or occur under circumstances that would approximate a random assignment. We also include studies where specific political or other situations result in quasi-experimental gender variation, such as policies or political events that increased women's political representation. We do not include natural experiments or policy initiatives that are not targeted at gender, even though they may have gendered effects. Examples of these include economic shocks, war, male out-migration, trade policies etc. Since the focus of the chapter is on natural experiments, we exclude experimental interventions introduced by researchers rather than by governments or other policy actors.

The rest of the chapter is structured as follows: Sections 2 and 3 examines the impact of political gender quotas, which have been implemented by more than 100 countries around the world. We discuss the typical methods used to analyze their effects, which are based either on randomized assignment (i.e. a comparison of treatment and control groups) or difference-in-differences (DiD) approach using staggered implementation of such policies. As we discuss each method, we discuss the underlying assumptions needed for identification of treatment effects and describe the common methods of testing for these. Section 4 examines natural experiments generated by the political setting, where many papers use a regression discontinuity design (RDD) to examine the effects of women's leadership. Section 5 discusses other gender-based policies that have been analyzed using DiD and RDD techniques. Section 6 concludes with a brief summary and directions for future research.

2. Political Gender Quotas

2.1. Political gender quotas around the world

According to the aggregated database of the International Institute for Democracy and Electoral Assistance (IDEA), more than 100 countries have some type of gender quota in their political systems. These are all aimed at increasing women's representation in political office and can be of three broad varieties: (a) reserved seats that specify a

percentage of seats in the legislature that must be filled by women, (b) candidate quotas that specify a certain percentage of electoral candidates to be women, or (c) political party quotas. Quotas of type (a) and (b) are typically mandated via constitutional amendments or specific legislation, while type (c) is usually voluntarily chosen by political parties. Developing countries with reserved seats (quotas of type (a)) include Egypt, Haiti, Burundi and local governments in India. Developing countries with candidate quotas (b) include Indonesia, Colombia, Bolivia and Paraguay. Countries have voluntary party quotas for women include El Salvador, Kenya and Mali. Currently, 27 countries have some type of reserved seats in the single or lower house of the national parliament, 61 countries have candidate quotas, and 23 countries possess voluntary political quotas.[3] Relatively few countries (22) have gender quotas for the upper chamber. At the sub-national level, 50 countries have mandated candidate quotas and 27 have reserved seat quotas.

It may be useful to consider here a few mechanisms by which gender quotas may affect overall development or women's outcomes. The implementation of gender quotas can lead to an increase in the number of women holding political office: as of 2015, countries that had gender quotas of types (a) or (b) had, on average, 24.7 percent women's legislative representation, whereas those without quotas had an average of 18.4 percent, a statistically significant difference (Hughes et al., 2019). Such numerical or descriptive representation may or may not translate into "substantive representation" i.e. incorporation of women's interests into policy decisions. This would depend on whether women indeed have different preferences from men, whether women in elected office share these different preferences, and whether the political system allows individual legislators or leaders some leeway to exercise their preferences in policy outcomes.[4] Finally, gender quotas may change the "symbolic representation" of women, namely how women are viewed by society as a whole. Gender quotas may result in altered social views about the gendered aspects of public life, may result in updated beliefs about women leaders' competence, or

[3] Data from "Gender Quotas Database," International Institute for Democracy and Electoral Assistance, https://www.idea.int/data-tools/data/gender-quotas/database, accessed March 2022.

[4] For instance, the basic two-candidate model of Downsian competition predicts that the policy chosen will be that preferred by the median voter, regardless of the preferences of the candidates (Downs, 1957). Other models allow candidate preferences to have a role in policy choices (Osborne & Slivinsky, 1996; Besley & Coate, 1997).

may result in backlash or resentment against the implementation of a quota (Franceschet et al., 2012).

Below we describe some of the research that has analyzed the substantive and symbolic effects of gender quotas in developing countries. The experimental analysis typically arises from features of the quota's implementation in specific countries. Most papers use either a randomized assignment of quotas across different places or a quasi-random analysis using variation over time and space. We describe the typical analysis below and discuss the conditions under which quota implementation can be treated as a natural experiment.[5]

2.2 Analysis of gender quotas as a randomized experiment

A well-analyzed example of randomized assignment of gender quotas arises from India's implementation of a gender quota in local governments, which was mandated by a 1993 constitutional amendment. The amendment required all local councils to have 33% seats in each local council to be reserved for women, as well as reserving 33% of village council head positions for women. Importantly, in most states, the villages that were required to have women chairpersons were chosen randomly from a list of all villages. This enabled researchers to examine the effects of the quota by simply comparing outcomes in village councils that were required to have women chairs ("treatment" areas) to outcomes in village councils that were not required to have women chairs ("control" areas).

Chattopadhyay and Duflo (2004) collected unique data on the provision of several types of public goods (water facilities, road condition, irrigation facilities etc.) across 256 villages in the states of West Bengal and Rajasthan. Their hypothesis was that local councils headed by women would respond more to the preferences of women, which they document by examining what men and women had complained about to the local council in previous years. They find that women in both these states were more concerned than men about the quality of drinking water facilities, that women were more concerned than men about road quality in West Bengal but not in Rajasthan, and that there were no differences between men and women regarding irrigation concerns.

[5] Clayton (2021) surveys a wider range of evidence on the effects of gender quotas, including non-experimental studies and qualitative research.

In order to interpret the randomized assignment of women council heads as a natural experiment, the authors needed to verify two important conditions. The first is that this policy actually made a difference to the gender of the council head. This would be violated, for instance, if a large fraction of non-reserved villages also elected women heads (say, because quotas resulted in a big change in voters' views of women leaders in other locations). Chattopadhyay and Duflo (2004) verify that this does not happen: only 6.5% of unreserved villages in West Bengal and 1.7% in Rajasthan have women chairpersons. The second condition is that the villages selected for women council heads are ex-ante similar to those that were not, meaning that the randomized assignment resulted in "balanced" treatment and control groups. The authors show that reserved and unreserved villages are very similar on many characteristics prior to the implementation of the gender quota, including population, literacy rates and water provision (Table 1, panel A).

[Insert Table 1 here]

After verifying these conditions, Chattopadhyay and Duflo (2004) compare public goods provision across reserved and unreserved villages, as a measure of the impact of the gender quota. They find that reserved villages invest more in public goods that are more closely linked to the concerns of women (Table 1, panel B). For instance, reserved villages in West Bengal are found to have, on average, 24 newly built or repaired drinking water facilities, compared to only 15 in unreserved villages (columns 1 and 2). In Rajasthan, the corresponding figures are 7.3 for reserved villages and 4.7 for unreserved ones (columns 5 and 6). These differences are statistically significant at the 5% level of significance (columns 3 and 4 show the difference and associated standard error for West Bengal, and columns 7 and 8 for Rajasthan). This is consistent with the drinking water concerns expressed by women in both states. Further, reserved villages in West Bengal (where women expressed more concern over roads) invest significantly more in road maintenance compared to unreserved villages. Such a difference is not seen in Rajasthan, where women were not more concerned about roads than men. Similarly, neither state shows a difference in irrigation provision, an issue on which there is no gender-based preference difference.

While these are convincing causal estimates of the quota policy, it is less clear that these are attributable primarily to leader's gender rather than other correlates of being a woman leader in this environment e.g. women are typically less educated than men, have

lower financial resources, are more likely to be first-time office holders and are less likely to want to run for re-election. Chattopadhyay and Duflo's detailed survey allows them to control for these characteristics, with the effect of the quota remaining unchanged, suggesting that it is indeed the leader's gender that matters. Another issue they devote attention to is establishing whether these changed public goods results are a result of women's different preferences rather than women's different leadership styles (perhaps women leaders are more responsive to citizen complaints). Their survey enables them to track responsiveness to both male and female citizen complaints, and their results strongly indicate that it is not responsiveness in general but women-specific preferences that drive the results.

How far do these results generalize? The evidence from different parts of India is mixed in this regard. Deininger et al. (2020) find a positive effect of gender quotas on women's participation on a national public works program. Duflo and Topalova (2004) use a nationwide survey of public goods and citizen satisfaction across 2,304 randomly selected villages in 24 states. When comparing villages where the chairperson position was reserved for women with villages where it was not reserved, they find higher levels of public goods in the women-reserved villages (particularly drinking water and education facilities), and citizens report paying fewer bribes in women-reserved villages. Despite these positive aspects, citizens are less satisfied with their local governments when there is a woman chairperson. The authors do not distinguish the mechanisms behind this result but propose alternatives such as the fact that maybe women's performance is worse in unobservable aspects, or the fact that women politicians also have some characteristics that make them more likely to be judged unfavorably. It is also possible that pre-existing social norms may make citizens averse to having women leaders ("taste-based" discrimination), or may react negatively to the fact of having a mandated quota.

Evidence from other studies have highlighted some negative effects of the local government gender quota. Gajwani and Zhang (2015) analyze data from the state of Tamil Nadu and find that fewer roads and schools are built in villages with women chairpersons; they also document that women leaders are less knowledgeable about the political system and have fewer interactions with upper-level officials. In a similar vein, Bardhan et al. (2010) use data from 16 districts of West Bengal to find that women's reservation has no

impact on drinking water or roads and results in worse targeting of benefits to needy households. Ban and Rao (2008) examine data from four states and find that women leaders provide more education facilities, a dimension on which there is no gender difference in expressed preferences; there is, however, no difference in their provision of water and sanitation facilities that are strongly preferred by women. Rajaraman and Gupta (2012) survey four other states and find women's reservation increases building construction, but not spending on water or sanitation. Deininger et al. (2014) use a nationwide sample to document lower quality of water, sanitation and education provision in women-reserved areas, but an increased willingness to contribute to public goods. Results from all these papers cannot be compared to Chattopadhyay and Duflo (2004) due to differences in the samples used. In other words, India is a very large country with large variations across regions, which means that these widely differing results are likely due to differences in the social context (how educated or empowered women are in the local areas), differences in the expressed preferences of women relative to men, or political factors that may give lower or higher agency to women chairpersons over and above women councilors. A meta-analysis of the effect of gender quotas would thus be needed in order to claim external validity.

2.3. Dynamic effects of gender quotas

A different concern with such "randomized experiment" designs is that they cannot reveal the dynamic effects over time, which includes women leaders' learning progress, changes to women's political participation, the changing views of voters, changes in the types of men and women elected to political office, and changes to social norms about women's roles in public life.[7] All of these will shape whether the effects of gender quotas become cumulative or attenuate over time. A few recent papers have begun to shed light on these dynamics.

[7] Political parties in Spain, for instance, placed women in areas where they were unlikely to win, thereby undermining the effectiveness of candidate quotas (Casas-Arce and Saiz, 2015). Party-specific quotas in developed countries have been shown to increase the average quality of male candidates (Besley et al., 2017; Baltrunaite et al., 2014).

Afridi et al. (2017) track the time path of women leaders' effectiveness using a sample of villages from Andhra Pradesh state. They find that women leaders' inexperience leads to higher citizen dissatisfaction and more corruption experienced at the beginning of their terms, but that there is no difference between women and men at the end of the five-year terms. Chaudhuri et al. (2022) similarly find that experienced female politicians are as likely to display dishonest behavior as male politicians in an experimental game, but that inexperienced women are less likely to be dishonest. These results suggest that cross-sectional correlations showing lower corruption under women leaders (Dollar, Fisman & Gatti, 2001; Swamy et al., 2001; Brollo & Troiano, 2016) may not last over the longer term as women politicians become more experienced.

The results are mixed when it comes to the effect of gender quotas on women's political participation. Chattopadhyay and Duflo (2004) find a higher fraction of women attending village council meetings when there is a woman chairperson in West Bengal, but not in Rajasthan (Table 1, panel C). On a different measure of participation, there is no statistically significant difference in the frequency of complaints by women or men. Many of the other studies cited in Section 2.2 find a higher participation of women in village meetings when there is a woman chairperson, but Iyer and Mani (2019) emphasize that such increases are able to close less than half of the gender gap in political participation.

Turning to electoral measures of political participation, Beaman et al. (2009) find that women's political candidacy increases in villages that were reserved for women in two consecutive terms, but not after a single term. They also document a decline in measured statistical discrimination amongst voters, but not in taste-based discrimination. However, analysis of a randomized gender quota in Lesotho finds an increase in women's electoral success even after the quota is withdrawn, but this is mostly attributable to the exit of male incumbents (Clayton & Tang, 2018). In fact, having a quota-mandated female representative reduces several dimensions of women's self-reported engagement with local politics (Clayton, 2015). Bhavnani (2009) use data from Mumbai to find an increase in candidacy after just one term of women's reservation, but re-analysis by Sekhon and Titiunik (2012) find much smaller effects. The latter paper points to an important statistical consideration, namely that gender quotas in a nearby area can have spillover effects (e.g.

women candidates often switch areas to contest from a reserved one rather than an unreserved one), so it may not be correct to treat these units as "independent" of each other.

Such spillover effects of quotas can potentially extend to the longer term and to other groups. The evidence base on such effects is small and inconclusive. Some papers have examined the effects of gender quotas on the aspirations and outcomes of the next generation. Beaman et al. (2012) found positive effects of quotas on educational and political aspirations of adolescent girls after a long (ten-year) exposure to women leaders. This is similar to cross-sectional evidence on greater political interest and engagement in countries with more women legislators (Wolbrecht and Campbell, 2007), but a DiD analysis of corporate board quotas in Norway finds only "mixed support" for the view that the reform affected the decisions of younger women (Bertrand et al., 2019).

A common concern with quotas or other affirmative action policies is whether prioritizing gender representation would lead to lower diversity on other dimensions. In the case of India, if the female representatives are likely to be from higher castes, then caste diversity would suffer due to the introduction of the gender quota. Two studies from India use the time-varying nature of seat reservation to find opposite results on this question: Karekurve-Ramachandra and Lee (2020) find more high-caste representation in gender-reserved seats, while Cassan and Vandewalle (2021) find more low-caste representation in such seats.[8]

3. Using Panel Data Analysis to Analyze Gender Quotas

A different empirical strategy to examine the effect of quotas or other gender-based policies relies on staggered implementation of such policies over time and space. Iyer et al. (2012) provide an example of this methodology. They examine India's gender quota, not by comparing places with or without a woman leader, but by examining outcomes in states before and after the implementation of the gender quota. Note that this captures the effect of the quota at all levels, namely in village, intermediate, and district councils and chairperson positions. In contrast, the papers described earlier capture only the effect of a

[8] Cross-sectional studies find that gender quotas improve the diversity of elected representatives along other dimensions (Barnes & Holman, 2020).

woman chairperson, over and above the presence of women as village councilors; those papers also do not take into account the role that may be played by quotas at the upper levels. Iyer et al. (2012) examine crimes against women as the primary outcome of interest, using a two-way fixed-effects regression specification:

$$(1) \qquad Y_{st} = a_s + b_t + gGenderQuota_{st} + X_{st}'\mathbf{d} + e_{st}$$

In equation (1), Y_{st} measures crimes against women in state s and year t, and the main explanatory variable is $GenderQuota_{st}$, a dummy that equals one if state s has implemented the gender quota in local elections on or before year t. a_s are fixed effects for each state (i.e. a dummy for each state), which essentially control for any time-invariant state characteristics such as geography, history or long-standing gender norms. b_t are fixed effects for each year, which control for any nationwide changes in year t that affect all states. The coefficient g is therefore identified purely by using the fact that different states implemented the gender quota in different years, so that each state can be compared to itself before and after such implementation, using the changes over the same period in non-quota states as a proxy for what the changes would have been in state s in the absence of the quota.

The staggered implementation of a policy across different areas at different times can be interpreted as a quasi-experimental variation under certain conditions. First, since the comparison is outcome changes in quota-adopting areas to changes in non-adopting areas, it should not be the case that quota-adopting areas were already experiencing faster or slower changes in outcomes even prior to the quota. This is often called the "no differential pre-trends" assumption. The typical way to test this is to conduct a difference-in-difference analysis for several years restricting the sample to a period before the policy was implemented, or conducting an event-study type of analysis and showing the existence of effects only for the period after the policy. If the policy was indeed a "natural experiment," then we should see no significant differences prior to policy adoption in the states that would subsequently adopt the quota.

Second, the estimates may be biased if the adoption of gender quotas coincided with the adoption of other policies that also affected crimes against women. To investigate

this, researchers need to understand the policy context and the reasons behind the variation in timing across states e.g. Iyer et al. (2012) cite the pre-existing schedule of local elections, lawsuits filed on the basis of caste and budgetary concerns as reasons for staggered implementation. Researchers also need to control for other state-and time-varying covariates that might simultaneously affect Y_{st} and $GenderQuota_{st}$. This is denoted as X_{st} in equation (1). Iyer et al. (2012) control for economic variables (per capita state GDP, fraction working in farming, urbanization), indicators of women's social and economic progress (the female share of the population, women's literacy rates), and women's representation in state institutions (the fraction of women in the policy, whether the state has a woman Chief Minister).

Iyer et al. (2012) find that the introduction of gender quotas results in a significant increase in reported crimes against women, but not in gender-neutral crimes (like burglary or arson) or in crimes where the victims are male. Together with supporting survey data, they attribute this effect primarily to a greater propensity of women victims to come forward in the post-quota period, and a greater probability of police action (e.g. arrests) on their behalf. This effect is particularly interesting, given that the women elected via local gender quotas do not have direct jurisdiction over the police.

Finally, the two-way fixed-effects estimator may be biased in the presence of heterogeneous treatment effects. This estimator compares the change in outcomes for a state that implements the quota in year *t* to the change in outcomes for states do not implement the quota at time t. The latter category includes both states that implemented the quota after time t (i.e. "not-yet-treated" at time t), and also states that implemented the quota before time t ("already treated"). If early quota adopters have a larger effect (say) than later adopters, then using these "already treated" states as de facto controls for the late adopters may result in estimates that are biased and even of the opposite sign to the actual treatment estimates. Several recent papers have discussed methods to both quantify the extent of such bias and to construct alternative estimators that would be robust to such biases (de Chaisemartin & D'Haultfoeuille, 2020; Goodman-Bacon, 2021; Callaway & Sant'Anna, 2021; Sun & Abraham, 2021). Since these techniques have been developed very recently, many of the DiD papers surveyed here do not perform these tests.

Such two-way fixed-effects or difference-in-difference analyses have been used to examine the effects of gender quotas in a cross-country setting. The adoption of a gender quota has been documented to result in a substantial increase in health spending (Clayton & Zetterberg, 2018) and a reduction in maternal mortality (Bhalotra et al., 2022). Examining variation of quota adoption in Mexican states, where political parties were limited in the proportion of candidates of the same gender they could have in their lists and where quotas were not always effectively implemented, Zetterberg (2012) finds that female leadership does not influence women's self-reported interest in politics. Nor does he find any systematic impact on women's political engagement across 17 Latin American countries with varying quota experiences (Zetterberg, 2009). Studies using state-level variation in the timing of India's local gender quota have shown an increase in women's political candidacy at the (higher) state and national levels (Karekurve-Ramachandra, 2021; Maitra & Rosenblum, 2022). Some studies have used the randomized reservation of women at the district council level to examine the effects on sex ratios at birth (Kalsi, 2017), and women's candidacy at the state level (O'Connell, 2020).

4. Quasi-Random Assignment of Women Leaders

4.1. Close elections in a first-past-the-post system

As discussed earlier, while gender quotas have increased female political representation, they can also affect the nature of political competition, by affecting the way candidates, voters, and political parties behave. Thus, estimating the effects of quotas on policy includes the effect of not just women policymakers, but also these other changes. Some features of political systems can be used to identify purely the effects of changing the leader's identity because of quasi-random variation generated by the existing system without any new interventions. A widely-used feature in empirical research is the as-if-random nature of close elections in a first-past-the-post electoral system. In first-past-the-post (FPTP) systems, common in the U.S., U.K. and India, the winner must obtain a higher number of votes than the runner-up to win. When this vote difference is very small, it means that both the winner and runner-up had an almost equal probability of winning. Under some circumstances, who actually wins can be treated as a natural experiment.

A Regression Discontinuity Design (RDD) is the most common way to analyze such quasi-experimental variation. A typical example is estimating the causal effect of a woman's presence in political office on an outcome Y.[9] We describe the implementation of RDD in Bhalotra et al. (2018), who examine whether a woman's electoral victory results in greater female candidacy in the next election.

4.2. Identifying assumptions in RDD

To implement the RDD, Bhalotra et al. (2018) compute the "vote margin" i.e. the difference in the fraction of votes obtained by the female and the male candidates. This variable X is known as the *forcing, assignment or running variable.* While this variable is continuous (can range between -1 and +1), it has discontinuous consequences at the threshold of zero: when the vote margin exceeds zero, the woman candidate wins the election and when less than zero, the male candidate wins. The RDD compares outcomes across places where women narrowly win (vote margin>0) to outcomes in places where women narrowly lose (vote margin<0), after controlling for the direct effects of the running variable. The regression specification is:

$$(2) \qquad Y_i = a + g AboveThreshold_i + f(RunningVariable_i) + e_i$$

In equation (2), Y_i represents the outcome (e.g. future female candidacy) in place i, $AboveThreshold_i$ is a dummy that takes value one when the running variable (vote margin) is above the threshold (zero) and f is a continuous function of the running variable, which captures the possibility that places where women win by large margins are different from places where they lose by large margins.

When can the estimate g be interpreted similarly to a coefficient estimated in an experimental setting? [10] First, we need to verify that being above the threshold changes the electoral outcome. This is obvious in the case of elections, but may be violated in other cases (e.g. people who are above an eligibility threshold for government benefits may

[9] Other policies that vary sharply at a threshold can also be used in RDDs. Examples include being eligible for government benefits below a specific income threshold or cities above a certain size acquiring electronic voting machines (Fujiwara, 2015).

[10] For a more formal and detailed explanation please check https://mixtape.scunning.com.

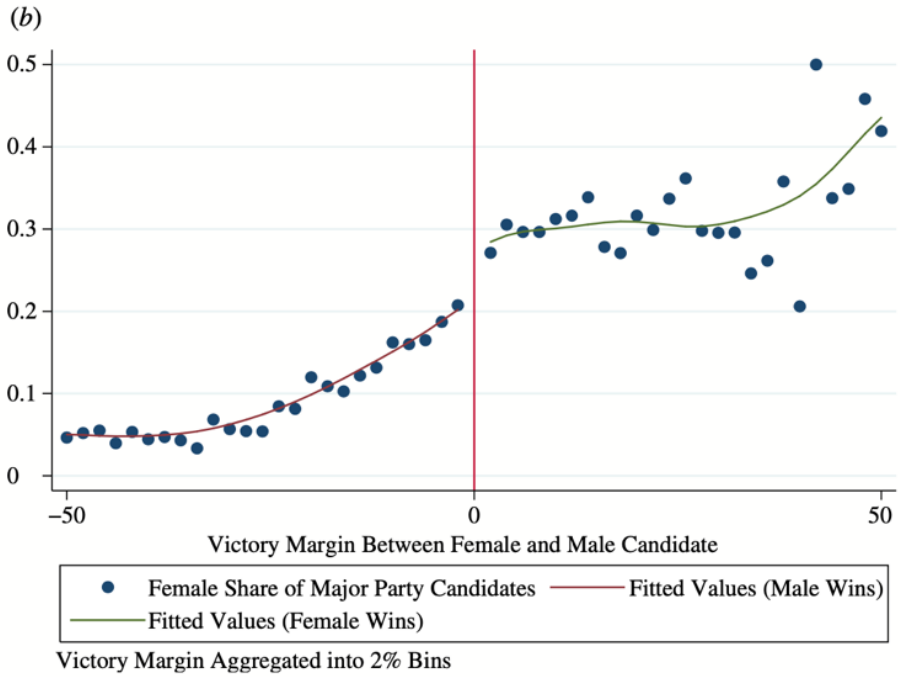
nevertheless be able to access them). Second, as we get closer to the discontinuity, we need to show that the probability of being above or below the threshold should be quasi-random. This means that male or female candidates should not have a differential advantage in winning close elections. This is often tested by examining whether the density of the running variable is continuous at the threshold (see McCrary (2008) for a formal test). Pre-determined covariates and prior outcomes should not vary discontinuously at the same threshold; if they do, the RDD estimate may be subject to omitted variables bias. Bhalotra et al. (2018) show that places where women won narrowly are similar along many dimensions to places where women lost narrowly: on demographic characteristics (population, urbanization, literacy etc.), on political histories (female candidates in the previous election, voter turnout), and on other characteristics of the candidate pool (education, wealth, party identity). The latter is very important if female and male candidates systematically differ on those characteristics.

After establishing these validity conditions, Bhalotra et al. (2018) run a RDD regression as in equation (2). They find, after controlling for a quadratic relationship with the vote margin, that the fraction of female major party candidates increases by 8.5 percentage points in places that saw a narrow victory by a woman over a man. Importantly, they find that the entirety of the effect is attributable to the winning women being more likely to contest for re-election. In fact, the share of major party candidates who are "new women" (i.e. those who did not run before) decreases by a statistically significant 2.3 percentage points, suggesting that any demonstration or role model effects of women's electoral victories do not manifest in the short run.[11] Their results are also robust to running local linear regressions around a very small bandwidth near the discontinuity.

A typical way to present the results of RDD is a graph like Figure 1 below. The x-axis shows the running variable, in this case the vote share difference between female and male candidates. The y-axis shows the outcome variable, here the fraction of female major party candidates. The vertical line represents the threshold, where a small change in the running variable results in a large change in the "treatment". We can see a clear jump or discontinuity in female candidacy at the threshold, which represents the RDD estimate.

[11] A similar RDD analysis of U.S. data also finds no entry of women candidates in nearby areas after a woman's electoral victory (Broockman, 2013).

Figure 1: Female victory and candidacy. Source: Bhalotra et al. 2018.



Using a similar RDD on Indian data, Baskaran et al. (2021) find that the presence of women legislators results in significantly higher economic growth, as measured by the growth of night lights. Brollo and Troiano (2016) find that narrowly elected female mayors in Brazil improve education and infant mortality in their cities and are less likely to be corrupt. Despite these positive aspects, they are significantly less likely to win re-election compared to narrowly elected male mayors.[12]

4.3. Using RDD as an IV

In some cases, the outcome variable is measured at a different level of aggregation than the RDD running variable. For example, in the case of India, development outcomes

[12] In contrast, Ferreira and Gyourko (2014) find that women mayors in the U.S. have a higher re-election probability than men, despite their presence not leading to any changes in municipal spending or crime rates. Hessami and Lopes da Fonseca (2020) review the evidence on the effects of female leadership in developed countries.

are often measured at the level of the administrative district, which has several elected state legislators. In this case, we cannot use the usual RDD, since the outcome cannot be matched to the presence or absence of a specific woman leader. However, we can make an RDD-based instrumental variables strategy as follows: regress the outcome to the aggregate share of elected women, and instrument this aggregate share with the fraction of women who are elected in close elections against a man. What we need for identification is a strong correlation between the fraction of women and the fraction of women politicians winning in close elections against men (akin to the usual IV "first stage" requirement). Since this instrument is constructed based on the existence of close mixed-gender elections, the exclusion restriction may be violated if the occurrence of close elections is correlated with the outcome variable. It is therefore important to control for the fraction of close mixed-gender elections when running this IV version of the RDD specification.

Formally, the specifications are as follows, where equation (3) is the second stage and equation (4) is the first stage:

(3) $Y_{idt} = \theta_d + \psi_t + \beta\, WL_{dt} + \lambda TC_{d,t} + X_{idt}\,\eta + \varepsilon_{idt}$

(4) $WL_{ds,t} = \pi_d + \rho_t + \kappa WC_{d,t-1} + \mu TC_{d,t-1} + X_{idt}\,\sigma + u_{d,t-1}$

where Y_{idt} is the outcome of individual i in district d and year t, and WL_{dt} is the fraction of constituencies in district d that elected women. The latter is instrumented with the share of constituencies in the district won by women in close elections against men in the same time period, WC_{dt}. The fraction of constituencies in the district in which there were close elections between women and men, TC_{dt}, is controlled for in the second stage (equation 3) and partialled out of the instrument in the first stage (equation 4). The coefficient of interest is β, which identifies the impact of female relative to male legislators. Other controls such as district and time fixed effects, polynomials in the margin of victory in each mixed-gender election, or individual characteristics, can be included to reduce standard errors, but are not essential to the identification.

Clots-Figueras (2012) uses this fuzzy RDD to find that the election of a female politician increases the probability that individuals living in urban areas attain primary education; however, the effects are not present in rural areas. She shows that her estimates are consistent with running RDD regressions in those districts that had only one mixed-gender election. Bhalotra and Clots-Figueras (2014) use a similar strategy to find that the

presence of women legislators leads to lower infant mortality; Anukriti et al. (2022) find that women legislators' presence results in divergent fertility preferences across men and women, and manifests in higher levels of domestic violence.

5. Quasi-Experimental Analyses of Other Gender-Based Policy Interventions

Several non-quota policies are amenable to a DiD analysis, provided that researchers can identify individuals or areas that were affected by the policies at different times. In the case of individuals, researchers usually examine cohorts that were exposed to specific policies and compare them to different cohorts (in the same place) that were not exposed to those policies. For instance, several states in India amended inheritance laws to give women a greater share of the family property; women whose fathers died prior to the passage of such laws form the unaffected cohorts, whose outcomes can be compared to women whose inheritances were subject to the new laws. Papers analyzing the impact of these laws find contrasting effects. Deininger et al. (2013) find that women's age of marriage and level of schooling rose after their enactment. However, Bhalotra et al. (2020) find that they led to an increase in female foeticide and neonatal mortality, while Anderson and Genicot (2015) find an increase in female and male suicide rates after these laws. La Ferrara and Milazzo (2017) also find the reverse result in Ghana, namely that education levels of boys in matrilineal communities fell after they became eligible for a higher level of inheritance.

Muralidharan and Prakash (2017) compare the outcomes of a bicycle grant program directed at secondary school girls in the Indian state of Bihar. Their triple-difference strategy compares the outcomes of affected cohorts of girls to those of unaffected cohorts, using boys and the neighboring state of Jharkhand as comparison groups. They find a large and significant increase in secondary school enrolment and completion for girls who benefited from this program, especially those who lived at an intermediate distance (5-13 km) from the schools. A similar program in Zambia reduced the number of school days missed by girls and improved their perceptions of being in control of their lives, but had no effect on school or grade progression (Fiala et al., 2022). Anukriti (2018) analyzes the effects of financial incentives that were aimed at reducing fertility and improving the

gender ratio at birth, and finds that the gender ratio worsened as fertility declined.[13] Amaral et al. (2021) find a substantial increase in the reporting of violence against women after women-only police stations are opened.

While RDD has been mainly used in the context of elections, it can also be used to analyze policies where there is a clear cutoff that determines eligibility for a particular policy. For example, Erten and Keskin (2018) use an RDD to estimate the effects of a compulsory education reform in Turkey that increased girls' education on domestic violence, and Bergolo and Galvan (2018) examine the effects of a cash transfer program targeted to women in Uruguay on women's employment and household decision-making.

6. Conclusions

Most findings on experimental and quasi-experimental impacts of gender-based policies do not generalize across settings, even within the same country. This is clearly the case for studies analyzing the random allocation of quotas for female politicians in local governments in India. This policy was implemented in a randomized setting that has been proven to be equivalent to an experimental approach, but this means it will suffer from many problems that arise in randomized controlled trials, which ensure internal validity, but may have large external validity and generalizability costs. Experimental approaches may also not good at tracking spillover effects across units of observation. This is the case because independence is often an assumption. Cluster-randomization can be one way to address this in experimental settings, but it may not be easily applied by policymakers. However, non-experimental studies are likely to suffer from worse problems, especially related to the existence of omitted variable bias and reverse causality.

The next stage of research should provide a more theory-driven analysis of gender-based policies that takes into account differences across contexts and can predict in which settings they are more likely to be effective. For example, a meta-analysis of India's gender quotas would be a clear opportunity to learn about this. Many public policies are not implemented in ways to allow for experimental or quasi-experimental impact evaluation; partnerships between academics and policymakers can be one way to increase our

[13] The link between reduced fertility and a male-biased gender ratio has also been explored in Ebenstein (2010) and Jayachandran (2017).

knowledge base of the effects of gender-based policies. Yet another big gap in the current knowledge set is due to the fact that most gender-based policies are only targeted towards women, while changing women's outcomes and particularly social norms may require some targeting towards men.

References

Afridi, F., Iversen, V., & Sharan, M.R. (2017). Women Political Leaders, Corruption and Learning: Evidence from a Large Public Program in India. *Economic Development and Cultural Change, 66*(1), 1-30.

Amaral, S., Bhalotra, S., & Prakash, N. (2021). *Gender, Crime and Punishment: Evidence from Women Police Stations in India*. CESifo Working Paper No. 9002.

Anderson, S., & Genicot, G. (2015). Suicide and property rights in India. *Journal of Development Economics, 114,* 64-78.

Anukriti, S. (2018). Financial Incentives and the Fertility-Sex Ratio Trade-Off. *American Economic Journal: Applied Economics*, *10*(2), 27-57.

Anukriti, S., Erten, B., Mukherjee, P. (2022). *Women's Political Representation and Intimate Partner Violence*. (Policy Research Working Paper 10113). Washington, DC : World Bank.

Baltrunaite, A., Bello, P., Casarico, A., & Profeta, P. (2014). Gender Quotas and the Quality of Politicians. *Journal of Public Economics, 118*, 62-74.

Ban, R., & Rao, V. (2008). Tokenism or Agency? The Impact of Women's Reservations on Village Democracies in South India. *Economic Development and Cultural Change*, *56*(3), 501-530.

Bardhan, P., Mookherjee, D., & Torrado, M. (2010). Impact of Political Reservations in West Bengal Local Governments on Anti-Poverty Targeting. *Journal of Globalization and Development, 1*(1), 1-38.

Barnes, T., & Holman, M. (2020). Gender Quotas, Women's Representation, and Legislative Diversity. *Journal of Politics*, *82*(4): 1271-1286.

Baskaran, T., Bhalotra, S., Min B. & Uppal, Y. (2021). *Women legislators and economic performance* (Warwick Economics Research Paper Series No. 1354). University of Warwick, Department of Economics.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., & Topalova, P. (2009). Powerful Women: Does Exposure Reduce Bias? *Quarterly Journal of Economics, 124*(4), 1497-1540.

Beaman, L., Duflo, E., Pande, R., & Topalova, P. (2010). Political Reservation and Substantive Representation: Evidence from Indian Village Councils. In S. Bery, B. Bosworth, & A. Panagariya (Eds.), *India Policy Forum 2010/11 Vol 7* (pp. 159-191). Brookings Institution Pres and NCAER.

Beaman, L., Duflo, E., Pande, R., & Topalova, P. (2012). Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India. *Science*, *335*(6068), 582-586.

Bergolo, M., & Galván, E. (2018). Intra-household Behavioral Responses to Cash Transfer Programs. Evidence from a Regression Discontinuity Design. *World Development*, *103*, 100–118.

Bertrand, M., Black, S., Jensen, S., & Lleras-Muney. (2019). A. Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway. *The Review of Economic Studies*, *86*(1), 191–239.

Besley, T., & Coate, S. (1997). An Economic Model of Representative Democracy. *Quarterly Journal of Economics, 112*(1), 85-114.

Besley, T., Folke, O., Persson, T., & Rickne, J. (2017). Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden. *American Economic Review, 107*(8), 2204-2242.

Bhalotra, S., Brulé, R., & Sanchari, R. (2020). Women's Inheritance Rights Reform and the Preference for Sons in India. *Journal of Development Economics*, *146*.

Bhalotra, S., Clarke, D., Gomes, J., & Venkataramani, A. (2022). *Maternal Mortality and Women's Political Power* (NBER Working Paper No. 30103). National Bureau of Economic Research.

Bhalotra, S., & Clots-Figueras, I. (2014). Health and the Political Agency of Women. *American Economic Journal: Economic Policy, 6*(2), 164-197.

Bhalotra, S., Clots-Figueras, I., & Iyer, L. (2018). Pathbreakers? Women's Electoral Success and Future Political Participation. *The Economic Journal, 128*(613), 1844-1878.

Bhavnani, R. (2009). Do Electoral Quotas Work After They Are Withdrawn? Evidence from a Natural Experiment in India. *American Political Science Review, 103*(1), 23-35.

Brollo, F. & Troiano, U. (2016). What happens when a woman wins an election? Evidence from close races in Brazil. *Journal of Development Economics, 122,* 28-45.

Broockman, D.E. (2014). Do Female Politicians Empower Women to Vote or Run for Office? A Regression Discontinuity Approach. *Electoral Studies,* 34, 190-204.

Callaway, B., & Sant'Anna, P.H.C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, *225*(2), 200-230.

Casas-Arce, P., & Saiz, A. (2015). Women and Power: Unpopular, Unwilling, or Held Back? *Journal of Political Economy, 123*(3), 641-669.

Cassan, G., & Vandewalle, L. (2021). Identities and public policies: Unexpected effects of political reservations for women in India. *World Development, 143*.

Chattopadhyay, R., & Duflo, E. (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica 72*(5), 1409-1443.

Chaudhuri, A., Iversen, V., Jensenius, F.R. and Maitra, P. (2022). Time in Office and the Changing Gender Gap in Dishonesty: Evidence from Local Politics in India. *American Journal of Political Science.* https://doi.org/10.1111/ajps.12733.

Clayton, A. (2015). Women's Political Engagement under Quota-Mandated Female Representation: Evidence from a Randomized Policy Experiment. *Comparative Political Studies, 48*(3), 333-369.

Clayton, A. (2021). How Do Electoral Gender Quotas Affect Policy? *Annual Review of Political Science, 24*(1), 235-252.

Clayton, A., & Tang, B. (2018). How women's incumbency affects future elections: Evidence from a policy experiment in Lesotho. *World Development, 110*, 385-393.

Clayton, A., & Zetterberg, P. (2018). Quota Shocks: Electoral Gender Quotas and Government Spending Priorities Worldwide. *Journal of Politics, 80*(3), 916-932.

Clots-Figueras, I. (2012). Are Female Leaders Good for Education? Evidence from India. *American Economic Journal: Applied Economics, 4*(1), 212-244.

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review, 110*(9), 2964-2996.

Deininger, K., Goyal, A., & Nagarajan, H. (2013). Women's inheritance rights and intergenerational transmission of resources in India. *Journal of Human Resources, 48*(1), 114-141.

Deininger, K., Jin, S., Nagarajan, H., & Xia, F. (2014). Does Female Reservation Affect Long-Term Political Outcomes? Evidence from Rural India. *The Journal of Development Studies, 51*(1), 32-49.

Deininger, K., Nagarajan, H., & Singh, S. (2020). Women's political leadership and economic empowerment: Evidence from public works in India. *Journal of Comparative Economics, 48*(2), 277-291.

Dollar, D., Fisman R., & Gatti R. (2001). Are women really the 'fairer' sex? Corruption and women in government. *Journal of Economic Behavior and Organization, 46*(4), 423-429.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy, 65*(2), 135–150.

Duflo, E., & Topalova, P. (2004). Unappreciated Service: Performance, Perceptions, and Women Leaders in India. *Framed Field Experiments,* 1-22.

Ebenstein, A. (2010). The Missing Girls of China and the Unintended Consequences of the One Child Policy. *Journal of Human Resources, 45*(1), 87-115.

Erten, B., & Keskin, P. (2018). For Better or for Worse? Education and the Prevalence of Domestic Violence in Turkey. *American Economic Journal: Applied Economics, 10*(1), 64-105.

Ferreira, F., & Gyourko, J. (2014). Does Gender Matter for Political Leadership? The Case of U.S. Mayors. *Journal of Public Economics, 112*, 24-39.

Fiala, N., Garcia-Hernandez, A., Narula, K., & Prakash, N. (2022). *Wheels of Change: Transforming Girl's Lives with Bicycles* (CESifo Working Paper No. 9865). CESifo Munich.

Franceschet, S., Krook, M.L. & Piscopo, J. (2012). Conceptualizing the impact of gender quotas. In Franceschet, Susan, Mona Lena Krook and Jennifer M. Piscopo (eds.) *The impact of gender quotas*. Oxford: Oxford University Press.

Fujiwara, T. (2015). Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil. *Econometrica*, 83(2), 423-464.

Gajwani, K., & Zhang, X. (2015). Gender and Public Goods Provision in Tamil Nadu's Village Governments. *The World Bank Economic Review, 29*(2), 234–261.

Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics, 225*(2), 254-277.

Hessami, Z., & Lopes da Fonseca, M. (2020). Female Political Representation and Substantive Effects on Policies: A Literature Review. *European Journal of Political Economy, 63*.

Hughes, M., Paxton, P., Clayton, A., & Zetterberg, P. (2019). Global Gender Quota Adoption, Implementation, and Reform. *Journal of Comparative Politics*, *51*(2), 219-238.

Iyer, L. & Mani, A. (2019). The Road Not Taken: Gender Gaps Along Paths to Political Power. *World Development, 119*, 68-80.

Iyer, L., Mani, A., Mishra, P., & Topalova, P. (2012). The Power of Political Voice: Women's Political Representation and Crime in India. *American Economic Journal: Applied Economics*, *4*(4), 165-193.

Jayachandran, S. (2017). Fertility Decline and Missing Women. *American Economic Journal: Applied Economics*, *9*(1), 118-39.

Kalsi, P. (2017). Seeing Is Believing- Can Increasing the Number of Female Leaders Reduce Sex Selection in Rural India? *Journal of Development Economics*, *126*, 1–18.

Karekurve-Ramachandra, V. (2021). Gender Quotas and Upward Political Mobility in India. Quarterly Journal of Political Science, *forthcoming*.

Karekurve-Ramachandra, V., & Lee, A. (2020). Do gender quotas hurt less privileged groups? Evidence from India. *American Journal of Political Science*, *64*(4), 757-772.

La Ferrara, E., & Milazzo, A. (2017). Customary Norms, Inheritance, and Human Capital: Evidence from a Reform of the Matrilineal System in Ghana. *American Economic Journal: Applied Economics*, *9*(4), 166-185.

Maitra, P. & Rosenblum, D. (2022), Upstream effects of female political reservations, *European Journal of Political Economy*, *71*(102061).

McCrary, J. (2008). Manipulation of The Running Variable in The Regression Discontinuity Design: A Density Test. *Journal of Econometrics,* 142(2), 698-714.

Muralidharan, K., & Prakash, N. (2017). Cycling to School: Increasing Secondary School Enrollment for Girls in India. *American Economic Journal: Applied Economics*, *9*(3), 321-350.

O'Connell, S. (2020). Can quotas increase the supply of candidates for higher-level positions? Evidence from local government in India. *Review of Economics and Statistics*, *102*(2), 65-78.

Osborne, M.J., & Slivinski, A. (1996). A Model of Political Competition with Citizen-Candidates. *The Quarterly Journal of Economics*, *111*(1), 65–96.

Rajaraman, I., & Gupta, M. (2012). Public Expenditure Choices and Gender Quotas. *Indian Growth and Development Review*, *5*(2).

Sekhon, J., & Titiunik, R. (2012). When Natural Experiments are Neither Natural nor Experiments. *American Political Science Review, 106*(1), 35-57.

Sun, L., & Abraham, S. (2021). Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics*, *225*(2), 175-199.

Swamy, A., Azfar O., Knack S. & Lee Y. 2001. Gender and corruption. *Journal of Development Economics*, *64*(1), 25-55.

Wolbrecht, C., & Campbell, D. (2007). Leading by Example: Female Members of Parliament as Political Role Models. *American Journal of Political Science, 51*(4), 921-939.

Zetterberg, P. (2009). Do Gender Quotas Foster Women's Political Engagement? Lessons from Latin America. *Political Research Quarterly*, *62*(4), 715–730.

Zetterberg, P. (2012). Political Engagement and Democratic Legitimacy in Mexico. In S. Franceschet, M.L. Krook, & J.M. Piscopo (eds), *The Impact of Gender Quotas* (pp. 173-189). Oxford Academic.

Table 1: Analysis of a Gender Quota as a Randomized Experiment

	West Bengal state				Rajasthan state			
	Chairperson post				Chairperson post			
	Reserved	Unreserved	Difference (2) - (1)	Standard error	Reserved	Unreserved	Difference (6) - (5)	Standard error
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Characteristics of reserved and unreserved villages, 1991 census								
Population	974	1022	-49	(75)	1249	1564	-315	(212)
Female Literacy Rate (fraction)	0.35	0.34	0.01	(0.01)	0.05	0.05	0.00	(0.01)
Male Literacy Rate (fraction)	0.57	0.58	-0.01	(0.01)	0.28	0.26	0.03	(0.03)
Fraction of villages with tap water	0.05	0.03	0.01	(0.03)	0.12	0.09	0.03	(0.06)
Panel B: Gender quota effect on public goods investments								
# of drinking water facilities newly built or repaired	23.83	14.74	9.09	(4.02)	7.31	4.69	2.62	(0.95)
Condition of roads (1 if in good condition)	0.41	0.23	0.18	(0.06)	0.9	0.98	-0.08	(0.04)
# of irrigation facilities newly built or repaired	3.01	3.39	-0.38	(1.26)	0.88	0.9	-0.02	(0.06)
Panel C: Gender quota effect on political participation								
% women participants in village council meetings	9.8	6.88	2.92	(1.44)	20.41	24.49	-4.08	(4.03)
Have women filed a complaint in the last 6 months	0.20	0.11	0.09	(0.05)	0.64	0.62	0.02	(0.10)
Have men filed a complaint in the last 6 months	0.94	1.00	0.06	(0.06)	0.95	0.88	0.073	(0.06)

Notes: Figures from Chattopadhyay and Duflo (2004). Columns 1 and 2 reflect average values in villages with reserved chairperson positions and villages without such reservation, for the state of West Bengal; columns 5 and 6 provide corresponding figures for the state of Rajasthan. The standard error is based on a t-test for equality of means across the sample of reserved vs unreserved villages.