

Powerful Women: Does Exposure Reduce Bias?

Lori Beaman, Raghavendra Chattopadhyay, Esther Duflo,
Rohini Pande and Petia Topalova*

March 13, 2008

Abstract

Many countries have introduced political affirmative action for women to improve female representation, often with the explicit hope of reducing perceived voters' bias against women as policy makers. This paper examines whether such mandated exposure can affect voter bias. We combine experimental and survey data on voter attitudes with exogenous variation in exposure to female leaders across Indian villages to show that voters do appear to be biased, but this bias is reduced by exposure. For example, voters, particularly men, rate a leadership speech more negatively when the speaker's voice is experimentally manipulated to be female (rather than male). However, this bias is, on average, absent or even reversed in villages which have had a female leader. Women appear to be somewhat less biased initially, but their opinions of the leaders on the speech is unaffected by their exposure. Voter evaluations of actual leaders mirror the experimental findings. Villagers rate first-time women leaders worse than, but second-time women leaders at par with, male leaders.

*The authors are from RWJ Berkeley, IIM Calcutta, MIT, Harvard and IMF respectively. The views expressed in this paper are those of the authors and do not implicate the International Monetary Fund, its management, or Executive Board. We thank MIT, YCIAS Yale, Ash Center Harvard and UNICEF for funding, Catherine Lee for exceptional research assistance, and Prasad Chakraborty, and the team of SRG investigators for fieldwork. We also thank Tim Besley and Shawn Cole for comments.

1 Introduction

In July 2006, women accounted for just under 17 percent of parliamentarians world-wide and a woman was the head of government in only seven countries (UNICEF, 2007). These gender disparities in leadership do not reflect institutionalized discrimination against women in politics. Women can vote, support candidates and run for office in almost every country. The possibility that the explanation lies in voter discrimination has motivated affirmative action for women in politics, either by law or by political parties, in over seventy countries (Dahlerup, 2006).

While political affirmative action has increased female representation in politics (Jones, 2004), and, in many cases, altered policy decisions (Chattapodhyay and Duflo (2004), Powley (2007)) we know little about how they have affected voter attitudes. In this paper we exploit random variation in the exposure of Indian villagers to female leaders, induced by the institutional features of political affirmative action, to examine this question. Such evidence can both improve our understanding of how public opinion is formed and, more specifically, identify whether a temporary affirmative action program can cause permanent gains for women in politics.

A large literature on public opinion formation suggests that voter awareness on specific issues is, in general, low (Zaller, 1992); as a result, stereotypes or other group-centered attitudes play an important role in shaping voter preferences, often at the unconscious level (see, for instance, Huddy and Terkildsen (1993), Kahn (1994), Golebiowska (2001), Sigelman et al. (1995), and on unconscious influences Mendelberg (2001), Valentino et al. (2002)). Further, a growing body of evidence suggests that sex-biased attitudes seem to lead to gender discrimination in hiring and leader evaluations (see, e.g. Goldin and Rouse (2000), Eagly and Karau (2002), Duflo and Topalova (2004)).

But can exposure to female leaders alter voter attitudes? If gender stereotypes reflect incorrect voter beliefs about female ability to lead (Arrow (1973), Phelps (1972)), then exposure may be an important way of altering these beliefs. But, if voter beliefs determine candidate electability, then exposure to female leaders may not be achieved without affirmative action. Further, if male and female politicians have different policy preferences

(Chattapodhyay and Duflo (2004), Rehavi (2008)), then affirmative action can play an important role in ensuring equal political voice for men and women (Pande, 2003).

However, the desirability of affirmative action may be questioned on two fronts. First, sex-biased attitudes may reflect not beliefs but preferences for male leaders (Becker, 1957). One reason for a voter distaste for female leaders may be societal norms of prescribed behavior which associate leadership activities with men (Akerlof and Kranton (2000), Eagly and Karau (2002)). In such a situation mandated exposure to women leaders can alter voter attitudes only if exposure challenges, and changes, underlying social norms. If not, such mandates are likely to be ineffective or worse still, by violating voters' sense of self and others (for instance, it may be perceived as reducing the value of a traditionally male activity (Goldin (1990)) may precipitate a backlash, resulting in even lower approval ratings for women politicians (Rudman and Fairchild (2004)).

It is also possible that affirmative action policies per se have a perverse effect on voter opinion. Voters may perceive affirmative action as reducing the incentives of women politicians to invest in job-specific skills. In such situations mandatory increases in the number of women politicians may cause voters to perceive women as less competent, and women politicians may, in turn, respond by lowering their investment in job skills and putting in less effort (Coate and Loury (1993)). Affirmative action may also directly breed voter resentment with having their leadership choice curtailed, and this may be reflected in worse evaluations of female leaders.¹

In assessing the relative importance of these arguments, a central piece of evidence is how exposure to female leaders alters voters' beliefs about women leaders. A failure of voters to update their beliefs about women's ability to lead, or indeed a worsening of public opinion towards women leaders, strongly militates the view that mandated increases in women politicians can break down negative stereotypes about women leaders.

A number of empirical studies focus on how support for ethnic parties is affected by neighborhood composition (Giles and Hertz (1994)). Others examine how exposure to members of a social group affects attitudes towards intergroup socialization. While many find that exposure improves intergroup attitudes, these studies are unable to deal with the

¹Thernstrom and Thernstrom (1997) express a similar concern in the context of racial affirmative action.

concern that those who are more tolerant of other groups are more likely to associate with members of other groups (Pettigrew and Tropp, 2000). One important exception is Boisjoly et al. (2006) who exploit random roommate assignment in a large US university to show that students' political opinions become more aligned with the interests of the social groups to which their (randomly assigned) roommates belong.

Similarly, a positive correlation between voters' attitude and the characteristics of their elected representative would likely reflect in part reverse causality. Consequently, causal evidence on the existence of bias and how it is shaped by exposure mainly comes from experimental studies where respondents are randomly exposed to either male or female leaders (for instance, through speeches or vignettes, e.g. Eagly and Karau (2002)). Most of these studies, however, remain laboratory based. Furthermore, while they establish the presence of a bias, they provide limited guidance on whether exposure to women leaders is likely to alter gender stereotypes on leadership.

The contribution of this paper is to provide evidence from a large-scale natural randomized experiment in India on how female leadership alters voter attitudes. Random differences in exposure were induced by political affirmative action (Chattapodhyay and Duflo, 2004) – since 1998 one third of village council leader positions have been randomly set aside (or “reserved”) for women; in reserved councils only women can run for the position of leader.

Our analysis makes use of an array of experimental and survey measures of voter attitudes. We measure voters' explicit preferences by villagers' general opinion of female and male leaders. We also measure their implicit preferences (or unconscious mental attitudes) in two ways. First, by computer-based Implicit Association Tests (IATs) – each test is a different double-categorization task which measures unconscious bias in relating men and women to different categories (either normative or occupational). Second, by a hypothetical leadership speech and vignette where we randomize the gender of the leader and the respondent was asked to rate the leader on a variety of dimensions. Implicit and explicit measures potentially vary in the extent to which they reflect prevalent social norms and villagers' judgement of the competence of specific leaders (Betrand et al. (2005)).

Both explicit and implicit attitudes show, on average, a significant dislike for female leaders – particularly among male villagers. For example, on a scale of 1 to 10, the average

man ranks female leaders one whole point below male leaders. Similarly, by measuring the respondent's normative attitude towards female leaders using the IAT, we find that men are significantly more likely to associate male leaders, relative to female leaders, with notions of "good" and with leadership tasks. Villagers also have a low opinion of female leadership ability in villages that have never been exposed to women: relative to men who heard a man deliver the leader speech or listen to a scenario where the man made a decision, those who had to rate a woman rated her on average 0.074 standard deviation lower. The point estimate is lower for women but not statistically different.

Exposure to a female leader does not affect either villagers' explicit attitudes, as measured by a simple ranking "ladder," or their implicit distaste for female leaders in general, as measured by the IAT associating male and female leaders with concepts of good and bad. However, in the IAT associating male and female names with domestic versus leadership tasks, exposure significantly reduces the positive association between male names and leadership, relative to female names and leadership. We interpret this as evidence that exposure to female leaders causes men to recognize that women can be leaders, even if the association remains undesirable. Clear cut evidence on this is provided by the speech and vignette experiments where exposure to a female leader completely erases bias: in villages where the leader position was reserved for women at any point, the rating of hypothetical leaders in the speech and vignette does not depend on leader's gender. Since assignment of political reservation is random across villages, this provides causal evidence that exposure reduces bias.

Finally, we examine the evaluation of actual leaders. We ask villagers to judge their leader along an array of dimensions, including general effectiveness. Consistent with the experimental data, we find evidence that greater exposure improves evaluation of female leaders. Specifically, in villages where the leadership position is reserved for the first time in the current electoral cycle, i.e. since 2003, male villagers give leaders significantly lower ratings (relative to villages where the leader position has never been reserved). Further, and again consistent with experimental data, this difference in ratings is absent for villages where the leader position is reserved for women for the *second time* in 2003. While it is possible that female leaders in villages which are reserved for the first time in 2003 are

differentially selected or undertake different policies than male leaders or female leaders elected in the second round of reservation, we do not find *prima facie* evidence in support of this hypothesis. First time female leaders have similar characteristics and deliver more public goods, at a lower price, than their male counterparts. They do marginally better than women leaders in places reserved for the second time and all women leaders appear equally biased in favor of women in their policy actions.

The rest of this paper is structured as follows. Section 2 describes the field setting of our study. Section 3 describes the data we use and our basic empirical strategy. Section 4 compares voter attitudes towards male and female leaders, and examines how reservation alters these attitudes. Section 5 concludes.

2 Background: Political Reservation

Electoral quotas for women usually require that women must constitute a certain number or percentage of the members of either a candidate list or an elected body such as the parliamentary assembly.² However, a number of studies show that requirements on candidate lists may not significantly increase actual female representation in politics (Frechette et al. (2006), Jones (2004)). As a result, progressively more gender quotas are being introduced using reserved seat systems, and increasingly women elected on reserved seats quota systems are not appointed, but elected like in Jordan, Uganda and Rwanda.

In India, political reservation for women was introduced at the local level in 1993. Universal franchise is in place in India since Independence, and India has had a number of very prominent female leaders (Indira Gandhi, Sonia Gandhi, Mayawati, Mamata Banerjee, and now Pratiba Pathil, the first female President are a few examples). However, the share of women politicians at the national level and in most Indian states has consistently remained under 10 percent. A 1993 constitutional amendment simultaneously devolved significant

²These quotas are usually mandated either by the constitution (like in Burkina Faso, Nepal, the Philippines and Uganda), or by electoral law (as in many parts of Latin America, as well as, for example, in Belgium, Bosnia and Herzegovina, Serbia and Sudan). In other countries, including Argentina, Bolivia, Ecuador, Germany, Italy, Norway and Sweden, a number of political parties have some type of quota.

powers to village councils (Gram Panchayats) and required that one-third of village council leader (Pradhan) positions, be reserved for women.³ Pradhan positions were also reserved for the two disadvantaged minorities in India, Scheduled castes (SC) and Scheduled tribes (ST), in proportion to their population share in the district. This led to a dramatic rise in female leadership in local government. The number of village-level female elected leaders is now close to 40 percent.

This paper focuses on West Bengal, a state where the Panchayat has been an effective body of governance since 1978 with regular contested elections. The Panchayat Constitution Rule of West Bengal, which was modified in April 1998 (Government of West Bengal, 1998), provides detailed rules for the implementation of reservations mandated in the constitutional amendment. For each election, GPs are first randomly assigned to three groups: Reserved for SC, Reserved for ST, and Unreserved. Within each group, GPs are then ordered by their serial numbers and every third GP in each list is reserved for a woman. In 1998, GPs starting with number 1 on each list were reserved for a woman, and in 2003 GPs starting with number 2 on each list were reserved. This assignment rule has three consequences. First, GP are randomly assigned to be reserved both for women and for SC and ST. Second, the randomization for women is implicitly stratified by SC/ST and by administrative blocks (since GP serial numbers start with the block ID).⁴ Second, a given GP may be reserved twice in a row – for instance, if it was the first on the list in 1998 and the second on the list in 2003.

To confirm that the rules were followed, we obtained GP serial numbers and the tables in the electoral law and reconstructed the reservation list. When we compared it to the list of reserved GPs, we found that the rule held, with no exception.⁵ Table 1 provides descriptive statistics on GP reservation. Of the 56 GPs reserved in 1998, 20 were also reserved in the 2003 election. 35 GPs were reserved for the first time in 2003 and 74 GPs have never been reserved. The reservation policy is rigorously implemented. All the Pradhans in GPs

³A Gram Panchayat typically consists of multiple villages (in West Bengal the average GP has 10-12 villages).

⁴For administrative purposes Indian districts are subdivided into blocks, the administrative level just above the village.

⁵The same was true in 1998, see Chattapodhyay and Duflo (2004)

reserved for women are indeed female.⁶

Overall, the fraction of female Pradhans in Birbhum stands at 43 percent. Election of women from non-reserved GPs is thus low, but has increased over time. However, *prima facie* the increase is unrelated to the reservation policy. Between 1998 and 2003 the fraction of women Pradhans elected from never reserved GPs increased from 7% to 16%; this is very comparable to the fraction of women elected in 2003 from GPs that were only reserved in 1998 (which is 14%). Turning to re-election, 17 of the 55 women elected on reserved seat ran again, and 4 of them were reelected. Of these, 3 were re-elected from the 20 GPs which continued to be reserved in 2003, and only 1 was re-elected on an unreserved seat.⁷ In comparison, out of 72 men, 19 run again, and 4 got re-elected. So the share of men who run again is lower than that of women (in part because many of them are prevented to run because of rules), as is the unconditional probability of being re-elected. The probability of being elected conditional on running on an unreserved seat is however much higher.

In Table 2 we use 1991 census data to examine whether there are any systematic differences between the villages in our sample, based on the reservation status of the GP they belong to: Never Reserved, Only Reserved in 1998, First Reserved 2003 and Reserved in 1998 and 2003. We observe only one significant difference across villages in these four categories: there are more tarmac roads in villages which were reserved for women both in 1998 and 2003 than in those reserved only in 2003.

3 Data and Empirical Strategy

3.1 Data

Our data comes from Birbhum district in West Bengal, a middle-income Indian state. Birbhum is among the poorest districts in this state with close to half its population below the poverty line in 1999.⁸ Between June and December 2006 and May and November 2007 we

⁶The only exception is one reserved GP which, due to political disturbances, did not have a Pradhan at the time of our survey.

⁷Two of the five women elected from unreserved GPs in 1998 ran again, and none were re-elected in 2003.

⁸Birbhum ranked 14th out of West Bengal's 17 districts on its gender development index in 2004.

surveyed the 165 GPs in this district. We randomly selected three villages in each GP, and collected detailed household and village-level data in the 495 villages.

In each GP we interviewed the village leaders (Pradhans) in 1998 and 2003, their spouses and a random sample of 15 households per village. We administered a household survey to the most knowledgeable household member (usually male), and an adult module to a prime aged male and female respondent in the household (the Appendix describes the sampling procedure). Overall we conducted 7,425 household surveys, and 6,717 male and 6,780 female adult modules.

We also administered computer-based Implicit Association Tests (IATs), described in more detail below, to a random subset of 5 households per GP.⁹ Overall, we have 4,378 IAT respondents spread across 1,968 households. Finally, we collected data on the quantity and quality of public good provision in every sampled village (on this, see the data Appendix). Our surveys collected multiple measures of attitudes towards women leaders, and we briefly describe these below.

• **Explicit Attitudes**

To ascertain respondents' explicit opinion about women leaders in general, we asked "on a ladder which has steps from 1 to 10, how do you feel about [person]", where [person] was (differently) a male leader, a female leader, a male villager and a female villager.¹⁰

We also asked villagers to evaluate their actual Pradhan on multiple dimensions (using the same ladder based 1-10 scale). Questions include "Do you think the Pradhan has done a good job looking after the needs of your village?" and "How would you rank the effectiveness of the current Pradhan?"

• **Implicit Association Tests**

Our first measure of implicit attitudes towards female leaders are Implicit Association Tests

⁹In each household we conducted IATs with adults between the ages of 15 and 45 who were present at the time of the enumerator's visit. We set a threshold target of IATs with two household members (one male and one female), and households were visited multiple times to attain this target. In 171 households we had only one participant despite multiple visits. Seventy-seven GPs have a total of 10 respondents since villages within those GPs were used for piloting IATs and therefore excluded from our main sample.

¹⁰This measure is adapted from the "Feeling Thermometer" which is widely used in the psychology literature [ref].

(IAT) (Greenwald et al., 1998). These are computerized tests which aim to measure attitudes and stereotypes of which respondents may not be explicitly cognizant. To achieve this, it uses a double-categorization task to measure the strength of the association between two concepts. Words from two different series (for example a series of first names and a series of adjectives) appear on the screen, and the respondents sort them in two categories (e.g. female and male names, and adjectives evoking good or bad things). In some tasks, they must put male names and good attributes on the left, and female names and bad adjectives to the right; in other tasks, the categories are switched (with males names and bad attributes on the right, and female names and good attributes on the left).

IATs rely on the assumption that this sorting task should be easier the stronger the association between the two concepts. An automatic association, or implicit stereotype, can be detected by comparing response time across concept pairs.¹¹ We use the standard measure for IAT prejudice, the D-measure, which is the difference in the average response time in the two blocks divided by the standard deviation of response time in those two blocks (Greenwald et al. (2003)). The D-measure is normalized such that a higher D measure indicates a stronger implicit stereotype.

Since 1998, when the first IAT results were published, IAT based bias measures have been widely used in a diverse array of disciplines, including various subfields in psychology, neuroscience and market research (B.Nosek et al., 2007). In economics, Rooth (2007) has shown that the D-measure for racial bias of Swedish employers correlates positively with the rate at which they call back applicants of different races. To the best of our knowledge, the only other paper which investigates the malleability of IAT is Dasgupta and Asgari (2004) who show that a higher proportion of female professors is associated with college students' implicit beliefs about women professionals. However, interpretation of these results is limited by endogenous selection into women's colleges.

We implemented three IATs. The first assesses the associational strengths between categories of male and female names and attributes of good (nice, delicious, etc) and bad

¹¹Specifically, for each test, there are two test blocks of interest: a first that associates say, male names and positive attributes and female names and negative attributes; and a second that reverses this and associates say, male names and negative attributes and female names and positive attributes.

(nasty, horrible etc). IATs with this structure are widely used in social psychology to measure group prejudice, such as racism (Banaji, 2001). The second IAT measures the association between male and female *politicians* and attributes of good and bad. Politicians are represented by pictures of either men or women giving speeches, leading crowds, etc. Finally, the third IAT examines the association between male and female names and domestic (e.g. cooking, tending the animals, farming) and leadership activities (such as meeting, panchayat, etc). Rudman and Kilianski (2000) have, for instance, used such a stereotype based IAT to show that respondents associate female names much more with family than career.

We adapted these IATs so that they did not assume either literacy or familiarity with computers. Prompts were either audio or represented by pictures on the screen.¹² In the Appendix we show an IAT screen and a picture of a participant playing the game (for more details on the tests, and a full list of actual prompts, see Beaman et al. (2008)). Since our pilots suggested that older respondents struggled with the IAT we only conducted IATs with respondents between the ages of 15 and 45.¹³

• Evaluation of Hypothetical Leaders

Our second class of measures uses experiments of the “Goldberg paradigm” type to assess bias in the perception of the effectiveness of women as leaders (Matland (1994); Huddy and Terkildsen (1993); Eagly and Karau (2002)), which have been used to measure the extent of biases against female leaders in the US.

In the speech experiment, the respondent heard a short tape-recorded leader speech, which was inspired by an actual speech delivered at a village meeting. In the speech, the Pradhan responds to a villager complaint by requesting villagers to contribute money and effort for local public goods.¹⁴ The speech was recorded in three female and three male voices, and respondents were randomly assigned one recording. The respondent was told that this was a speech by a Pradhan in a village meeting in another district, and asked to assess the Pradhan along multiple dimensions. These included overall effectiveness, whether

¹²IATs have been similarly adapted for children, see Baron and Banaji (2006).

¹³We excluded respondents with an average response time over 6 seconds during the first test block or less than 65% average correct responses. We dropped all responses for which the respondent took more than 10 seconds to categorize the prompt (standard response time are measured in milliseconds).

¹⁴The English translation is in the Appendix.

he/she took the right decision in the matter at hand, whether he/she could be trusted to take the villagers’ interest, etc.

We conducted a similar exercise with “vignettes”: the surveyor described a situation where resources were scarce and the Pradhan had to choose whether to invest in a drinking water or an irrigation project. Vignettes varied along two dimensions: the Pradhan’s choice and the Pradhan’s gender. Each respondent was read a randomly selected vignette. The gender of the leader assigned to each respondent was the same on both the speech and the vignette.

As the different versions of the speech and vignette varied only along the gender of the leader, and participants were randomly assigned either a male or female leader, we examine whether, holding the actual action, villagers, on average, ranked female leaders lower than male leaders.

3.2 Empirical strategy

Given the randomized setting, our empirical strategy is relatively straightforward. The sample has four types of GPs: First Reserved 2003, Reserved 1998 and 2003, Only Reserved 1998, and Never Reserved. Randomization of reservation status implies that we can study its reduced form effect by comparing the means of outcomes of interest across GPs with different reservation status. Since all reserved GPs, but relatively few of the unreserved and previously reserved GPs, have a female Pradhan (Table 1), this reduced form effect is similar to that obtained by instrumenting for Pradhan’s gender by reservation status of GP. Thus, in what follows we focus on this reduced form analysis.

The unit of observation in our villager regressions is respondent i resident in GP g . Throughout we report results separately for male and female villagers. Our regression specifications vary across outcomes. For the difference in explicit bias, feeling towards male and female leaders, IAT D-measure of leader bias, villager evaluations of own Pradhan and the Pradhan’s action and characteristics, we report two specifications. The first simply compares the outcomes of interest across ever reserved GPs and never reserved GPs,

$$y_{ig} = \beta R_g + X_{ig}\gamma + \alpha_b + \epsilon_{ig} \quad (1)$$

where R_g is an indicator variable for the GP being currently or previously reserved (from now on, ever reserved). The second separates different types of reserved GPs and unreserved GPs

$$y_{ig} = \beta_2 R_{g2} + \beta_{2and1} R_{g2and1} + \beta_1 R_{g1} + X_{ig} \gamma + \alpha_b + \epsilon_{ig} \quad (2)$$

R_{g2} is an indicator for the GP being reserved for the first time in 2003 (during the second round of reservation), R_{g2and1} is an indicator for the GP being reserved in 2003 and 1998, and R_{g1} is an indicator for the GP being reserved in 1998, but not in 2003.

All regressions include a set of respondent controls (X_{ig}): age, household size, education, caste, religion and proxies for household wealth constructed using a principal component analysis.¹⁵ α_b is a block fixed effect. Standard errors are clustered by GP in all specifications.

In the case of vignettes and speech, we are interested in whether male and female leaders are judged differently, and whether this varies with reservation induced exposure to a female leader. Let F_{ig} indicate whether respondent i was presented with a "female" stimulus (for example, was asked their opinion of a female leader in general, heard the speech in a female voice, or was described the vignette with a female leader). We estimate:

$$y_{ig} = \delta F_{ig} + \lambda(R_g * F_{ig}) + \mu R_g + X_{ig} \gamma + \alpha_b + \epsilon_{ig} \quad (3)$$

and

$$y_{ig} = \delta F_{ig} + \lambda_2(R_{g2} * F_{ig}) + \lambda_{2and1}(R_{g2and1} * F_{ig}) + \lambda_1(R_{g1} * F_{ig}) + \sum_k R_k \lambda_k + X_{ig} \gamma + \alpha_b + \epsilon_{ig}, \quad (4)$$

where $k \in 2, 2and1, 1$ controls for the main effect of different reservation categories.

In equation (3) the coefficients of interest are δ which captures bias towards female leaders in unreserved GPs, and λ which indicates whether current or past exposure to a female leader changes the level of bias. In equation (4), we are interested in λ_2 , λ_{2and1} and λ_1 , and how they differ from each other.

Finally, on many issues of interest we ask respondents multiple questions. Often the expectation is that within some groups (or "families") of outcomes, the coefficients of the variables of interest should go in the same direction. To avoid drawing inferences based

¹⁵Regressions without controls variables are very similar to the results reported here, and are available from the authors.

on selected outcomes, we also report effects that average across all the outcomes within a family, following Kling et al. (2007). Specifically, for each outcome we construct a normalized transformation where we subtract the mean for the never reserved GPs and divide by the standard deviation. We estimate standardized effects using a seemingly unrelated regression (SUR) system, where we account for correlation across outcomes and average across outcomes to obtain an average effect for the “family” of outcomes.

4 Results

4.1 Preferences for Female Leaders

In table 3, we first investigate the preferences for female leaders in general (as opposed to the evaluation of the competencies of a specific female leader).

Starting with explicit attitude in column 1 and 2, we see that villagers are more favorable towards male Pradhans, and are not shy about admitting this. In unreserved villages, males villagers rank female leaders -0.89 points lower than male leaders. For female villagers, the difference is smaller (-0.354) but still significant.

This distaste for female leaders is not ameliorated by exposure. The coefficient of “ever reserved” is negative and insignificant for both men and women (panel A), as are all the reservation category dummies in panel B. This result is, at first blush, somewhat discouraging: especially since most of the literature suggests that respondent are *more* biased than what they are willing to explicitly admit Bertrand et al. (2005). If this is also true for Indian villages, then the evidence on explicit attitudes suggests that reservation policy will, at best, do nothing and may even lead to a backlash.

On the other hand, villagers’ willingness to express their relative distaste for female leaders may simply reflect the prevalence of social norms, which consider bias permissible. In column (3) to (6) we report the results of two IATs designed to measure implicit preferences for female and male villagers (columns (3) and (4)), and female and male leaders (columns (5) and (6)). In columns (7) and (8), we present results of an IAT which capture villagers’ stereotypes about female leaders.

The outcome y_{ig} is the “IAT effect” (or D-measure): the normalized difference in the mean response time between the “non-stereotypical” and “stereotypical” test blocks. The stereotypical block in the first two IATs places male names (or pictures of male leaders) and adjectives evoking good attributes on one side of the screen, and female names (or pictures of female leaders) and adjectives evoking negative attributes on the other side.

A positive value for y_{ig} thus indicates an implicit bias against female villagers or leaders (since it indicates that the respondent is slower in associating female leaders or villagers and good things, than female leaders or villagers and bad things).

We estimate equations (1) and (2) for the three measures. The mean of the unreserved sample indicates the level of bias, and the reservation indicators how it varies with exposure to female leaders. A negative coefficient on the reservation dummies represents a reduction in the underlying implicit bias.

Both sexes show strong same gender preference. Men are significantly more likely to associate good with male names.¹⁶ In contrast, women associate female names with positive attributes. Not surprisingly, neither is affected by reservation. In columns (3) and (4) we observe very similar results for implicit bias towards male and female *leaders*. Both genders exhibit same-sex preference, and this is unaffected by reservation.

The similarity between these two IATs may reflect the fact that in both IATs, villagers perceived mostly the gender of the character, and did not associate the pictures of female leaders with leadership. In that sense, the second IAT may have failed to capture the “leadership” dimension. Thus, just like explicit preference, which showed a fairly strong and persistent dislike for female leaders, implicit bias for female leaders does not seem to be affected by exposure based on these two IATs.

¹⁶This finding is in contrast with what is typically found in developed countries, where these tests exhibit a “women are wonderful” effect (Eagly and Mladinic, 1989).

4.2 Perception of female competencies

4.2.1 Implicit Association Test: Women and Leadership

The third IAT, however, also presented in table 3 (columns (7) and (8)), paints a very different picture relative to the taste IATs discussed above.

In this IAT, the stereotypical block places on one side of the screen male names and leadership activities, and female names and domestic activities on the other side. Non-stereotypical blocks reverse this association. While each person is assigned both stereotypical and non-stereotypical blocks the ordering of these blocks is randomized. A positive value of y_{ig} indicates that the respondent associates women with domestic activities and men with leadership activities.

Both genders are biased against women in leadership activities in never reserved GP. In those GPs, respondents are faster in associating women with domestic activities than with leadership actions (men are 0.1 standard deviation faster, and women are 0.15 standard deviation faster, on average, and in both cases, the baseline D measure is significant in the never-reserved villages). Exposure to a female leader as a result of the reservation policy, however, significantly reduces this association for men. There is no such impact for women, however. If anything, the stereotype associating women and domestic activities increased for those women exposed to a female Pradhan for the first time in 2003.

4.2.2 Speech and Vignette

The results of the IAT associating women and leadership suggest that exposure to female leaders induced by the reservation may have reduced the stereotype linking men with leadership activities, at least among men. This suggests that men may have learnt that women can be capable leaders.

To examine this further, we analyze evaluations of speech and vignette descriptions of hypothetical *specific* leaders undertaking typical Pradhan activities. In both cases the respondent is asked to judge a specific action by the Pradhan. We conjecture that in judging the Pradhan the villager will base his or her assessment on own previous experience – whether similar decisions or actions taken by his or her Pradhans in comparable situations had been

good decisions.

In this case we may expect respondent's exposure to female Pradhans to influence whether (and how) the gender of the hypothetical Pradhan in the speech and vignette colors his or her judgement. Consider the situation where all villagers initially believe that women make worse leaders, but villagers who are exposed to female Pradhans update their beliefs based on their actual Pradhan performance. Then we would expect to see that villagers in never reserved villages are less likely to approve the hypothetical female Pradhan even when she takes the same decision as the hypothetical male Pradhan. However, if actual female Pradhans are relatively good leaders, this bias will be diminished in ever reserved villages.

In Table 4 we report the overall leader evaluation, averaging across both the speech and irrigation vignette (recall that respondents were exposed to the same gender for the speech and the vignette). In panel A we examine (using equation (3)) the effect of current or past reservation on overall evaluation for male and female respondents respectively. We see that men in never reserved villages rate a Pradhan 0.074 standard deviation lower if she is a woman (with a standard error of 0.029). The bias is lower and insignificant for women (0.049, with a standard error of 0.034), but we cannot reject the hypothesis that women and men have the same bias.

Reservation has a large impact on men's opinions: the coefficient on the interaction between female Pradhan and ever reserved is 0.122, and is strongly significant. If anything, adding the coefficients on female pradhan and its interaction with the reservation indicator, the bias is now reversed, though the resulting pro-female bias is not strongly significant. In contrast, there is no impact of reservation on female's opinion of the female Pradhans.

In panel B, we disaggregate these results by type of reservation (present and past). The results suggest that no single category of reservation differs significantly in its effect on leader evaluation for male respondents: all coefficients are positive, and the point estimates are similar in magnitude, and statistically indistinguishable. It seems that having been exposed at some point to at least one female leader is sufficient to erase the bias against female leaders, and that the effect persists even after the woman has been replaced by a man. Once again, there is no perceptible effect for women.

In columns (4) through (12) we provide results for the averages across groups of

questions which enter the overall assessment. For men, the same effects (bias against women, and preference reversal) are present for all the individual measures, and significant for most of them. Once again, the estimated coefficients on the different reservation categories are very similar in magnitude. For women, exposure to female leaders does not seem to affect any of the measures.

4.3 Perceptions of Actual Leaders

The experimental evidence strongly suggests that villagers, and, in particular, men, exhibit significant bias against female leaders, and this colors their judgment of leader actions. However, prior exposure to a female Pradhan attenuates such bias. We now examine whether similar forces are at work in villager evaluations of their actual Pradhan.

4.3.1 Overall Rating

We asked each respondent to evaluate different dimensions of their Pradhan’s performance on the same 1 to 10 ladder. Many of the questions were purposefully chosen to be similar to those asked about the hypothetical Pradhan in the speech and vignette.

Table 5 shows that villagers are significantly more dissatisfied with their Pradhan in GPs which were reserved for women for the first time in 2003. Column (1) examines the average across various Pradhan satisfaction measures for male respondents. The first row shows that Pradhans in GPs that are first reserved for a woman in 2003 score 0.21 standard deviations less than Pradhans in unreserved GPs; this coefficient is strongly significant. Column (2) shows a negative, but smaller, effect for female villagers (-0.10 standard deviation, with a standard error of 0.06). These results echo those in Duflo and Topalova (2004) (see also Duflo (2005)), which, using nation-wide data for India, showed that villagers are more dissatisfied with female leaders in general.

Looking across individual measures (columns (3)-(10)), the coefficients of interest are very similar across satisfaction measures – female Pradhans in GPs first reserved in 2003 are ranked lower on general effectiveness, ability to look after village needs or respondent’s own needs, and in preparing the list of beneficiaries for the “below the poverty line” (BPL)

program.

Moreover, consistent with the experimental data, prior exposure to a female Pradhan affects a villager's opinion of current female Pradhan. The second row in Table 5 shows that the villager's opinion of female Pradhans in (reserved) GPs that were also reserved in 1998 is statistically undistinguishable from that of Pradhans in unreserved GPs. This is analogous to the result in the speech and vignette, where (male) respondents exposed to a female stimulus ranked leaders in the speech or vignette lower in unreserved GPs, but not in ever reserved GPs. Contrary to the speech and vignette results, this is now true for female respondents as well. The difference between the coefficient of the appreciation of the Pradhan in the GP reserved for the first and the second time are significant, at the 1% level in the male regression, and the 8% level in the female regression.

Finally, row 3 shows that (mostly male) leaders in GPs that were reserved in 1998, but are not in 2003 are evaluated similarly to those in other unreserved GPs.

4.3.2 Do Overall Ratings Reflect Performance?

These results are consistent with the hypothesis that villagers (men more so than women) start with negative opinions of female leaders, but update their (performance based) opinion in response to seeing women leaders. While the first leader does not benefit from it, likely because people are usually internally consistent in how they rate specific people, both hypothetical leaders and future actual leaders do. Nevertheless, we need to be cautious in interpreting these results, since, unlike the hypothetical leaders in the speech and vignette, male Pradhans and second time female Pradhans may, in reality, be better leaders than first-time female Pradhans.

Tables 6 and 7 present some suggestive evidence that, *prima facie*, this does not appear to be the case.

In Table 6 we examine whether public good provision and the allocation of BPL cards suggests worse performance by female Pradhans. In column (1) we analyze the quantity of public goods that have been built or repaired in the villages since the last election. Across all goods, the average number of repairs or new constructions in villages in currently reserved GPs is 0.21 standard deviations higher. The detailed results are presented in the appendix.

As in Chattapodhyay and Duflo (2004), investments in drinking water and sanitation, irrigation and roads are higher in currently reserved GPs. Differently from them, we find an insignificant effect on education and a marginally significant positive effect on investment in public health facilities. If anything, Pradhans in GPs reserved for the first time in 2003 seem more effective than Pradhans in never reserved GPs.

In GPs reserved for the second time, the quantity of public goods provided is still somewhat higher than in never reserved GP, but the difference is smaller and insignificant (0.096, with standard error of 0.08). The coefficient is not statistically different from that of GP reserved for the first time.

Women Pradhans may invest in more, but lower quality, public goods. However, column (2) shows no significant differences in the quality of public good provision for GPs which are currently reserved, previously reserved or twice reserved. This is also reflected in the villager's satisfaction with each individual good (which is different from what Duflo and Topalova (2004) found in an Indian-wide data set: they found that on average, villagers were less likely to be satisfied with the public goods in their villages when women were in charge). Furthermore, this is true for each individual goods (results in appendix)¹⁷ Once again, there is also no difference between women elected in GPs reserved for the first time and women elected for the second time.

Since these goods are mainly financed by contributions from the State Government, the contrast between villagers' (especially male) negative evaluation of female Pradhans and the fact that female Pradhans seem to have invested more, with no discernible reduction in quality, is unlikely to be explained by men resenting a "big government" approach. There are, however, (at least) two other ways in which villagers may have to pay for these goods: voluntary contributions and bribes. In our household survey there were two specific public goods for which bribes and voluntary contributions are well identified. The first is whether the household had to pay a bribe to receive a BPL card. The second is how much they had

¹⁷The one exception is the BPL list, as we saw in Table 7: men, in particular, resent the way female leaders allocate them. However, since there is a BPL quota there is, in a sense, no way to satisfactorily allocate BPL cards in a way that pleases everyone. This suggests that dissatisfaction with BPL list construction is likely to reflect general disapproval.

to pay (in bribes and voluntary contributions) for drinking water. In column (3) we see that on average, individuals in currently reserved GPs (for the first or the second time alike) are much less likely to have paid a bribe, and this effect is driven by BPL card. There is no difference in payments for water maintenance.

In summary, women leaders provide more public goods, of equal quality, at a lower price. While we may have failed to measure some critical aspect of Pradhan performance, the weight of the evidence is consistent with the experimental findings. It appears that villagers rate more negatively female leaders even though they perform at least as well as the average male Pradhan. Column 7 shows that women take different decisions than men, and that their decisions are biased towards women's preferences. This column reproduces the specification in Chattopodhyay and Duflo (2004): it uses the difference in male and female villagers' complaints to construct a measure of female preferences, and provides a summary measure of whether the allocation of the public goods is aligned more closely with the preferences of female villagers. Column (7) shows that women invest more in goods preferred by women. This may explain why male voters prefer male Pradhans, even though they recognize that women are as effective and competent.

Female pradhans in GPs reserved in both elections seem to be just as likely to reflect female preferences in their policy choices as women elected in GPs reserved for the first time. This is surprising since, during the second round of elections, parties may field different candidates or voters may select different types of women, among the pool of possible candidates, once they have realized that women's decisions reflect the policy preferences of women. However, this does not seem to be the case, and female leaders in twice reserved GPs continue to put greater weight on women's priorities.¹⁸

To further investigate possible selection of female pradhan in twice reserved GPs, we compare some salient leader characteristics, by reservation status in Table 7. Unsurprisingly, male and female leaders differ: women are younger, less educated, more likely to come from landless households and have a lower wealth index.¹⁹

¹⁸Surprisingly, this is also true for GPs that were formerly (but not currently) reserved.

¹⁹The wealth index is constructed from a principal component analysis of the household assets. See Appendix for detailed description.

However, Table 7 does not reveal any systematic differences between women leaders in first and second time reserved GPs. In particular, very few women elected in second time reserved GPs have any Pradhan experience, and they do not have significantly more experience as GP council members. While this may seem surprising, recall that the reason why some GPs are reserved twice in a row for a woman is because they have changed reservation category.²⁰ The person who was elected in 1998 is thus either ineligible for reelection in 2003 if the seat moved from the general to unreserved category, or is unlikely to be elected if the seat moved from the SC to general category (since SC and ST members are rarely elected from general seats). As a result, different women (all first time leaders) were typically elected in GPs reserved in 1998 and 2003.

In summary, the evaluation of the actual Pradhans appears to closely mirror the evaluation of the hypothetical ones. The first time men are exposed to a female Pradhan (either real or hypothetical), they dislike her general performance. The second time, this dislike disappears. The effects for women go in the same direction, but are attenuated and insignificant. While we can not fully rule out the possibility that men’s dislike for Pradhans in GPs reserved for the first time and their favorable evaluation of Pradhans elected in GPs reserved for the second time just reflects different performances on their part, the consistency between these two sets of results does suggest that exposure to female leaders has reduced the bias.

4.4 Robustness Checks and Alternative Explanations

Table 8 examines a few alternative explanations for the results, other than gender bias per se.

A first possibility is that women are ranked lower than men, not because of their gender, but because they are new and inexperienced. In the vignette and speech experiments, respondents may correctly assume that the female Pradhans are also new and inexperienced, which would explain their judgement. To assess whether this could explain our results, in panel I, we regress the respondent’s evaluation of the Pradhan on whether the Pradhan was

²⁰For example, they were reserved for SC in 1998 and were the second in the list for this group, and they are in the “general category” in 2003, and are the second (or fifth) in the list in that group.

elected to any seat in the GP (not necessarily Pradhan) for the first time in 2003 (about 60% of males were previously elected, while, as shown in Table 7, the experience of women leaders is significantly lower). There is no correlation between whether the Pradhan is new and their general evaluation.

A second possibility (along the line of Thernstrom and Thernstrom (1997)) is that voters have a low opinion of women leaders because they were elected on a reserved seat, not because of their gender. This, however, cannot explain the absence of bias in the case of women elected in GP reserved twice in a row, unless voters get used to the idea of reservations. In this case, voters should rank other reserved Pradhans lower than unreserved Pradhans, but their ranking would improve in places which have been reserved for women previously. In panel II, we show that while voters do rank SC Pradhans lower than non SC Pradhans (it is quite possible that they are discriminated against too), their opinion is not influenced by past reservations for women. Similarly in panel III we see that the negative evaluation of first time female leaders is unaffected by whether the GP was previously reserved for SC.

Finally, in Panel IV, we investigate whether past reservation for SC lessens voters' bias in the speech and vignette. This is not the case. It is really exposure to women that makes a difference, not exposure to reservation in general.

4.5 Pradhan's experience

The evidence we presented suggests that the first cohorts of female Pradhans elected in reserved GPs may have encountered significant bias, which was substantially reduced for the subsequent cohort of female leaders (in GPs reserved for the second time). This would suggest that the job of the Pradhan elected for the first time would have been more difficult. In this subsection we present suggestive evidence that this may have been the case.

In Table 9 we use data from the Pradhan surveys to examine whether women leaders elected from GPs reserved for the first time in 2003 had a different personal experience than the women in the twice reserved GPs. We first asked women leaders whether they believed that gender caused problems for them in discharging their duties as Pradhans. Column 1 reports the means of the answer to this variable for Pradhans in GPs reserved for the first and second time, and a test for their difference. Forty-one percent of the first time female

Pradhans reported that their gender created problems. Only 10 percent of the second time Pradhans did. This large and significant difference may be linked to the barriers that first time women face.²¹

In column 2, we report the result of a regression of Pradhan’s satisfaction with life (on a scale from 1 to 10, using the standard phrasing for this question on self reported happiness), on dummies for whether the GP is reserved in 2003, and reserved in 1998 and 2003. Female Pradhans in GPs first reserved for women in 2003 are significantly less happy with their life than Pradhans elected in unreserved seats (the coefficient is -0.47, and the mean for men is 3.7). This difference completely disappears for Pradhans elected in GPs reserved for the second time (the coefficient becomes a positive 0.379). The difference between the satisfaction of the Pradhans in these two types of GP is strongly significant.

Finally, column (3) reports the results from a similar regression, which looks at Pradhan’s plans to run for re-election. The evidence goes in the same direction, though it is much noisier. Women elected in GPs reserved for the first time in 2003 are less likely to plan to run again. The point estimate is large but not significant. This difference disappears for Pradhans in twice reserved GPs, though again the estimate is noisy enough that we cannot reject equality.

Overall, this data is suggestive that bias may take a toll on the first cohort of female leaders. However, this improves for the second cohort of women. This finding is in line with the bias results.

5 Conclusion

This paper shows that voters are biased against female leaders. In villages that had no experience with female Pradhans in the past, hypothetical leaders performing the same action are rated lower if they are female.

Political reservation, which forces voters to be exposed to a female leader, does not

²¹Although this is somewhat speculative, difficulties encountered by female Pradhans may also explain why, in Table 3, after one round of reservation, women were significantly *less* likely to associate female with leadership activities, but this was not the case in places that had been reserved twice.

alter villagers' stated opinion of women leaders. They also judge their own female leaders harshly in places that are reserved for the first time, despite the fact that women, if anything, appear to do a better job. This may explain why very few female leaders were re-elected once their seat became unreserved in the second election cycle after reservations were introduced.

The implicit bias about female leaders is, however, affected by past experience, at least for male voters. In places which have been previously reserved for a woman, the difference in the evaluation of hypothetical males and females leaders entirely disappears for men. Women have a lower initial bias, but this bias is not affected by reservation.

Moreover, this is reflected in the judgement of actual female leaders: women elected in villages that are reserved for the second time are judged to perform at par with men, both by male and female villagers.

Note that the general evaluation of the leaders' competence may not translate into higher votes for women if reservations were to be removed, especially since a general dislike appears to persist for the idea of having female leaders. It is too early to tell whether female Pradhans are more likely to be elected after reservation ends in villages that have been reserved twice in a row, since the second election cycle is still in process.²²

However, while this paper does not establish whether temporary reservations would increase female representation permanently, taken together, the weight of the evidence suggests that political reservations, or more generally quota systems, may play an important role in reducing bias against female leaders.

²²Furthermore, in many ways, this is not the ideal experiment to investigate this question, since the electoral rules prevent many existing Pradhans to run again (they are prevented to run if their seat as Pradhan or their seat as council member is reserved for category they don't belong to). Moreover, since the Pradhan election is indirect, re-election as Pradhan may reflect manipulations in the village council, rather than voters' preference for a particular Pradhan.

References

- Akerlof, George and Rachel Kranton, "Identity and Economics," *The Quarterly Journal of Economics*, 115 (2000), 715–154.
- Arrow, Kenneth, *The Theory of Discrimination*, Discrimination in Labor Markets (Princeton: Princeton University Press 1973) 3–33.
- Banaji, Mahzarin, "Implicit attitudes can be measured," in III J. S. Nairne (Eds.) H. L. Roediger, ed., "The nature of remembering: Essays in honor of Robert G. Crowder," (Washington, DC: American Psychological Association 2001).
- Baron, A and Mahzarin Banaji, "The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood," *Psychological Science*, 17 (2006), 53–58.
- Beaman, Lori, Raghav Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova, "Gender Bias and the Implicit Association Test: Evidence from Rural India," (2008), mimeo, Harvard and MIT.
- Becker, Gary, *The economics of discrimination* (Chicago: Chicago University Press 1957).
- Betrand, Marianne, Dolly Chugh, and Sendhil Mullainathan, "Implicit Discrimination," *American Economic Review*, 95 (2005), 94–98.
- B.Nosek, Anthony Greenwald, and Mahzarin Banaji, "The Implicit Association test at Age 7: A Methodological and Conceptual Review," in J. A. Bargh, ed., "Automatic Processes in Social Thinking and Behavior," (2007).
- Boisjoly, Johanne, Greg Duncan, Michael Kremer, Dan Levy, and Jacque Eccles, "Empathy or Antipathy? The Consequences of Racially and Socially Diverse Peers on Attitudes and Behaviors," *American Economic Review*, 96 (2006), 1890–1906.
- Chattapodhyay, Raghavendra and Esther Duflo, "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," *Econometrica*, 72 (2004), 1409–1443.

- Coate, Stephen and Glenn Loury, “Will Affirmative-Action Policies Eliminate Negative Stereotypes?” *American Economic Review*, 83 (1993), 1220–1240.
- Dahlerup, Drude, *Women, Quotas and Politics* (New York: Routledge 2006).
- Dasgupta, Nilanjana and Shaki Asgari, “Seeing is Believing: Exposure to Counterstereotypic Women Leaders and its Effect on the Malleability of Automatic Gender Stereotyping,” *Journal of Experimental Social Psychology*, 40 (2004), 642–658.
- Duflo, Esther, “Why Political Reservations,” *Journal of the European Economic Association*, 3 (2005), 668–678.
- Duflo, Esther and Petia Topalova, “Unappreciated Service: Performance, Perceptions, and Women Leaders in India,” *Mimeo*.
- Eagly, A.H. and S.J. Karau, “Role congruity theory of prejudice toward female leaders,” *Psychological Review*, 109 (2002), 573–598.
- Eagly, A.H. and A. Mladinic, “Gender stereotypes and attitudes toward women and men,” *Personality and Social Psychology Bulletin*, 15 (1989), 543–558.
- Frechette, Guillame, Francois Maniquet, and Massimo Morelli, “Incumbents’ Interests, Voters’ Bias and Gender Quotas,” (2006), cORE Discussion Paper No. 2006/83.
- Giles, Michael and Kaenan Hertz, “Racial Threat and Partisan Identification,” *American Political Science Review*, 88 (1994), 317–326.
- Goldin, Claudia, “A Pollution Theory of Discrimination,” (1990), mimeo, Harvard.
- Goldin, Claudia and Cecilia Rouse, “Orchestrating Impartiality: The Impact of ”Blind” Auditions on Female Musicians,” *American Economic Review*, 90 (2000), 715–741.
- Golebiowska, Ewa, “Group Stereotypes and Political Evaluations,” *American Politics Research*, 29 (2001), 535–65.
- Government of West Bengal, *The West Bengal Panchayat (Constitution) Rules, 1975*, Department of Panchayats & Rural Development (1998).

- Greenwald, A.G., M.R. Banaji, and B.A. Nosek, "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm," *Journal of Personality and Social Psychology*, 85 (2003), 197–216.
- Greenwald, A.G., D. McGhee, and J.L.K. Schwartz, "Measuring individual Differences in Implicit Cognition: The Implicit Association Test," *Journal of Personality and Social Psychology*, 75 (1998), 1464–1480.
- Huddy, Leonie and Nayda Terkildsen, "Gender Stereotypes and the Perception of Male and Female Candidates," *American Journal of Political Science*, 37 (1993), 119–147.
- Jones, Mark, "Quota Legislation and the Election of Women: Learning from the Costa Rican Experience," *Journal of Politics*, 66 (2004), 1203–1223.
- Kahn, Kim, "Does Gender Make a Difference? An Experimental Examination of Sex Stereotypes and Press Patterns in Statewide Campaigns," *American Journal of Political Science*, 38 (1994), 162–195.
- Kling, Jeffrey, Jeffrey Liebman, and Lawrence Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*.
- Matland, Richard E., "Putting Scandinavian Equality to the Test: An Experimental Evaluation of Gender Stereotyping of Political Candidates in a Sample of Norwegian Voters." *British Journal of Political Science*, 24 (1994), 273–292.
- Mendelberg, Tali, *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality* (Princeton: Princeton University Press 2001).
- Pande, Rohini, "Can Mandated Political Representation Provide Disadvantaged Minorities Policy Influence? Theory and Evidence from India," *American Economic Review*, 93 (2003), 1132–1151.
- Pettigrew, T.F. and L.R. Tropp, *Does Intergroup Contact Reduce Prejudice: Recent Meta-analytic Findings*, *Reducing Prejudice and Discrimination* (Mahwah, NJ: Lawrence Erlbaum Associates 2000) 93–114.

- Phelps, Edmund, "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62 (1972), 659–661.
- Powley, Elizabeth, "Rwanda: The Impact of Women Legislators on Policy Outcomes Affecting Children and Families," *Background Paper, State of the World's Children*.
- Rehavi, Marit, "Sex and Politics: Do Female Legislators Affect State Spending?" *mimeo, Berkeley*.
- Rooth, Dan-Olof, "Implicit Discrimination in Hiring: Real World Evidence," *IZA Discussion Paper*, 2764.
- Rudman, Laurie and Kimberly Fairchild, "Reactions to Counterstereotypic Behavior: The Role of Backlash in Cultural Stereotype Maintenance," *Journal of Personality and Social Psychology*, 87 (2004), 157–176.
- Rudman, Laurie and Stephen Kilianski, "Implicit and Explicit Attitudes Toward Female Authority," *Personality and Social Psychology Bulletin*, 26 (2000), 1315–1328.
- Sigelman, Carol, Lee Sigelman, Barbara Walkosz, and Michael Nitz, "Black Candidates, White Voters: Understanding Racial Bias in Political Perceptions," *American Journal of Political Science*, 39 (1995), 243–65.
- Thernstrom, S. and A. Thernstrom, *America In Black And White: One Nation, Indivisible* (New York, NY: Simon and Schuster 1997).
- UNICEF, *The State of The World's Children: Women and Children: The Double Dividend of Gender Equality* (New York: UN 2007).
- Valentino, Nicholas, Vincent Hutchings, and Ismail White, "Cues That Matter: How Political Ads Prime Racial Attitudes During Campaigns," *American Political Science Review*, 96 (2002), 75–90.
- Zaller, John, *The Nature and Origins of Mass Opinion* (Cambridge: Cambridge University Press 1992).

A Appendix

A.1 Sampling Procedure

Birbhum has 165 GPs. In each GP we selected three villages – two of these were the randomly selected two villages which were studied in Chattapodhyay and Duflo (2004). From the remaining villages in each GP we randomly selected one additional GP.

We obtained household listings for each of our sample villages from one of three census listings: the economic census list (undertaken by the Central Statistical Organization), the Total Sanitation Campaign household list and the tax assessment list prepared by the GP office. A village census was undertaken for constructing each of these two lists in 2005-06. We then used the circular systematic sampling technique - where the first household is identified using a randomly generated number and the rest at equal interval (the interval being $N/15$) - to draw a total of 15 households from the list for each village.

A.2 Variable description

A. Vignette

Note: There are four versions of the vignette - Male (Tapan Das), Female (Sandhya Das), Invest in Irrigation and Invest in Water.

READOUT: Now we will read you a short description of the Gram Pradhan of village CHANDI in district South 24 Parganas in West Bengal. We will ask you a couple of questions about what you think the Pradhan should have done. There are no right or wrong answers. Please answer each in terms of your own reactions.

Vignette

Pradhan Tapan Das [*Pradhan Sandhya Das*] has been serving his [*her*] Panchayat for ten months. As the end of the year approaches, there is only a limited amount of money remaining in the budget. Yet, villagers have been pressing him [*her*] to make improvements in two major areas: irrigation and drinking water. There was enough money to make investments in only one area. Prior to making a decision, Pradhan Tapan Das [*Pradhan Sandhya Das*] consulted with villagers at the Gram Sabha. Many people expressed frustration that there was still no safe drinking water available in the village.

Many people, especially children, were getting sick. Others were upset about the quality of the irrigation system. Poor irrigation system meant that, in dry years, many people lost their crops. Shri Tapan Das [*Shrimati Sandhya Das*] considered the demands carefully, and wondered what to do. On the one hand, Shri Tapan Das [*Shrimati Sandhya Das*] knew the health cost of bad water quality. Yet, wouldn't everyone be better off with better irrigation. After careful reflection Pradhan Tapan Das [*Pradhan Sandhya Das*] decided to invest in irrigation improvement [*drinking water*].

B. Speech

Note: There are two versions of the speech - a male voice recording and a female voice recording
READOUT: Now we will play a tape-recorded speech from the Gram Sabha meeting of Gram Panchayat Labhpur in district West Dinajpur in West Bengal. We will ask to rank the effectiveness of this speech on a scale from 1 to 10. There are no right or wrong answers. Please answer each in terms of your own reactions.

Recording

VILLAGER: The tube well of our Kumarpara is not functioning. The repairing job of the tubewell in your locality has been done partially, but the same work at Nutangram has been completed.

PRADHAN: For repairing of tubewells maximum amount of funds of the Panchayat is being drained out. As a result of which, other works can't be done. From the next stage you, the people, should take mental preparations that the minor repairing jobs of the tubewells won't be done by the Panchayat. I mean that if the work involves a large amount of money, e.g. if a pipe is needed then it involves the money above Rs.250, Rs.300, this type of works will be done by the Panchayat. But for the minor repairing jobs the people have to take initiative to collect subscriptions to do this. In the future, the plan of the Panchayat will be "plans with equal sharings" ("Samobhagi Parikalpana"). The Government won't provide all the money. The Government will provide some amount of money and the rest have to be borne by the people either by giving labor or helping financially. In this way the work of the Panchayat have to be done. Suppose a village road has to be constructed, then the people of the village will do the earthen work and the Panchayat will supply the morram. Therefore the people will now share the jobs, which the Panchayat did mostly. Then the total work can be made with a success. So in the next stage that preparation have to be taken. I would now like all villagers to approve the village budget.

B. Public Good Provision

The water and sanitation quantity variable represents the average effect of reservation on a dummy for whether a tubewell was built, a tubewell was repaired, a sanitation pit was built, a sanitation pit repaired. Water and Sanitation quality variables include: handpumps are perennial, provide clean water, no stagnant water, have drainage and sanitation – no stagnant water, drainage facility. The Irrigation variable represents the average effect of reservation on a dummy for whether an irrigation pump was built or repaired. The Roads quantity variable is a dummy for whether a metal road was built or repaired since 2003. The Roads quality variables include: condition of road (1-5) and number of potholes in 100m. The Transport quantity variable is the number of transportation related infrastructure (bus stop, bus service, taxi). The Transport quality variables are: whether there is a bus stand and if bus stand has shelter. The Schools and other education facilities quantity variable represents the average effect of reservation on a dummy for whether any educational facility was built, a dummy for whether such facility was repaired, a dummy for whether there is a creche and an indicator for a CE Center/CE Library. Educational facilities include: SSK, Anganwadi, primary schools, middle schools, libraries and secondary schools. The Schools and other education facilities quality variable represents the average effect of reservation on a dummy for whether all primary schools have drinking water, latrines, blackboards and reading and math test scores. Health: Quantity variable represents the average effect of reservation on a the number of health facilities, a dummy for whether a health facility was built, a dummy for whether a health facility was repaired (0 if no health facility existed), and number of doctors. Quality variable represents the average effect of reservation on an indicator for facility having tap or handpump water and an indicator for having a labor room. Fair Price Shop Quality measures include: whether prices displayed, no bad behavior of shop keeper, and no complaint against shop.

Table 1. Fraction Women Pradhans in Currently and Previously Reserved and Unreserved GP

	Total GP	First Reserved 2003	Reserved 1998 and 2003	Only Reserved 1998	Never Reserved
Fraction Female in 2003 election	0.434	0.982 ¹	1	0.139	0.162
Fraction Female in 1998 election	0.376	0.086	1	1	0.068
N	165	35	20	36	74

Notes:

1 One GP, Ayas, does not currently have an elected Pradhan.

2 "First Reserved 2003," "Reserved 1998 and 2003," "Only Reserved 1998," and "Never Reserved" are indicator variables for GPs reserved for a female pradhan for the first time in 2003, in both 1998 and 2003, only in 1998, and not reserved in either election, respectively.

Table 2. Comparison of Villages in 1991

Dependent Variable	First	Reserved	Only	Never	Diff: (1)	Diff: (3)
	Reserved	1998 and	Reserved			
	2003	2003	1998	Reserved	and (2)	and (4)
	(1)	(2)	(3)	(4)	(5)	(6)
Total Population	1,323 (1320)	1,311 (1434)	1,195 (1324)	1,369 (1569)	-136.048 (196.574)	19.418 (194.309)
Household Size	5.424 (0.621)	5.177 (0.429)	5.385 (0.575)	5.479 (2.390)	0.099 (0.115)	-0.012 (0.142)
Literacy	0.385 (0.119)	0.362 (0.108)	0.401 (0.119)	0.371 (0.136)	-0.019 (0.025)	0.018 (0.018)
Fraction of Women Literate	0.283 (0.125)	0.274 (0.106)	0.305 (0.111)	0.274 (0.133)	-0.033 (0.027)	0.021 (0.017)
Sex Ratio Under 6	1.093 (0.593)	1.028 (0.308)	1.054 (0.223)	1.036 (0.223)	0.183 (0.088)	0.026 (0.026)
Share of SC / ST Population	0.419 (0.279)	0.486 (0.244)	0.454 (0.226)	0.471 (0.267)	-0.007 (0.044)	-0.022 (0.033)
Percentage of Irrigated Land	0.583 (0.337)	0.579 (0.358)	0.617 (0.334)	0.536 (0.355)	-0.125 (0.064)	0.040 (0.049)
Village has a Bus or Train Stop	0.279 (0.451)	0.457 (0.504)	0.300 (0.461)	0.262 (0.441)	-0.189 (0.117)	0.054 (0.059)
Pucca Road to the Village	0.107 (0.310)	0.356 (0.484)	0.186 (0.391)	0.198 (0.400)	-0.296 (0.091)	-0.004 (0.055)
Number of Health Facilities	0.183 (0.413)	0.196 (0.453)	0.137 (0.346)	0.200 (0.778)	0.014 (0.095)	0.010 (0.066)
Village has Tube Well	0.905 (0.294)	0.929 (0.261)	0.904 (0.296)	0.963 (0.190)	-0.027 (0.025)	-0.010 (0.022)
Village has Hand Pump	0.119 (0.325)	0.000 (0.000)	0.101 (0.303)	0.054 (0.226)	0.000 (0.000)	0.014 (0.012)
Village has Well	0.427 (0.497)	0.543 (0.504)	0.426 (0.497)	0.475 (0.501)	-0.003 (0.089)	-0.014 (0.075)
Village has Community Tap	0.063 (0.243)	0.065 (0.250)	0.043 (0.203)	0.018 (0.134)	0.027 (0.025)	-0.029 (0.021)
Total Number of Schools	1.192 (0.893)	1.261 (0.880)	1.314 (0.965)	1.147 (0.897)	-0.133 (0.192)	0.213 (0.119)

Notes:

- 1 Standard deviations, below the means, in columns (1)-(4).
- 2 Tests in columns (5)-(6) are based on regressions with block FE and standard errors clustered by GP.
- 3 Source of data is 1991 Census of India.
- 4 Reservation variables are defined in the notes of Table 1.

Table 3: Implicit Association Test Measure of Implicit Bias and Ladder

	Ladder: Female Pradhan		Male/Female Names and Good/Bad IAT		Male/Female Politician and Good/Bad IAT		Leadership/Domestic and Male/Female IAT	
	Relative to Male Pradhan		Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Ever Reserved	-0.231 (0.176)	-0.081 (0.161)	-0.001 (0.032)	0.006 (0.042)	-0.008 (0.034)	-0.015 (0.037)	-0.070 (0.030)	0.022 (0.041)
Panel B								
First Reserved 2003	-0.257 (0.262)	-0.126 (0.205)	-0.039 (0.042)	0.020 (0.051)	-0.005 (0.049)	0.010 (0.049)	-0.089 (0.040)	0.104 (0.053)
Reserved 1998 and 2003	-0.054 (0.218)	-0.016 (0.205)	0.039 (0.041)	0.044 (0.068)	0.004 (0.052)	-0.008 (0.052)	-0.024 (0.045)	-0.079 (0.067)
Only Reserved 1998	-0.306 (0.217)	-0.091 (0.212)	0.011 (0.047)	-0.048 (0.051)	-0.020 (0.044)	-0.043 (0.051)	-0.080 (0.039)	-0.021 (0.050)
Test: 2003 = both 1998 and 2003 = 1998	0.578	0.885	0.301	0.299	0.908	0.636	0.390	0.032
Mean of Never Reserved Sample	-0.890 (0.062)	-0.354 (0.059)	0.134 (0.025)	-0.157 (0.026)	0.093 (0.027)	-0.079 (0.025)	0.110 (0.021)	0.150 (0.027)
N	3511	3669	510	408	554	510	477	357

Notes:

- 1 The dependent variables in Columns (1)-(2) are Feeling Towards Leader is the response on a scale of 1-10 of the question "Please show us how you feel towards a x Pradhan" where x is either Male or Female. The specification includes individual fixed effects and the Mean of the Never Reserved Sample reflects the average relative ranking of Female Pradhans to Male Pradhans. Standard errors are clustered by GP.
- 2 In columns (3)-(8), trials with a latency greater than 10000 ms and any respondent with either an average response time less than 6000ms for the first test block or an average percent correct <65% for either test block are removed.
- 3 Columns (3)-(8) include block fixed effects and individual controls. The individual controls are age, age squared, household size, religion, caste, education and sex of respondent, wealth (pca), landholdings, and an indicator for survey period. Standard errors are clustered by GP.
- 4 Dependent variables in columns (3)-(8) are the IAT D measure, defined as the difference in average response latencies between the two test blocks divided by the standard deviation of latencies in the two blocks. All stimuli are described in Data section.
- 5 Ever Reserved is an indicator for whether a GP was reserved for a female Pradhan in either 1998, 2003 or in both elections. All other reservation variables are as defined in Table 1.
- 6 The p-value from a Wald test of the equality of the coefficients on First Reserved in 2003, Reserved 1998 and 2003 and Only Reserved 1998 is reported.
- 7 The sample size in columns (3)-(8) reflect that we administered the IAT to approximately 5 households within each village.

Table 4. Female Leadership and Prejudice: Speech and Vignettes

	Cares about villagers'											
	Average Effect		Pradhan is Effective		welfare		Approve of Pradhan		Perform Duties Well		Agree with Pradhan	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A												
Female Pradhan	-0.074	-0.049	-0.066	-0.051	-0.072	-0.027	-0.056	-0.084	-0.073	-0.002	-0.143	-0.088
	(0.029)	(0.034)	(0.034)	(0.041)	(0.035)	(0.038)	(0.036)	(0.041)	(0.038)	(0.042)	(0.054)	(0.055)
Female Pradhan * Ever Reserved	0.122	0.002	0.131	0.014	0.158	-0.016	0.073	0.020	0.111	-0.011	0.122	-0.044
	(0.040)	(0.042)	(0.046)	(0.050)	(0.047)	(0.047)	(0.048)	(0.050)	(0.052)	(0.053)	(0.072)	(0.073)
Test: Female Pradhan + Female Pradhan * Ever Reserved	0.048	-0.047	0.065	-0.037	0.085	-0.044	0.016	-0.063	0.038	-0.013	-0.022	-0.133
	(0.027)	(0.025)	(0.032)	(0.030)	(0.032)	(0.029)	(0.032)	(0.029)	(0.035)	(0.031)	(0.048)	(0.048)
Panel B												
Female Pradhan * First Reserved 2003	0.142	-0.030	0.152	-0.022	0.192	-0.059	0.042	0.022	0.171	-0.074	0.072	-0.014
	(0.048)	(0.051)	(0.058)	(0.057)	(0.059)	(0.053)	(0.054)	(0.060)	(0.064)	(0.066)	(0.086)	(0.090)
Female Pradhan * Reserved 1998 & 2003	0.129	0.045	0.143	0.038	0.166	0.034	0.105	0.086	0.071	0.035	0.172	-0.018
	(0.066)	(0.064)	(0.066)	(0.078)	(0.078)	(0.079)	(0.070)	(0.067)	(0.092)	(0.082)	(0.107)	(0.095)
Female Pradhan * Only Reserved 1998	0.104	0.006	0.108	0.030	0.125	-0.011	0.086	-0.026	0.078	0.030	0.147	-0.115
	(0.056)	(0.053)	(0.065)	(0.068)	(0.065)	(0.063)	(0.071)	(0.065)	(0.071)	(0.062)	(0.103)	(0.105)
Test: FP* 2003 = FP* (both 1998 and 2003) = FP* 1998	0.824	0.502	0.821	0.627	0.646	0.442	0.638	0.319	0.406	0.248	0.490	0.376

Notes:

- 1 Average evaluation coefficients in columns (1)-(2) are from a seemingly unrelated regression (SUR) which includes dependent variables listed in columns (3)-(10) and Appendix Table 1.
- 2 All regressions include individual controls as defined in Table 3, and standard errors are clustered by GP. The regressions also include: in Panel A, Ever Reserved and in Panel B, First Reserved 2003, Reserved 1998 or 2003, and Only Reserved 1998 (see Table 1 and Table 3 notes for definitions). Female leader is an indicator for when the voice delivering the speech was female or if the female Pradhan described in the vignette was female.
- 3 Dependent variables in columns (4)-(7) and (10)-(11) are originally on a scale of 1 to 10, while dependent variables in columns (8)-(9) and (12)-(13) are indicator variables, then normalized by the mean and standard deviation of the never reserved sample. Columns (4)-(11) show the average evaluation coefficients are from a seemingly unrelated regression (SUR) which includes dependent variables listed individually in Appendix Table 1. The dependent variable in columns (12)-(13) is Agree with Pradhan in the Vignette experiment which equals one if the villager agrees with statement and then normalized by the mean and standard deviation of the never reserved sample.

Table 5. Female Leadership and Pradhan Evaluation

	Average Effect		Pradhan is effective		Pradhan did a good job					
	Male	Female	Male	Female	looking after village needs		looking after your needs		making BPL lists	
					Male	Female	Male	Female	Male	Female
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
Female Pradhan (GP reserved only in 2003)	-0.212 (0.058)	-0.104 (0.061)	-0.208 (0.065)	-0.105 (0.071)	-0.212 (0.061)	-0.126 (0.064)	-0.221 (0.063)	-0.062 (0.067)	-0.206 (0.066)	-0.122 (0.060)
Female Pradhan (GP reserved 1998 and 2003)	-0.009 (0.061)	0.015 (0.053)	0.012 (0.068)	-0.015 (0.064)	0.005 (0.069)	-0.025 (0.061)	0.003 (0.062)	0.068 (0.058)	-0.056 (0.067)	0.031 (0.054)
Unreserved (GP previously reserved 1998)	0.020 (0.081)	0.015 (0.054)	-0.022 (0.087)	-0.007 (0.063)	-0.027 (0.083)	-0.019 (0.064)	0.014 (0.078)	0.049 (0.056)	0.115 (0.092)	0.035 (0.056)
Test: 2003 = both 1998 and 2003 = 1998	0.004	0.115	0.018	0.386	0.014	0.253	0.004	0.157	0.005	0.024
Test: 2003 = both 1998 and 2003	0.008	0.081	0.052	0.220	0.044	0.166	0.008	0.128	0.002	0.025
N			6530	6315	6590	6363	6471	6323	6246	5893

Notes:

- 1 Average evaluation coefficients are the result of a seemingly unrelated regression (SUR) including the four dependent variables listed in columns 3 through 10.
- 2 All regressions include individual controls as defined in Table 3, and standard errors are clustered by GP.
- 3 All dependent variables are originally on a scale of 1 to 10, and then normalized by the mean and standard deviation of the never reserved sample.
- 4 The p-values from Wald tests of the equality of the coefficients on First Reserved in 2003, Reserved 1998 and 2003 and Only Reserved 1998; and First Reserved in 2003 and Reserved in 1998 and 2003 are reported.
- 5 Reservation variables are as defined in Table 1.

Table 6. Pradhan Performance: Public Goods, Bribes and Satisfaction

	Average Public Good Provision		Average Bribes	Average Satisfaction		Alignment with Female Preferences
	Quantity (1)	Quality (2)	(3)	Male (5)	Female (6)	(7)
Female Pradhan (GP reserved 2003 only)	0.211 (0.069)	-0.020 (0.042)	-0.072 (0.033)	0.068 (0.053)	0.004 (0.056)	0.521 (0.279)
Female Pradhan (GP reserved 2003 and 1998)	0.096 (0.080)	0.002 (0.044)	-0.073 (0.029)	-0.024 (0.069)	0.065 (0.068)	0.659 (0.358)
Unreserved (GP previously reserved 1998)	0.035 (0.060)	-0.032 (0.040)	-0.026 (0.039)	-0.030 (0.067)	0.075 (0.057)	0.563 (0.243)
Test: 2003 = both 1998 and 2003 = 1998	0.051	0.808	0.405	0.231	0.387	0.942
Test: 2003 = both 1998 and 2003	0.240	0.681	0.983	0.241	0.394	0.730

Notes

- ¹ Average coefficients are from seemingly unrelated regressions (SUR). Appendix tables A2, A3 and A4 provide results for each variable in the SUR, and data appendix provides variable descriptions. All dependent variables are originally on a 1-10 scale, and then normalized by the mean and standard deviation of the never reserved sample.
- ² The sample in Columns (1) and (2) regressions are 495 villages, while columns (3)-(6) regressions use the villager sample. All regressions include block fixed effects, and standard errors are clustered by GP. Villager sample regressions individual controls as defined in Table 3.
- ³ The p-values from Wald tests of the equality of the coefficients on First Reserved in 2003, Reserved 1998 and 2003 and Only Reserved 1998; and First Reserved in 2003 and Reserved in 1998 and 2003 are reported.
- ⁴ Column 7 tests whether there is more investment in reserved GPs in goods mentioned more frequently by women, as measured by formal complaints to the GP during 6 months in 2000 (Chattopadhyay and Duflo, 2004). Column 7 shows the coefficients from the interaction of reservation status and the average difference between fraction of requests for good *i* from women and from men. Goods analyzed include: drinking water, road improvement, education, irrigation, and other.
- ⁵ Reservation variables are as defined in Table 1.

Table 7. Pradhan Characteristics

Dependent Variable	Coefficients on:				
	First Reserved	Reserved 1998 and 2003	Only Reserved 1998	Mean of Never Reserved	Diff: (1) and (2)
	(1)	(2)	(3)	(4)	(5)
Age	-5.780 (1.890)	-6.692 (2.305)	0.718 (2.112)	41.194 (1.146)	0.721
Female	0.844 (0.047)	0.775 (0.061)	-0.025 (0.073)	0.181 (0.046)	0.158
Educational Level	-2.329 (0.719)	-1.209 (0.792)	0.550 (0.641)	10.278 (0.377)	0.229
Married	-0.184 (0.089)	0.018 (0.084)	-0.026 (0.064)	0.889 (0.037)	0.085
Number of Children	-0.367 (0.263)	0.003 (0.315)	-0.383 (0.261)	2.167 (0.165)	0.327
SC	0.101 (0.105)	0.025 (0.131)	0.053 (0.105)	0.389 (0.058)	0.615
ST	-0.054 (0.058)	-0.048 (0.087)	-0.027 (0.063)	0.125 (0.039)	0.941
Muslim	-0.058 (0.079)	-0.064 (0.093)	-0.047 (0.086)	0.222 (0.049)	0.958
Landless	0.068 (0.100)	0.023 (0.118)	-0.076 (0.083)	0.236 (0.050)	0.758
Wealth Index: Quartile 1	0.170 (0.100)	0.089 (0.121)	-0.073 (0.086)	0.208 (0.048)	0.574
Wealth Index: Quartile 2	-0.056 (0.097)	-0.183 (0.103)	0.099 (0.099)	0.319 (0.055)	0.312
Wealth Index: Quartile 3	-0.001 (0.099)	0.075 (0.114)	-0.057 (0.093)	0.264 (0.052)	0.569
GP Experience	-0.431 (0.162)	-0.085 (0.215)	0.110 (0.270)	1.611 (0.109)	0.129
Other Political Experience	-0.195 (0.091)	-0.081 (0.130)	-0.104 (0.109)	0.417 (0.059)	0.416
Average Experience Effect	-0.419 (0.136)	-0.126 (0.183)	-0.051 (0.190)	0.086 (0.094)	0.146
Spouse Ever Elected to Panchayat	0.006 (0.009)	0.047 (0.054)	0.032 (0.029)	0 (0)	0.438
Previous Pradhan was Re-elected in 2003	-0.077 (0.040)	0.097 (0.086)	-0.047 (0.047)	0.069 (0.030)	0.040
Previous Pradhan Ran for Re-election in 2003	-0.067 (0.083)	0.095 (0.116)	0.118 (0.095)	0.221 (0.051)	0.211

Notes:

- ¹ Columns 1 through 3 are the coefficients on indicator variables for each reservation status type. Column 4 is the constant, representing the average for that variable in never reserved GPs. All regressions include are block fixed effects and standard errors are clustered at the GP level.
- ² Average characteristics coefficients are the result of a seemingly unrelated regression where the dependent variables are normalized by the mean and standard deviation of the never reserved sample.
- ³ Reservation types are defined in the notes of Table 1.
- ⁴ Wealth Index is based on a principal components analysis using the number of household assets.
- ⁶ GP experience is the number of times Pradhan was elected to the GP.
- Other Political Experience is an indicator for whether the Pradhan has or holds other political office, including: booth member,
- ⁷ member of higher panchayat, MP, MLA, youth party president, district / block level posts of party, Mahila Samiti.
- ⁸ Average experience effect includes both GP Experience and Has or Holds Political Office.

Table 8. Robustness

	Average effect	
	Male	Female
	(1)	(2)
I. Evaluation of New Pradhans		
New Pradhan in 2003 or after	0.016 (0.082)	0.028 (0.072)
II. Evaluation of SC Pradhans		
GP Reserved for SC in 2003	-0.184 (0.066)	-0.085 (0.057)
GP Reserved for SC in 1998	-0.143 (0.056)	-0.061 (0.053)
GP Reserved for SC in 2003 * Reserved for Woman in 1998	0.102 (0.137)	-0.037 (0.112)
III. Evaluation of Female Pradhans and SC Reservation		
First Reserved 2003 for Woman	-0.223 (0.077)	-0.098 (0.080)
Reserved 1998 and 2003 for Woman	0.034 (0.082)	-0.026 (0.071)
Only Reserved 1998 for Woman	0.055 (0.098)	0.037 (0.070)
Reserved for Female in 2003 * Previously Reserved for SC	0.044 (0.135)	-0.020 (0.132)
IV. SC Reservation on Female Leadership: Speech and Irrigation Vignette Results		
Panel A		
Female Pradhan	-0.069 (0.038)	-0.060 (0.038)
Female Pradhan * Ever Reserved SC	-0.008 (0.041)	0.024 (0.044)
Panel B		
Female Pradhan * Reserved for SC in 2003	0.026 (0.049)	0.027 (0.048)
Female Pradhan * Reserved for SC in 1998	-0.034 (0.048)	0.036 (0.053)

Notes

- 1 All regressions include individual controls as defined in Table 3, and standard errors are clustered by GP.
- 2 Average evaluation coefficients are the result of a seemingly unrelated regression (SUR) including the four dependent variables listed in columns (3)-(10) of Table 7 in I, II and III, and the seven dependent variables listed in columns (1)-(7) of Table A1 in IV.
- 3 All regressions include individual controls as defined in Table 3, and standard errors are clustered by GP.
- 4 Sample in Panel A contains only never reserved GPs.
- 5 Panel III also includes Previously Reserved for SC, Reserved for Female in 1998*Previously Reserved for SC, and Reserved for Female in 1998 and 2003*Previously Reserved for SC.
- 6 Reservation types are defined in the notes of Table 1.

Table 9: Pradhan Opinions

	Gender Caused		
	Problem to do Duties	Satisfied with Current Life	Plan to re-run for Pradhan
	(1)	(2)	(3)
First Reserved 2003	0.412 (0.086)	-0.470 (0.217)	-0.154 (0.102)
Reserved 1998 and 2003	0.100 (0.069)	0.379 (0.262)	-0.058 (0.123)
Only Reserved 1998		-0.148 (0.216)	-0.025 (0.101)
Test: 2003 = both 1998 and 2003	0.039	0.006	0.502
Mean of Never Reserved Sample		3.708 (0.113)	0.708 (0.054)
N	54	161	161

Notes

- ¹ Gender Caused Problem to do Duties an indicator variable. Satisfied with Current Life is on a scale of 1 to 5. Plan to re-run for Pradhan is also an indicator.
- ² Columns 2 and 3 include block fixed effects.
- ³ Reservation variables are as defined in Table 1.
- ⁴ In Column 1, First Reserved 2003 and Reserved 1998 and 2003 are the mean of the variable for those GPs.
- ⁵ The p-value from a Wald test of the equality of the coefficients on First Reserved in 2003 and Reserved 1998 and 2003 is reported.

Table A1. Female Leadership and Prejudice -- Speech and Vignette Results

	Speech							Irrigation Vignette			
	Pradhan is effective	Addressed villager satisfactorily	Cares about villagers' welfare	Will allocate BPL cards well	Villager approves pradhan's budget	Pradhan will get resources by lobbying	Pradhan will collect villagers' share well	Pradhan is Effective	Cares about villagers' welfare	Agree with Pradhan	Would vote for Pradhan
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
I. Males											
Panel A											
Female Pradhan	-0.067 (0.040)	-0.111 (0.041)	-0.082 (0.042)	-0.034 (0.042)	-0.034 (0.042)	-0.065 (0.039)	-0.070 (0.042)	-0.065 (0.054)	-0.062 (0.054)	-0.143 (0.054)	-0.078 (0.060)
Female Pradhan * Ever Reserved	0.096 (0.054)	0.115 (0.055)	0.132 (0.055)	0.106 (0.057)	0.046 (0.057)	0.153 (0.053)	0.125 (0.053)	0.166 (0.072)	0.183 (0.075)	0.122 (0.072)	0.099 (0.077)
Panel B											
Female Pradhan * First Reserved 2003	0.156 (0.069)	0.143 (0.067)	0.204 (0.068)	0.199 (0.071)	0.049 (0.068)	0.193 (0.071)	0.182 (0.070)	0.148 (0.095)	0.180 (0.096)	0.072 (0.086)	0.036 (0.092)
Female Pradhan * Reserved 1998 & 2003	0.101 (0.080)	0.072 (0.092)	0.099 (0.088)	0.071 (0.099)	0.062 (0.099)	0.133 (0.097)	0.137 (0.086)	0.185 (0.093)	0.232 (0.112)	0.172 (0.107)	0.149 (0.106)
Female Pradhan * Only Reserved 1998	0.036 (0.071)	0.114 (0.074)	0.082 (0.071)	0.043 (0.077)	0.038 (0.076)	0.131 (0.067)	0.074 (0.067)	0.179 (0.101)	0.167 (0.105)	0.147 (0.103)	0.134 (0.108)
Test: FP* 2003 = FP* (both 1998 and 2003) = FP* 1998	0.317	0.763	0.258	0.175	0.974	0.725	0.374	0.786	0.915	0.490	0.382
N	6715	6716	6715	6714	6551	6715	6716	3173	3173	3172	3116
II. Females											
Panel A											
Female Pradhan	-0.050 (0.045)	-0.027 (0.047)	-0.051 (0.046)	0.022 (0.044)	-0.098 (0.046)	-0.069 (0.045)	-0.053 (0.043)	-0.053 (0.058)	-0.003 (0.055)	-0.088 (0.055)	-0.069 (0.058)
Female Pradhan * Ever Reserved	0.002 (0.055)	0.021 (0.057)	0.010 (0.057)	-0.043 (0.056)	0.073 (0.058)	0.038 (0.058)	0.018 (0.057)	0.026 (0.072)	-0.043 (0.070)	-0.044 (0.073)	-0.033 (0.076)
Panel B											
Female Pradhan * First Reserved 2003	-0.043 (0.069)	-0.036 (0.075)	-0.044 (0.069)	-0.111 (0.068)	0.075 (0.079)	-0.019 (0.072)	-0.035 (0.077)	-0.002 (0.083)	-0.075 (0.082)	-0.014 (0.090)	-0.031 (0.087)
Female Pradhan * Reserved 1998 & 2003	0.022 (0.077)	0.078 (0.081)	0.061 (0.088)	-0.007 (0.096)	0.075 (0.078)	0.073 (0.099)	0.048 (0.081)	0.053 (0.101)	0.006 (0.105)	-0.018 (0.095)	0.098 (0.105)
Female Pradhan * Only Reserved 1998	0.042 (0.067)	0.052 (0.066)	0.042 (0.069)	0.008 (0.067)	0.079 (0.070)	0.081 (0.067)	0.060 (0.067)	0.017 (0.102)	-0.064 (0.095)	-0.115 (0.105)	-0.131 (0.108)
Test: FP* 2003 = FP* (both 1998 and 2003) = FP* 1998	0.483	0.360	0.386	0.242	0.998	0.392	0.473	0.851	0.907	0.376	0.369
N	6771	6772	6771	6770	6372	6770	6770	2869	2869	2869	2740

Notes

¹ 1 Dependent variables in columns (1)-(4) and (6)-(9) are originally on a scale of 1 to 10, while dependent variables in columns (5) and (10)-(11) are indicator variables, then normalized by the mean and standard deviation of the never reserved sample. In the vignette experiment, Effectiveness of leader and cares about villager welfare are originally on a scale of 1 to 10, while Agree with Leader and would vote for Leader are indicator variables which equal one if the villager agrees with statement, then normalized by the mean and standard deviation of the never reserved sample. In the speech experiment, Effectiveness of leader and cares about villager welfare are originally on a scale of 1 to 10, while Agree with Leader and would vote for Leader are indicator variables which equal one if the villager agrees with statement, then normalized by the mean and standard deviation of the never reserved sample.

² 2 All regressions include individual controls as defined in Table 3, and standard errors are clustered by GP. The regressions also include: in Panel A, Ever Reserved and in Panel B, First Reserved 2003, Reserved 1998 or 2003, and Only Reserved 1998 (see Table 1 and Table 3 notes for definitions). Female leader is an indicator for when the voice delivering the speech was female or the Pradhan in the vignette was female.

³ 3 Female Pradhan is a dummy variable indicating that the voice delivering the speech was female or the Pradhan depicted in the vignette was female.

Table A2. Female Leadership and Performance: Satisfaction and Bribes

	Average Effect: Water and Sanitation		Average Effect: Transport		Average Effect: Schools & Other educational facilities		Average Effect: Public Health Facilities		Average Effect: Fair Price Shop/BPL	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Bribe										
First Reserved 2003	-0.040									-0.009
	(0.031)									(0.004)
Reserved 1998 and 2003	-0.043									-0.008
	(0.028)									(0.003)
Only Reserved 1998	-0.006									-0.005
	(0.037)									(0.003)
Test: 2003 = both 1998 and 2003 = 1998	0.556									0.676
Test: 2003 = both 1998 and 2003	0.926									0.869
N	7404									7404
Mean of Never Reserved Sample	0.233									0.015
	(0.009)									(0.002)
Panel B: Satisfaction										
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
First Reserved 2003	0.010	0.012	0.061	0.016	0.007	-0.006	-0.004	-0.008	0.004	-0.003
	(0.015)	(0.012)	(0.048)	(0.048)	(0.009)	(0.012)	(0.010)	(0.010)	(0.020)	(0.018)
Reserved 1998 and 2003	-0.017	-0.012	0.017	0.084	0.010	0.011	-0.014	-0.005	-0.020	0.001
	(0.018)	(0.012)	(0.074)	(0.064)	(0.007)	(0.011)	(0.016)	(0.011)	(0.022)	(0.020)
Only Reserved 1998	-0.007	-0.008	-0.086	0.058	0.015	0.009	-0.001	0.003	0.029	0.023
	(0.016)	(0.014)	(0.077)	(0.055)	(0.009)	(0.009)	(0.010)	(0.009)	(0.017)	(0.017)
Test: 2003 = both 1998 and 2003 = 1998	0.353	0.157	0.117	0.458	0.703	0.427	0.745	0.554	0.129	0.391
Test: 2003 = both 1998 and 2003	0.194	0.107	0.582	0.283	0.759	0.247	0.600	0.835	0.383	0.863
N	6717	6780	695	725	3746	4016	5702	6047	3214	3397

Notes

¹ Average effect coefficients are the result of a seemingly unrelated regression (SUR) using dependent variables normalized by the mean and standard deviation of the never reserved sample. Details on the variables which are included in each SUR are available in the data appendix.

² All regressions in Panel B include individual controls as defined in Table 3. Regressions in Panel A include household size, religion, caste, education and sex of respondent, wealth (pca), landholdings and block fixed effects. All standard errors are clustered by GP.

³ Independent variables are defined in Table 1.

Panel A show regressions which are at the household level. The dependent variables in Columns 1 and 9 are indicators for whether either adult respondent paid a bribe for a

⁴ water connection or a BPL card respectively.

Table A3. Effect of Female Leadership on Public Goods Quantity

	N	Coefficients on:			Mean of Never Reserved
		GP reserved only in 2003	GP previously reserved in 1998	GP reserved in 2003 and 1998	
At Least One New Tubewell was Built	495	0.152 (0.066)	0.073 (0.063)	0.160 (0.088)	0.365 (0.482)
At Least One Tubewell was Repaired	482	0.208 (0.067)	0.130 (0.064)	0.080 (0.089)	0.628 (0.484)
At Least One Drainage/Sanitation Facility was Built	495	0.053 (0.067)	-0.113 (0.059)	0.052 (0.091)	0.428 (0.496)
At Least One Drainage/Sanitation Facility was Repaired	396	0.150 (0.067)	-0.017 (0.062)	0.032 (0.071)	0.178 (0.384)
At Least One Irrigation Pump was Built	495	0.137 (0.053)	0.005 (0.051)	-0.013 (0.050)	0.180 (0.385)
At Least One Irrigation Pump was Repaired	319	0.110 (0.092)	-0.078 (0.086)	-0.005 (0.123)	0.417 (0.495)
Number of metal roads built or repaired since 2003	495	0.274 (0.117)	0.046 (0.070)	0.079 (0.065)	0.118 (0.448)
Number of transportation related infrastructure (Bus Stop, bus service, taxi)	495	0.074 (0.175)	0.250 (0.160)	0.303 (0.225)	1.302 (1.201)
At Least One Educational Facility was Built	495	0.053 (0.042)	-0.030 (0.036)	0.026 (0.055)	0.117 (0.322)
At Least One Educational Facility was Repaired	465	0.165 (0.072)	0.039 (0.069)	0.001 (0.097)	0.296 (0.458)
At Least One Community Education Center	495	-0.007 (0.010)	0.030 (0.023)	-0.001 (0.009)	0.009 (0.095)
There is a NGO Child center/Creche	495	-0.045 (0.016)	-0.039 (0.021)	-0.027 (0.023)	0.045 (0.208)
Number of Health Facilities (PHC, Health SubCenter)	495	-0.025 (0.049)	0.027 (0.052)	-0.005 (0.084)	0.257 (0.468)
At Least One Health Facility was Built	495	0.011 (0.015)	-0.004 (0.014)	-0.018 (0.009)	0.014 (0.116)
At Least One Health Facility was Repaired (0 if no fac)	495	0.061 (0.023)	0.016 (0.016)	0.047 (0.024)	0.009 (0.095)
Number of Trained Dais, Untrained Dais and Private Doctors	495	-0.069 (0.232)	-0.158 (0.226)	0.384 (0.423)	1.014 (2.012)
Average Effect		0.211 (0.069)	0.035 (0.060)	0.096 (0.080)	

Notes

1 All regressions include block fixed effects, and standard errors are clustered by GP.

2 Reservation status defined in Table 1.

Average effect coefficients are the result of a seemingly unrelated regression using dependent variables normalized by the mean and standard deviation of the never reserved sample.