Powerful Women: Does Exposure Reduce Bias?

Lori Beaman, Raghabendra Chattopadhyay, Esther Duflo,

Rohini Pande and Petia Topalova[*]

Abstract

We exploit random assignment of gender quotas across Indian village councils to investigate whether having a female chief councillor affects public opinion towards female leaders. Villagers who have never been required to have a female leader prefer male leaders and, for the same stated performance, perceive hypothetical female leaders as less effective than their male counterparts. Exposure to a female leader does not alter villagers' stated distaste for female leaders. However, it significantly weakens stereotypes about gender roles in the public and domestic spheres and eliminates biased perceptions of female leaders' effectiveness, at least among male villagers. Female villagers exhibit less prior bias, but are also less likely to know about or participate in local politics; as a result their attitudes are largely unaffected. Consistent with our experimental findings, villagers rate first-time, but not second-time, women leaders as less effective than male leaders.

1 Introduction

In July 2006, women accounted for 17 percent of parliamentarians world-wide and a woman headed the government in only seven countries (UNICEF, 2007). These gender disparities do not reflect legal restrictions, however – women can vote, support candidates and run for office in almost every country. The possibility that the reason for this disparity lies in voter and party discrimination has, over the last two decades, led to more than a hundred countries introducing

affirmative action policies for women in public office, either by law or through voluntary actions of political parties (Krook (2005), Dahlerup (2006)).[1] While these policies have significantly increased female representation in politics (Jones, 2004), and often altered subsequent policy-making (Chattapadhyay and Duflo (2004), Powley (2007)), little is known about their impact on voter attitudes towards female leaders.

A large literature on public opinion formation suggests that low voter awareness on specific issues can cause group stereotypes to play an important role in shaping voter preferences (see, for instance, Zaller (1992), Huddy and Terkildsen (1993) and Mendelberg (2001)). A widely discussed reason for why group stereotypes may hinder the entry of women into politics is social norms of prescribed behavior which associate leadership activities with men (Akerlof and Kranton (2000), Eagly and Karau (2002) and Ingelhart and Norris (2003)). Holding leader effectiveness across genders constant, such norms can cause voters to favor, and elect, male candidates. Limited exposure to female leaders may, in turn, fuel biased perceptions of female leader effectiveness.

While, in many settings, exposure to members of another group creates "empathy"[2], the policy discourse and empirical evidence on whether mandated exposure to women leaders alters either social norms or perceptions of women's ability to lead remain mixed. Using international survey data Ingelhart and Norris (2003) document deep seated attitudes towards sex roles in public life across countries. They argue that these attitudes are difficult to change in the short run and quotas, by violating voters' sense of identity, may precipitate a backlash (also see Rudman and Fairchild (2004)). For instance, women's entry into politics may be perceived as reducing the value of a traditionally male activity (Goldin, 1990). This backlash may be accentuated if affirmative action breeds voter resentment for having their choice curtailed (Thernstrom and Thernstrom (1997) express similar concerns in the context of racial affirmative action).

However, independent of their impact on voter tastes, quotas, by exposing voters to a female leader, can improve the precision of voters' information about the expected effectiveness of future

---

[1]Details of quotas by country are available at http://www.quotaproject.org. We should note that another rationale for affirmative action is gender-specific policy preferences. If male and female leaders have different policy preferences (Chattapadhyay and Duflo (2004), Rehavi (2008)), and female turnout is relatively low, then affirmative action can help ensure equal political voice for men and women.

[2]For example Boisjoly et al. (2006) show that students who were randomly assigned to an African-American roommate are more likely to sympathize with African-Americans and affirmative action.

female leaders and therefore reduce statistical discrimination. Unless women make incompetent leaders (causing voters to update negatively) mandated exposure can improve the perception of women leaders' effectiveness, simply by reducing the risk associated with any member of an unknown group.[3]

This paper seeks to provide direct evidence on these issues. We first develop a simple voting model to show how, in an electoral setting, future male candidates may systematically benefit from an initial voter preference for male leaders. Specifically, greater exposure to male leaders improves voter ability to screen male candidates and, thereby, causes the expected effectiveness of male candidates to exceed that of their female counterparts. If voters are risk averse, then statistical discrimination against female leaders will reinforce the initial taste bias and cause the stereotype that women make ineffective leaders to persist. Political reservation, by increasing voter exposure to female leaders, can reduce statistical discrimination even when voter tastes are unaffected. We test these hypotheses using detailed survey and experimental data on Indian villagers' attitudes towards male and female village council leaders.

In general, measuring how exposure changes voters' attitudes vis-a-vis women leaders is difficult. Less biased voters are presumably also more likely to elect, and be exposed to, female leaders.[4] To address this concern, our empirical strategy exploits random variation in mandated exposure to female leaders across village councils in the Indian state of West Bengal. Since 1998, one third of village council leader positions in this state have been randomly set aside (or "reserved") for women at each election; in reserved councils only women can run for the position of leader. In our analysis, we compare villager attitudes towards hypothetical women leaders, and their actual leaders, across councils which have been reserved for women once, twice, or never. Random allocation of reservation implies that difference in voter attitudes across reserved and unreserved villages captures the causal effect of mandated exposure. In addition, the passage of a significant period of time since the introduction of political reservation (two electoral cycles, or ten years) makes it a good place to study the impact of exposure to female leaders on voter attitudes.

---

[3]One reason why quotas could lead to less competent female leaders is if, as in Coate and Loury (1993), they work less in anticipation of not being rewarded as much for their effort.

[4]This is true even with affirmative action, since in most countries, even when a law requires mandated representation of women, parties can still decide where to field women candidates (Frechette et al., ming).

An important innovation of our study is the collection, and use, of detailed survey and experimental data on voters' taste for female leaders and their perceptions of gender roles and female leaders' effectiveness. We examine explicit and implicit measures of voters' tastes (on the use of such measures see, for instance, Charles and Guryan (2007)), Bertrand et al. (2005) and Rooth (2007)). Explicit tastes are captured by voters' stated feelings towards the general idea of male and female leaders, and implicit feelings by Implicit Association Tests (IAT) (Nosek et al. (2007). Our taste IAT is a computer-based double-categorization task which examines the strength of association between images of (anonymous) male and female leaders and normative categories of good and bad. We also conducted an occupation IAT to measure the strength of association between men and women and leadership and domestic tasks. Finally, we collected voter perceptions of leader effectiveness by asking villagers to evaluate the effectiveness of hypothetical female and male leaders described through vignettes and speeches in which leader gender is experimentally manipulated.

Clearly, our measures of "taste" and "perception of effectiveness" remain imperfect: a villager's perceptions of a potential leader's effectiveness may influence his feelings towards the idea of male and female leaders; conversely, the performance rating of a leader may be colored by a villager's general distaste for female leaders. That said, our maintained hypothesis is that the general "feeling" variables remain more closely correlated with deep-seated, and harder to change, tastes while the specific "effectiveness" questions reflect voter perceptions of leader performance, and may respond more to mandated exposure.

Both explicit and implicit measures suggest that in villages that have never experienced political reservation, villagers, particularly men, dislike the idea of female leaders. On a scale of 1 to 10, the average man rates his general feeling towards female leaders one whole point below those for male leaders. IAT measures show that, relative to female leaders, men are significantly more likely to associate male leaders with notions of "good". In addition, female leaders are also perceived as less effective than male leaders. For instance, the average male villager ranked the same speech and vignette describing a leader's decision 0.074 standard deviations lower when the leader's gender was experimentally manipulated to be female (relative to male). The point estimate, while lower, is not statistically different for female villagers.

Mandated exposure to a female leader does not affect villagers' general preference for male

leaders. However, among male villagers it weakens the stereotype (as measured by an IAT) that men are associated with leadership activities and women with domestic activities. It also alters perceptions of female leader effectiveness. In the speech and vignette experiments we find that villagers who have been required to have a female leader rate hypothetical female and male leaders as equally effective. This reduction in bias is, however, absent among female villagers. We posit that a likely reason is the lower levels of political knowledge, and exposure to local politics, among women.

Finally, we examine villagers' evaluation of their actual leader along multiple dimensions, including general effectiveness. Consistent with the experimental data, we find that prior exposure improves the 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, leader rating by male villagers is significantly lower (relative to villages where the leader position has never been reserved). This difference in ratings is absent in villages where the leader position is reserved for women for the *second time* in 2003. We do not find *prima facie* evidence that female leaders in first time reserved villages have less experience, are differentially selected, or undertake different policies, than leaders elected in the second round of reservation. Rather, first time female leaders deliver more public goods and take less bribes, than their male counterparts.[5]

Taken together, our results suggest that while deep preferences and social norms are difficult to erode, affirmative action programs can play an important role in improving perceptions of female leader effectiveness and reducing statistical discrimination.

The rest of this paper is structured as follows. Section 2 provides a conceptual framework and Section 3 describes the data. Section 4 lays out the institutional context and our empirical strategy. Section 5 reports the results and Section 6 concludes.

2   Conceptual Framework

In most elections voters select a preferred candidate while remaining uncertain about the exact competence of the available candidates. We use a simple model of statistical discrimination (our

---

[5]They do marginally better than women leaders in places reserved for the second time. First and second time female leaders have similar characteristics, and all women leaders appear equally biased in favor of women in their policy actions.

formulation follows Aigner and Cain (1977)) to demonstrate that voter uncertainty may interact with an underlying voter distaste for female leaders such that female candidates, on average, are not elected and biased beliefs about their effectiveness as leaders persists. Political reservation can play an important role in reducing the bias in beliefs.

Consider a single village council leadership election with the outcome determined by plurality rule. Candidates differ in competence, e.g. ability to extract state resources for the village, resolve dispute etc. Candidate i's competence is given by η_i, where $\eta_i = \eta + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma_\eta^2)$. We assume that voters are risk averse and value candidate i's effectiveness as $f(\eta_i)$, where $f'(.) > 0$ and $f''(.) < 0$. Risk aversion is a standard assumption in many political economy models.[6] In the context of village council elections risk aversion is very plausible if we equate competence with the resources a leader may raise for the village. An important implication of risk aversion is that voters are likely to penalize politicians on which they have little information.

Voter j's utility from candidate i is defined as:

$$u_{ij} = f(\eta_i) + \alpha m_i + \omega_{ij}, \tag{1}$$

where m_i is a male dummy which equals one if the candidate is male. We assume $\alpha > 0$. This can be interpreted as reflecting taste discrimination against women or a dislike for the policies women implement.[7] $\omega_{ij} \sim N(0, \sigma_\omega^2)$ is a i.i.d random voter and politician specific taste shock.

Let $x_i \in M, F$ represent candidate i's gender. Before casting their vote, citizens observe candidate gender x, and a noisy signal of competence: $\tilde{\eta}_{ix} = \eta_i + \nu_{ix}$, where $\nu_{ix} \sim N(0, \sigma_{\nu x}^2)$. In general, a difference in the mean of the signals across genders (e.g. voters have prior beliefs that women tend to be less competent than men, and these beliefs are self-fulfilling) can be sufficient for statistical discrimination to arise. However, to focus on the possibility that risk aversion can cause women to suffer a "double penalty" (of statistical discrimination reinforcing taste discrimination), we only allow the variance of the signal to differ across genders.

Voters select the candidate who maximizes their expected utility. The expectation of η_i given

---

[6]The typical spatial voting model, which uses a quadratic utility function, implies risk aversion in voting. Risk aversion is also commonly invoked to explain electoral outcomes such as split-ticket voting and the punishment of ambiguous candidates.

[7]For example, Chattapadhyay and Duflo (2004) show that women invest in goods that female-preferred goods.Anticipating that, men may always prefer male leaders.

x and $\tilde{\eta_{ix}}$ is given by

$$E(\eta_i|\tilde{\eta_{ix}}, x) = \eta(1 - \gamma) + \tilde{\eta_{ix}}\gamma$$

where $\gamma = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_{\nu x}^2}$. The error term in this equation is given by:

$$e = E(\eta_i|\tilde{\eta_{ix}}, x_i) - \eta_i = \frac{\nu_i \sigma_{\nu x}^2 - \epsilon_i \sigma_\eta^2}{\sigma_{\nu x}^2 + \sigma_\eta^2},$$

which has variance

$$V(e) = \frac{\sigma_\eta^2 \sigma_{\nu x}^2}{\sigma_\eta^2 + \sigma_{\nu x}^2}$$

The variance is increasing in $\sigma_{\nu x}^2$. Since $f(.)$ is concave, $E[f(\eta_i)|\tilde{\eta_{ix}}, x_i]$, the expected leader effectiveness, is declining in the variance of the signal.

Consider the case where the signal on candidate of gender x is derived as the prediction of a model relating observed characteristics and actions X_i for previously observed leaders of gender x to their competence: $\tilde{\eta_{ix}} = \hat{\beta}_x X_i$. The coefficients of this model are gender specific.[8] If the model is estimated using N_x past leaders' observed actions, then the standard errors of $\hat{\beta}_x$ and the standard deviation of the resulting prediction error will decreases with $\sqrt{N_x}$. Hence the variance of the signal for a gender is decreasing in the number of elected candidates (of that gender).

Taste discrimination, or a preference for the policies selected by men, will imply, for equal expected competence, higher voter utility with a male leader (since $\alpha > 0$). Hence, female candidates will have been elected only in the few cases when the idiosyncratic taste shock was sufficiently favorable for her among a majority of the population. Relative to the male candidate, the ability to extract information from a signal about the female candidate will be weak and $\sigma_{\nu F}^2$ will be very high, compared to $\sigma_{\nu M}^2$.[9] This causes statistical discrimination: aside from any distaste for women leaders, risk averse voters will prefer the male candidate as the variance of their effectiveness is lower. This creates a vicious circle – a female candidate is less likely to be elected both because voters have a distaste for woman leaders, and because they are perceived

---

[8]This is a natural assumption, since men and women may differ in both their actions and characteristics. While an assertive man may be an effective leader, being assertive may be a liability for a woman.

[9]With a positive cost of running, presumably very few women will choose to run which will accentuate the signal-extraction problem for female candidates.

as too risky.[10]

Political reservation serves the purpose of increasing N_f. This improves the precision of voters' estimate of β_F, and reduces the variance of expected competence of future female leaders. For similar underlying levels of competence across genders, reservation-induced exposure should improve the relative evaluation, and electoral prospects, of women.

It is, however, possible to identify two countervailing forces. First, if political reservation worsens voters' prior about the competence of women leaders, then competent female candidates may respond either by under-investing in skills (Coate and Loury, 1993), or by simply choosing not to stand. In this case, political reservation will reduce both the variance and mean of the female distribution. Second, if voters react negatively to having their voting choices constrained (Thernstrom and Thernstrom, 1997) then political reservation may increase taste discrimination against women.

To summarize, if the pool of female candidates (and their investment in job-specific skills) is unchanged by political reservation, then:

Political Reservation and Taste Discrimination Taste discrimination, as captured by a greater preference for male leaders (relative to female leaders), will be unchanged or increase with political reservation.

Political Reservation and Statistical Discrimination Holding competence constant across genders, political reservation will reduce statistical discrimination as measured by the difference in the evaluation of female and male leader effectiveness.

Our model demonstrates that statistical discrimination is a natural by-product of taste discrimination if voters are risk averse. As mentioned earlier, these predictions are not sensitive to the source of statistical discrimination. Specifically, we obtain very similar predictions regarding the efficacy of political reservation if we assume that statistical discrimination occurs because voters have biased priors of average female leader effectiveness (rather than receive signals with

---

[10]The voter's decision problem resembles the classic two-armed bandit problem (Bergemann and Valmiki, 2008). The two arms are the male and female candidate pools, with unknown variance with respect to competence. Voters get new information on the distribution of competence within a gender only if they select a candidate from that gender. Since the observations the voter uses to gain information and reduce the variance of the signal are also his reward (in terms of selecting a leader), he must strike a balance between gaining rewards and gaining information. The median voter will combine information on the mean with beliefs about the variance of payoffs to maximize the present discounted value of his payoff. He will favor exploiting information (i.e. choose the male candidate arm) if he is sufficiently risk averse, faces a finite time horizon or has a sufficiently high discount rate.

different variance).[11]

In the remainder of this paper we use detailed data on voter attitudes to test our two predictions about the implications of political reservation. We start by describing our data.

3 Data

Our data comes from West Bengal, a middle-income Indian state, where elected village councils (Gram Panchayats (GP)) have been an effective elected body of governance since 1978. It is in charge of both constructing and maintaining village infrastructure (streetlights, roads and drinking water facilities) and choosing beneficiaries for most government schemes for the rural poor. The Pradhan has significant discretionary power and Between June 2006 and November 2007 we surveyed 495 villages spread across the 165 GPs in Birbhum district, one of the poorest districts in West Bengal (we randomly selected three villages per GP). In each surveyed village we collected data on the quantity and quality of public good provision, and administered household and individual surveys to a random sample of 15 households. These surveys, together with Implicit Association Tests (IATs), form the main sources of our voter attitude measures.

3.1 Measures of Voter Taste

We measure voter taste by their explicit and implicit feelings towards male and female leaders as expressed in survey responses and in IATs. We collected these data for a random sample of 5 households per village.[12]

To ascertain respondents' explicit taste for male and female leaders, we asked "on a ladder which has steps from 1 to 10, how do you feel about a [X]" where X was (separately) a female leader, a male leader, a female villager and a male villagers. This question is adapted from the "Feeling Thermometer" which is widely used in the political science literature. Feeling thermometer rankings have been used to both produce rank-orderings of parties or candidates

---

[11]Biased beliefs about average competence can be motivated in a number of ways. For example, female leaders tend to be less educated and poorer than male leaders. Voters may conclude that this reflects lower effectiveness.

[12]In each household we conducted IATs with available adults aged between 15 and 45. We set a threshold target of two IATs per household (one per gender), and have 4,378 IAT respondents spread across 1,968 households (in 171 households we had only one participant). We exclude IAT data from the one pilot village in each of seventy-seven GPs. Our survey data includes 7,182 survey responses across 2,926 households; we do not exclude the pilot sample for the survey questions.

for voters (for example, see Keller and Mirer (1974)) and to measure partisan affiliation (e.g. Alvarez (1990) and Weisberg (Weisberg)).

We complement the explicit taste measures with taste Implicit Association Tests (Greenwald et al., 1998). An IAT is a computerized test which aims to measure attitudes of which respondents may not be explicitly cognizant. Over the last ten years IATs have been widely used in a diverse array of disciplines (Nosek et al., 2007), including various subfields in psychology, neuroscience, market research and increasingly in economics.[13]

To measure implicit bias an IAT 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 into two categories (e.g. female and male names, and adjectives evoking good or bad attributes). Some tasks require the respondent to put male names and good attributes on the left, and female names and bad attributes to the right while in other tasks the categories are switched (i.e. males names and bad attributes on the right, and female names and good attributes on the left).

IATs assume that a stronger association between two concepts makes the sorting task easier. An automatic association can, therefore, be detected by comparing response time across concept pairs. Specifically, each test has two test blocks of interest: a "stereotypical" block that, for instance, associates male names and good attributes and female names and bad attributes; and a "non-stereotypical" block that reverses this association. IAT bias is given by the D-measure: the normalized difference in mean response time between the "non-stereotypical" and "stereotypical" test blocks (Greenwald et al. (2003)). A higher value of the D-measure indicates a stronger implicit stereotype.

To the best of our knowledge, we are the first to implement IATs in a field setting in a low income country. Our IATs used either audio or pictoral prompts and, therefore, did not assume literacy or familiarity with computers.[14] Appendix Figure 1 shows an IAT screen and a picture of a participant playing the game, and the IAT Appendix lists the set of prompts (for

---

[13]Rooth (2007) shows that the D-measure for racial bias of Swedish employers correlates positively with the rate at which they call back applicants of different races.

[14]Similar IATs have been used for children (Baron and Banaji, 2006). We used an age cut-off of 15-45 years to identify a respondent sample which was comfortable with playing IATs. We excluded respondents with an average response time of over 6 seconds in 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 any prompt. Finally, we randomize the ordering of stereotypical and non-stereotypical blocks across respondents.

more details, see Beaman et al. (2008)). A second innovation of our study is to exploit variation in respondent exposure to female leaders to examine the malleability of IAT bias. This is in contrast to most IAT studies which assume IATs are "hard-wired". We are only aware of one other field study which investigates this question, and we discuss this in section 3.2 below.

We administered two taste IATs; this class of IATs have been widely used to measure group prejudice, such as racism (Banaji, 2001). The first assesses the associational strength between male and female names and good (nice, delicious etc.) and bad (nasty, horrible etc) attributes. The second measures the association between these attributes and images of male and female politicians (pictures of either men or women giving speeches, leading crowds etc). In both cases the stereotypical block places male names or leader pictures and adjectives evoking good attributes on one side of the screen, and female names or leader pictures and adjectives evoking negative attributes on the other side.

3.2   Measure of Leadership Stereotype

Our model suggests that while political reservation is unlikely to reduce taste discrimination as measured by the taste IAT, it should make it more likely that voters associate women with leadership activities (by virtue of having seen women in action as leaders).

To examine changes in the gender stereotyping of occupations we administered an activity-based IAT. The IAT examined whether villagers exposed to reservation were less likely to associate women with domestic tasks and men with leaderships activities. Specifically, the IAT examined the association between male and female names and domestic (e.g. cooking, eating puffed rice etc) and leadership activities (such as meeting, public speaking etc). To avoid loading the results towards associating women with domestic activity, the domestic activities were chosen to depict gender neutral activities which were as likely to be performed by men as women.

Unlike the taste IATs, this IAT does not capture a value judgement but rather indicates whether women are considered potential leaders. Rudman and Kilianski (2000) used such an IAT to show that respondents in the US associated female names more strongly with family than with career. There is also some indication that responses to occupation IATs are influenced by the environment. Dasgupta and Asgari (2004) administered this IAT to American college students and found that students in colleges with a higher proportion of female professors were

11

more likely to associate women with professional activities. However, endogenous selection into college makes a causal interpretation of these results difficult.

We hypothesize that the main channel by which reservation influences voter beliefs is enhanced information on prospective women leaders. Associating women more easily with leadership activities would be a natural first step. Comparing outcomes across the taste and occupation IATs is also informative of whether exposure differentially affects preferences and stereotypes.

3.3 Measures of Leader Effectiveness

To measure bias in voter perception of leader effectiveness, we administered speech and vignette experiments of the "Goldberg paradigm" type.[15] Such experiments have been widely used in the US to assess bias in the perception of the effectiveness of women as leaders (see, for instance, Matland (1994), Huddy and Terkildsen (1993) and Eagly and Karau (2002)). These experiments were administered as part of a longer survey administered to a prime-aged male and female respondent in each of our surveyed household. Our sample consisted of a random sample of 15 households per village (of these, a random subset of 5 households were administered IATs). Overall we conducted 6,717 male and 6,780 female adult modules.

In the speech experiment the respondent heard a short tape-recorded leader speech which was adapted from an actual village council leader (now on, Pradhan) speech at a village meeting. Respondents were randomly assigned one of six speech recordings (three male and three female), and told that this was a speech by a Pradhan in a village meeting in another district. In the speech, the Pradhan responds to a villager complaint about a broken tubewell by requesting villagers to contribute money and effort for local public goods.[16] After hearing the speech the respondent was asked seven questions on different aspects of the Pradhan's perceived performance and overall effectiveness. These included whether the Pradhan addressed villager's concerns correctly, whether he/she would be good at collecting resources from villagers etc. We did not include any questions on whether the villager sympathized with the Pradhan; rather, all the questions were strictly about evaluating the Pradhan's action and his or her effectiveness.

We conducted a similar exercise with the "vignette". Each respondent heard a randomly

---

[15]Goldberg (1968) first used such experiments where gender of a protagonist is randomly varied to identify gender bias.

[16]A colorful literal translation, provided by our survey team, is in included the Appendix.

selected vignette, in which a situation of resource scarcity was described and the Pradhan had to decide 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.

The Pradhan gender was randomly varied across respondents, in such a way that a respondent was exposed to the same Pradhan gender in the speech and vignette. This allows us to combine a villager's responses across the speech and vignette and ask whether, holding actual action constant, villagers, on average, rank female leaders below male leaders.

Finally, we asked villagers to evaluate their actual Pradhan's effectiveness on multiple dimensions (using a ladder based scale of 1 to 10). Questions included "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?"

3.4   Discussion

Clearly, neither our "taste" nor our "perception of effectiveness" measure is a perfect empirical counterpart to the corresponding parameter in our simple model. Voters may associate female leaders with bad adjectives in general because they considered them incompetent. Equally, villagers may deem the woman leader giving the speech ineffective because they want to punish her for violating norms of prescribed behavior. On other words, akin to the utility function we posit, our measures of voter attitudes may capture both voter taste and perception of effectiveness. However, as long the relative weights of taste and statistical discrimination vary across these measures we would expect our taste measures to, on average, be less affected by reservation than our perception of effectiveness measures.

4   Institutional Context and Empirical strategy

As a precursor to the empirical analysis we describe the system of political reservation in West-Bengal and our empirical strategy.

4.1 Political Reservation in India

Electoral gender quotas usually require that women constitute a certain percentage of either party candidate lists or the elected body such as the legislature.[17] In India, a 1993 constitutional amendment introduced a system of political reservation for village council (Gram Panchayat) heads in which all leaders are elected, but only women can run in women reserved councils.[18] This led to a dramatic rise in local female leadership across Indian villages and the number of village-level female elected leaders is now close to 40 percent.

In West Bengal, the Panchayat Constitution Rule was modified in April 1998 to introduce reservation for women and two disadvantaged minorities, Scheduled castes (SC) and Scheduled tribes (ST) (Government of West Bengal, 1998). Two elections have been conducted with reservation: in 1998 and 2003.

The rules require that prior to an election GPs are randomly assigned to three lists: Reserved for SC, Reserved for ST, and Unreserved.[19] These lists are redone at every election to ensure that no GP features on the SC or ST list for two consecutive elections. In each list GPs are ordered by their serial numbers and every third GP is reserved for a woman. In 1998, every GP 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, random assignment to being reserved, for women, SC and ST. Second, implicit stratification of women randomization by SC/ST and administrative block (since GP serial numbers start with a block identifier; a block is the administrative unit below the district) Third, a 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 reconstructed the 2003 reservation list using GP serial numbers and the electoral law tables. We found that the rule held, with no

---

[17]These quotas are typically constitutionally mandated (Burkina Faso, Nepal, Philippines and Uganda), by electoral law (as in many parts of Latin America, and in Belgium, Bosnia and Herzegovina, Serbia and Sudan) or political parties have some type of quota (Argentina, Bolivia, Ecuador, Germany, Italy, Norway and Sweden). A growing concern that requirements on candidate lists may not increase actual female representation in politics (Jones, 2004)) has led to gender quotas being increasingly introduced using political reservation like Jordan, Uganda and Rwanda.

[18]India has had universal franchise since Independence, and many prominent female leaders (including Indira Gandhi, Mayawati, Mamata Banerjee, and now Pratibha Patil, the first female President). However, the average share of women politicians at the national and state level remains around 10%.

[19]A Gram Panchayat typically consists of multiple villages (in West Bengal the average is 10-12 villages). The extent of reservation for SC and ST is proportional to their population share in the district.

exception (The same was true in 1998, see Chattapadhyay and Duflo (2004)).

Figure 1 shows the patterns in GP reservation in our study area of Birbhum district in West Bengal. 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. All the Pradhans in GPs reserved for women are female, and the overall fraction of female Pradhans in Birbhum is 43 percent (one reserved GP, due to political disturbances, did not have a Pradhan at the time of our survey).

In non-reserved GPs few women are elected, but this number has increased over time. *Prima facie* the increase appears to be unrelated to the reservation policy. The fraction of women elected in 2003 from GPs that were only reserved in 1998 (14%) is comparable to the fraction of women Pradhans elected from never reserved GPs in 2003 (16%).

In Table 1 we use 1991 census data to check for systematic differences between 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. As expected given the randomization these variables are not jointly significant predictors of reservation assignment (see p-values in columns (5) and (6)), though three variables (sex ratio under 6, having a hand pump and having permanent (pucca) road) are significant at 10% or less in column (5). Our regressions include controls for these variables.

4.2 Empirical Strategy

Given the randomized setting, our empirical strategy is 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 allows us to 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 unreserved and previously reserved GPs, have a female Pradhan (Figure 1), this reduced form effect is similar to that obtained by instrumenting for Pradhan's gender by reservation status of GP.

The unit of observation in our regressions is villager i in GP g situated in block b. Throughout, we report separate results for male and female villagers. We have two sets of outcomes, and the regression specification differs across these. Our first set of outcomes are common across

15

respondents, and include the difference in a villager's feelings towards male versus female leaders, the IAT D-measures, and evaluations of the Pradhans. For these outcomes we report two specifications. The first simply compares outcomes of interest across ever reserved GPs and never reserved GPs,

$$y_{ig} = \beta R_g + X_{ig}\gamma + \alpha_b + \epsilon_{ig} \tag{2}$$

where R_{gb} is an indicator variable for the GP being currently or previously reserved (now on, ever reserved). The second separates reserved and unreserved GPs further:

$$y_{ig} = \beta_2 R_{g2} + \beta_{2and1} R_{g2and1} + \beta_1 R_{g1} + X_{ig}\gamma + \alpha_b + \epsilon_{ig} \tag{3}$$

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 only being reserved in 1998. Standard errors are always clustered by GP.

All regressions include a block fixed effects (α_b), and a set of respondent controls (X_{ig}): age, household size, education, caste, religion and proxies for household wealth constructed using a principal component analysis. We also include investgator gender fixed effects, a survey period indicator and three village controls (under 6 sex ratio, handpumps, and roads).[20]

Our second set of outcomes consist of the vignette and speech responses. Each respondent received a single vignette and speech about a male or female Pradhan. We are interested in whether male and female leaders are judged differently, and whether this varies with reservation status of the village. Let F_{ig} indicate whether respondent i was presented with a "female" stimulus (i.e. heard the speech in a female voice, or was described the vignette with a female leader). We estimate:

$$y_{ig} = \delta F_{ig} + \lambda(R_g * F_{ig}) + \mu R_g + X_{ig}\gamma + \alpha_b + \epsilon_{ig} \tag{4}$$

---

[20]Regressions without controls variables are very similar to those reported here, and are available from the authors.

and

$$y_{ig} = \delta F_{ig} + \lambda_2(R_{g2} * F_{ig}) + \lambda_{2and1}(R_{g2and1} * F_{ig}) + \lambda_1(R_{g1} * F_{ig}) + \sum_k R_k \mu_k + X_{ig}\gamma + \alpha_b + \epsilon_{ig}, \quad (5)$$

where the indicator variables R_k control for the main effect of different reservation categories.

In equation (4) 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 (5), 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 we expect the coefficients on the variables of interest to go in the same direction within a group (or "family") of outcomes. To avoid drawing inferences based on selected outcomes, we report effects that average across all outcomes within a family, following Kling et al. (2007). Specifically, for each outcome we construct a normalized transformation where we subtract the mean for 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.

To summarize, the regression tables report two sets of regressions. Panel A reports the coefficients on the reservation dummy from estimating either equation (2) or equation (4), and Panel B from estimating equation (3) or equation (5). The last row in each Table reports the baseline level of bias for the relevant attitude measure (usually the mean for the unreserved sample).

5 Results

5.1 Do Villagers know their Leader?

Our model of voter behavior is based on the premise that villagers' are aware of local politics, and able to gather information about leader actions. Therefore, we start our empirical analysis by examining villager awareness of local politics and exposure to their Pradhan.

Male villagers exhibit relatively high levels of knowledge, and involvement in local politics.

However, we observe a significant gender gap on every measure with the exposure of women being much more limited. The last row in Table 2 provides the means of variables measuring political awareness and activism in the sample of never reserved GP. In these GPs, 67% of the male, but only 33% of the female, villagers know their current Pradhan's name. A similar-sized gender gap exists for previous Pradhans (columns (1)-(4)). 53% of the male, but only 13% of female, villagers state having ever approached the Pradhan about their needs or village issues (column (5) and (6)). A large majority of men have heard of the village meeting (column (7)), even though a minority has actually attended one (34%). In contrast, only 56% of the women have heard of the meetings, and only 6% have attended.

Panel A shows the coefficient for the "ever reserved" variable in equation (2), and Panel B the coefficients for the reservation variables in equation (3). In columns (1) through (6) we see that both genders are less likely to have interacted with female Pradhans and to know their name. For example, men are between 11-12.5% less likely to know the name of their Pradhan if the GP is currently reserved, and between 10-14% less likely to know the name of the previous Pradhan if the GP was previously reserved.

This negative effect of reservation on villagers' interactions with the village leaders has several possible interpretations: female Pradhans may be less proactive, less likely to be the traditional village leader or, more generally, be a trusted village authority. That said, an important take away from Table 2 is that, even in reserved GPs, a majority of male villagers know who their Pradhan is, and have, at some point, seen him or her.

5.2   Do Villagers Prefer Male Leaders?

In Table 3 we examine differences in villagers' general feelings towards male and female leaders. As a reminder, we measure an explicit preference for male leaders by the difference between a villager's general rating of male and female leaders (on a 1-10 scale). We measure implicit preference for male leaders by the D-measure in the taste IATs. To benchmark a respondent's ranking of leaders, we also report the rating and IAT results for respondent preferences for male versus female villagers.

The last row of Table 3 presents summary statistics for never reserved GPs. Villagers in Birbhum are not shy about admitting explicit preferences. In columns (1) and (2) we see

that men rank male villagers 0.69 points higher than female villagers. Women also exhibit a significant, though much smaller, bias towards male villagers. For both genders, this bias is magnified in the case of leaders (columns (3) and (4)). Male villagers rate male leaders 1.44 points higher than female leaders. For female villagers the difference, while smaller (0.56), remains significant.

This distaste for female leaders is not ameliorated by exposure. Rather, the coefficient of "ever reserved" (Panel A, column (3)) is positive (0.22) and significant for male villagers, suggesting that their relative preference for male leaders (compared to female leaders) is strengthened in GPs which have been experienced a gender quota. The fact that this effect is strongest in GPs that were reserved only once (Panel B, see coefficient on reservation in 1998 or in 2003) leaves open the possibility that villagers get used to quotas over time.

Columns (5)-(8) report findings for the villager and leader taste IATs. Both IATs demonstrate a strong same gender preference among male and female villagers. Men are significantly more likely to associate good with male names while women associate female names with positive attributes.[21] Neither is affected by reservation.

These results are discouraging of the view that affirmative action can alter voter preferences. The evidence on explicit bias is striking, especially when contrasted to rich countries where explicit bias tends to be muted even when respondents exhibit strong implicit bias (Bertrand et al., 2005). A possible explanation is strong social norms that women should not be leaders. The finding that men are *more* biased in once reserved GPs, but only as measured by explicit measures, supports the "backlash" hypothesis. That is, forcing men to elect women leads them to state a dislike for female leaders as a protest against the quota system.

Social norms which militate against female leadership could also explain the fact that we observe consistent results across explicit and implicit measures for male, but not female, villagers. Women may be less biased towards female leaders (as captured by their implicit beliefs), while recognizing that ranking female leaders below male leaders is prescribed behavior. Finally, the two taste IATs evoked very similar results from both genders. It may be that in both cases villagers mostly perceived the gender of the character, and did not associate pictures of

---

[21]Our finding of same sex preference on the part of men differs from what is typically found in developed countries, where these tests exhibit a "women are wonderful" effect – and everyone associate women with good things (Eagly and Mladinic, 1989).

female leaders with leadership. In that sense, the second IAT may have failed to capture the "leadership" dimension.

5.3 Do Villagers Have Gender-Occupation Stereotypes?

Reservation does not appear to have not altered voter's taste for females leaders. We now use an occupation-based IAT to examine whether it made them more aware of the possibility that women can lead. The IAT tests whether villagers are more likely to associate leadership activities with men (and domestic activities with women).

The results, given in Table 4, paint a very different picture from the taste IATs. Both genders exhibit significant biased against women in leadership activities in never reserved GP. Specifically, respondents are faster at associating women with domestic activities than with leadership actions (on average, men are 0.1 standard deviation faster, and women 0.15 standard deviation faster). Exposure to a female leader, however, significantly reduces this association for men. This suggests that reservation-induced exposure may have reduced the strength of the stereotype linking men with leadership activities, at least among male villagers; they may have learned that women can be capable leaders. It is worth noting that this is, to the best of our knowledge, the first study that identifies a causal effect of a policy on implicit beliefs, as measured by an IAT.

We do not observe any impact of exposure for women. If anything, the stereotype associating women and domestic activities increased for those women exposed to a female Pradhan for the first time in 2003. We return to implications of this result in Section 5.6.

These results suggest a more nuanced view of how political reservation affects attitudes. While reservation does not make male villagers more sympathetic to the idea of female leaders, it makes them recognize that women can lead. The fact that male villagers begin associating women with leadership opens up the possibility that they may have also updated their perception of the competence of women leaders, which is what we turn to now.

5.4 Do Villagers Perceive Female Leaders as Less Effective?

5.4.1 Hypothetical Leaders

Table 5 reports villagers' evaluation of hypothetical Pradhans which were presented to them via a tape-recorded speech or a vignette. Recall that respondents were exposed to the same gender in the speech and the vignette. In both cases, the respondents answered multiple questions on the perceived effectiveness of the Pradhan.

In columns (1) and (2) we report the average coefficients, averaged across all questions in the speech and the vignette (Kling et al., 2007). In panel A we see that men in never reserved villages rate the effectiveness of a hypothetical female Pradhan 0.074 standard deviation below that of a male Pradhan (with a standard error of 0.029). While lower and insignificant for women (0.049, with a standard error of 0.034), we cannot reject the hypothesis that both genders are equally biased.

Reservation alters male evaluations: the coefficient on the interaction between female Pradhan and ever reserved is 0.122, and is strongly significant. Adding the coefficients on female Pradhan and its interaction with reservation suggests that political reservation completely erases, and indeed reverses, this bias (though the resulting pro-female bias is not strongly significant). In panel B, we see that the impact of reservation is similar across all reservation categories: all coefficients are positive, with similar sized and and statistically indistinguishable point estimates. The results suggest that exposure, at some point, to a female leader is sufficient to erase male villager's statistical discrimination against female leaders, and this effect persists even after the woman has left office. In contrast, reservation has no impact on leadership evaluation of female villagers.

Columns (4)-(12) consider each separate effectiveness question which enters the overall assessment.[22] We see that no single variable drives the results. Rather, the results are remarkably consistent across outcomes. On each outcome men in never reserved GPs evaluate the hypothetical female Pradhan more negatively than her male counterparts. This bias is absent in GPs which have been ever-reserved. In the case of women villagers, reservation has no impact.

Why are women's evaluations of the effectiveness of female leaders unaffected by reservation?

---

[22]They are still averaged across the speech and vignette questions. Appendix Table A1 reports question-wise results for the speech and vignette.

One possibility is the much lower exposure of female villagers to local politics (see Table 2). For the presence of female leaders to alter an individual's perception of their competence, she needs to have observed them in action. If women are largely inactive in politics then it is unsurprising that reservation does not affect their evaluations. Table 6 provides evidence which supports this hypothesis. We continue to consider the average response to the speech and vignette questions as the relevant outcome variable. We expand our set of explanatory variables to include the interaction of the reservation effect with whether the villager knows the Pradhan's name. Columns (1)-(3) present the OLS results. While noisy, the estimates suggest that even for men, the effect is entirely driven by those who know the Pradhan. To address the concern of reserve causality (less biased villagers are also more likely to know the Pradhan's name), columns (4)-(6) report IV regressions where we instrument whether a villager knows the Pradhan's name with a set of variable indicating geographic and social proximity with the Pradhan (same village, same caste) and political knowledge (read the newspaper, knows the name of legislator and has heard of village meetings). The IV estimates are significantly larger for both genders (though still noisily estimated).

Taken together, our speech and vignette results are consistent with the predictions of a simple model of statistical discrimination in which villagers update their initial prior that women leaders have less effective only after hey have seen a woman leader in action. The results suggest a rapid updating of voter beliefs about women's effectiveness, at least among men. This suggests that political reservation can, even if it does not change voter preferences, alter the future electoral fortunes of women (assuming voters place enough weight on competence). And even if electoral outcomes are unchanged, political reservation can enhance welfare by improving the information available to villagers. While the speech and vignette experiments are appropriate for measuring pure bias, they may not perfectly predict how reservation will influence the evaluation of real leaders in situations where the information available about leaders may differ with the gender of the leader.

5.4.2   Actual Leaders

In our survey we asked each respondent to evaluate their Pradhan's performance on four different dimensions using the same 1 to 10 ladder which was used for the vignette and speech experiments.

The questions were purposefully chosen to be similar to those asked about the hypothetical Pradhan in the speech and vignette.

In Table 7 we consider both the average across the four questions, and question-wise responses as the outcome. In all cases we find that villagers evaluate their Pradhan as significantly less effective in GPs first reserved in 2003. Column (1) examines the average outcome 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, standard error of 0.06). These results echo Duflo and Topalova (2004) who, using nation-wide data for India, showed that villagers are more dissatisfied with female leaders in general (also see Duflo (2005)).

The coefficients of interest are very similar across individual outcomes (columns (3)-(10)). Female Pradhans in GPs first reserved in 2003 are ranked lower on general effectiveness, ability to look after village, or respondent's, needs, and in preparing the list of beneficiaries for the "below the poverty line" (BPL) program.

However, consistent with the experimental data, the second row in Table 7 shows that villagers' evaluation of female Pradhans in twice reserved GPs is statistically indistinguishable from that of Pradhans in unreserved GPs. This is analogous to the speech and vignette results, where (male) respondents exposed to a female stimulus considered the hypothetical female Pradhans to be less effective than their male counterparts in unreserved, but not ever reserved, GPs. There is weaker evidence that the evaluation of female respondents is also affected (the difference between the coefficient on Pradhan appreciation in first and second time reserved GPs is significant, at the 1% and 8% level in the male and female regression respectively). Finally, row 3 shows that (mostly male) leaders in GPs that were only reserved in 1998 are evaluated at par with leaders from never reserved GPs.

These results are consistent with the view that exposure to a female leader causes villagers to update their prior on average female effectiveness. It is worth noting that the first cohort of women leaders fails to benefit from this updating, possibly because voters tend to "stick" with their original assessment (Mullainathan and Washington, ming). In contrast, the second cohort of female leaders does not appear to suffer from statistical discrimination. Needless to

say, we need to be caveat this interpretation of these results with the observation that a possible l alternative explanation is that, 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.

Are First-time Women Leaders less Effective?

In Table 8 we present evidence that, prima facie, male Pradhans do not appear to outperform female Pradhans. In column (1) we see that, across all public goods, the average number of repairs or new constructions since the last election is 0.21 standard deviations higher in villages in reserved GPs. If anything, Pradhans in GPs first reserved in 2003 are more effective than Pradhans in never reserved GPs. In twice reserved GPs public good provision is somewhat higher than in never reserved GPs. While statistically indistinguishable from the effect in GPs first reserved in 2003, the effect is smaller and insignificant (0.096, with standard error of 0.08).

A second possibility is that women Pradhans invest in more, but lower quality, public goods. However, in column (2) we do not see any significant differences in the quality of public good provision across currently reserved, previously reserved or twice reserved GPs. Once again, the performance of women elected in first-time reserved GPs is indistinguishable from women elected from GPs reserved for the second time.

Since public goods are mainly financed by State Government funds, the contrast between villagers' (especially male) negative evaluation of female Pradhans and the fact that female Pradhans seem to invest 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 we can identify bribe payments and voluntary contributions for two specific goods. The first is whether the household paid a bribe to receive a BPL card. The second is bribes and voluntary contributions for drinking water provision. In column (3) we see that on average, individuals in currently reserved GPs (for both the first and second time reserved) are 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 columns (5) and (6) we examine self-reported villager satisfaction with public goods. We consider villager satisfaction averaged across all public goods as the outcome variable, and we

observe very similar levels of satisfaction across GPs in different reservation categories. [23]

The evidence suggests that women leaders provide more public goods, of equal quality, at a lower effective price. While we may have failed to measure some critical aspect of Pradhan performance, the weight of the evidence is consistent with our experimental findings. That is, villagers rate female leaders in villages reserved for the first time more negatively even though they perform at least as well as the average male Pradhan.

Finally, we examine whether female leaders decisions are biased towards women's preferences. This is motivated by the fact that women leaders invest in different types of goods than men (see Appendix Table A2).We follow Chattapadhyay and Duflo (2004) and use the difference in male and female villagers' complaints in public meetings to construct a measure of female preferences. We then examine whether public good allocation is aligned more closely with female villager preferences. Column (7) shows that female leaders invest more in women-preferred goods. At first pass, this suggests an alternative explanation for both the apparent taste discrimination against female leaders and the observed backlash in male preferences for female leaders in Table 3. However, this explanation is difficult to square with the fact that the policy choices of female Pradhans in first and second reserved GPs are equally pro-woman.[24] And we had seen in table 3 that the "backlash" effect was particularly strong for GP reserved only once.

To further investigate possible differences in female Pradhan selection in first time and second time reserved GPs, we compare salient leader characteristics by reservation status in Table 9. Male and female leaders differ: women are younger, less educated, more likely to come from landless and less wealthy households. However, we do not see 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 GPs which are reserved twice in a row for a woman have usually

---

[23]The results are very similar when we examine individual satisfaction measures. The one exception is the BPL list, men, in particular, resent the way female leaders allocate them. However, the fact that there is a fixed BPL quota implies that there is no way to allocate BPL cards in a way that pleases everyone. We should also note that our satisfaction results differ from Duflo and Topalova (2004): in an Indian-wide data set they found that, on average, villagers were less likely to be satisfied with public good provision when women. were in charge.

[24]This is somewhat surprising since one may expect parties to field different candidates (or voters to select different types of women) once they realize that female leader's decisions reflect the policy preferences of women. Surprisingly, this is also true for previously (but not currently) reserved GPs.

changed reservation category. For example, the GP may have been in the general category in 1998 but be reserved for SC in 2003 or vice-versa. The incumbent is thus either ineligible for re-election in 2003 (if the seat moved from the general to SC category), or is unlikely to be elected (if the seat moved from the SC to general category, since SC candidates are typically not elected from general seats). As a result, incumbents are rarely re-elected in twice reserved GPs.

In summary, the evaluation of the actual Pradhan closely mirrors that of the hypothetical ones. The first time men are exposed to a female Pradhan (either real or hypothetical), they consider her ineffective. This difference disappears the second time around. The effects for women, while similar, are attenuated and insignificant. While we can not fully rule out the possibility that performance differences underlie 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, the consistency between these two sets of results strongly points to the possibility that exposure to female leaders reduced bias.

5.5   Robustness Checks and Alternative Explanations

In Table 10 we conduct a number of robustness checks to examine possible alternative explanations for our results (other than gender bias).

A first possibility is that female leaders are evaluated as worse than male leaders, not because of their gender, but because they are new and inexperienced. In the vignette and speech experiments, respondents who have never had a female leader may assume that female Pradhans are new and inexperienced, and therefore likely to be worse. In Panel I of Table 10 we show that a respondent's evaluation of a Pradhan is uncorrelated with whether the Pradhan is new (i.e has not been elected to any GP position prior to 2003; about 60% of male Pradhans have previous experience and the experience of women leaders is significantly lower (see Table 9)).

Another possibility, which we alluded to earlier, is that voters have a worse evaluation of women leaders simply because they were elected on a reserved seat. This cannot explain the absence of bias for women elected from twice-reserved GPs, unless we assume that voters get used to the idea of reservation. If this assumption is correct, then we would expect that voters rank other reserved Pradhans lower than unreserved Pradhans, but their ranking of reserved

26

Pradhan would improve in GPs which were previously reserved for women. In Panel II we show that while voters rank SC Pradhans lower than non SC Pradhans (potentially because they are also discriminated against), previous reservation for women does not affect this. Similarly Panel III shows 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 previous reservation for SC lessens voter 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.

5.6   Do Women Leaders React to Voter Bias?

Reservations ensure that women get elected. But if there is a significant amount of discrimination against them, this may make their job more difficult. In Table 11, we report results based on surveying Pradhans' which suggests that their satisfaction with their job correlates with the how well they are received by voters.

We first consider Pradhan's stated satisfaction with life (on a scale from 1 to 5, using the standard phrasing for this question on self reported happiness). In column (1) of Table 11 we see that Pradhans in GPs first reserved for women in 2003 are significantly less satisfied than Pradhans elected from unreserved seats (the coefficient is -0.47, and the mean for men is 3.7). This difference is absent for Pradhans elected in GPs reserved for the second time; moreover, the difference in stated satisfaction levels of female Pradhans in first and second time reserved GPs is strongly significant.

In column (2) we restrict the sample to women who were elected as Pradhan from reserved GPs in either 1998 or 2003 (for Pradhans elected in 1998 we use survey data from Chattapadhyay and Duflo (2004)). We see that, relative to women elected from second-time reserved GPs, women leaders elected from first time reserved GPs (in either 1998 or 2003) report their gender as creating more on-job difficulties. Forty-one percent of the first time female Pradhans reported that their gender created problems. Only 10 percent of the second time Pradhans did.

Finally, in column (3) we expand the sample to all Pradhans elected in either 1998 or 2003, and examine Pradhan plans to run for re–election. Incumbents in GPs reserved for the first time in either 1998 or 2003 are 11% more likely to plan to not rerun. This difference is absent

for Pradhans in twice reserved GPs.

Overall, these findings suggest that bias is reflected in lower levels of life and job satisfaction for first time female leaders. This is not the case for second time leaders. Although this is somewhat speculative (especially in the light of any other effect for women), this may also explain why, in Table 4, 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.

6   Conclusion

We have provided evidence that voters prefer male leaders and have biased priors on the effectiveness of female leaders leaders. Male villagers rate their feeling towards female leader more than a point below that towards male leaders; in never reserved villages hypothetical leaders performing the same action are evaluated as less effective if they are female.

We have argued that, in large part, the overall "feeling" for male leaders reflected "taste" discrimination, or a deeply rooted social norm against the idea that it is appropriate for women to lead. Such social norms may be hard to alter, even ten years after reservation was introduced. In contrast, we may expect that the effectiveness rating of a leader in a speech or a vignette is easier to change in so far as it reflects statistical discrimination (possibly due to, or reinforced by, the lack of exposure to female leaders).

Indeed, political reservation, which forces voter exposure to a female leader, does not alter villagers' explicit and implicit dislike for female leaders, but, at least among males, it completely erases the bias in perceived effectiveness.

We fail to find any evidence in favor of the alternative explanation that negative feelings towards the idea of female leaders reflects worse performance by female Pradhans. If anything, female Pradhans appear to outperform male Pradhans on many dimensions of performance: they provide more public good at a lower price. Moreover, the performance rating of the actual female Pradhan closely tracks that of hypothetical leaders: the first cohort of female leaders is rated below men (possibly reflecting statistical discrimination), but the second cohort is rated as par with men. Another alternative explanation, which we are unable to disentangle from

taste discrimination, is that men dislike female Pradhans because they implement pro-woman policies.

The difference in our findings on how reservation impacts perceived effectiveness of specific leaders and overall taste for female leaders makes it difficult to predict whether a temporary program of reservation will translate into a permanent improvement in the electoral fortunes for women. The results of the 2003 election paint a mixed picture. Seventeen of the 55 women elected on reserved seat in 1998 ran for re-election in 2003. However, only 4 were re-elected. Of these, 3 were re-elected from the 20 GPs which continued to be reserved in 2003. Only one was re-elected on an unreserved seat.[25] In comparison, 19 of the 72 male Pradhans ran for re-election and 4 got re-elected. In sum, the fraction of Pradhans who stand for re-election, and the unconditional probability of being re-elected, is lower for male than for female Pradhans (in part, because many men are prevented to run because of rules). However, the probability of re-election conditional on contesting an unreserved seat is much higher for male candidates.

To conclude on a more positive note, taken together, the weight of the evidence suggests that exposure matters and political reservations, or more generally quota systems, may play an important role in reducing biased perceptions of female leader effectiveness.

References

Aigner, D. and G. Cain (1977). Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review 30*, 175–187.

Akerlof, G. and R. Kranton (2000). Identity and Economics. *The Quarterly Journal of Economics 115*(3), 715–154.

Alvarez, M. (1990). The Puzzle of Party Identification - Dimensionality of an Important Concept. *American Politics Quarterly 18*, 476–91.

Banaji, M. (2001). Implicit attitudes can be measured. In I. . J. S. N. E. H. L.Roediger (Ed.), *The nature of remembering: Essays in honor of Robert G. Crowder*. Washington, DC: American Psychological Association.

Baron, A. and M. Banaji (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science 17*(1), 53–58.

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova (2008). Gender Bias and the Implicit Association Test: Evidence from Rural India. Mimeo, Harvard and MIT.

Bergemann, D. and J. Valmiki (2008). Bandit problems. In S. Durlauf and L. Blume (Eds.), *The New Palgrave Dictionary of Economics*.

---

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

Bertrand, M., D. Chugh, and S. Mullainathan (2005). Implicit Discrimination. *American Economic Review 95*(2), 94–98.

Boisjoly, J., G. Duncan, M. Kremer, D. Levy, and J. Eccles (2006). Empathy or Antipathy? The Consequences of Racially and Socially Diverse Peers on Attitudes and Behaviors. *American Economic Review 96*, 1890–1906.

Charles, K. and J. Guryan (2007, December). Prejudice and the economics of discrimination. NBER Working Paper 13661.

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

Coate, S. and G. Loury (1993). Will affirmative-action policies eliminate negative stereotypes? *American Economic Review 83*(5), 1220–1240.

Dahlerup, D. (2006). *Women, Quotas and Politics*. New York: Routledge.

Dasgupta, N. and S. Asgari (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology 40*, 642–658.

Duflo, E. (2005). Why political reservations. *Journal of the European Economic Association 3*(2-3), 668–678.

Duflo, E. and P. Topalova (2004). Unappreciated service: Performance, perceptions, and women leaders in india. *Mimeo*.

Eagly, A. and S. Karau (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review 109*, 573–598.

Eagly, A. and A. Mladinic (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin 15*(4), 543–558.

Frechette, G., F. Maniquet, and M. Morelli (forthcoming). Incumbents interests, voters bias and gender quotas. *American Journal of Political Science*.

Goldberg, P. (1968). Are women prejudiced against women? *Transaction 5*, 28–30.

Goldin, C. (1990). A pollution theory of discrimination. Mimeo, Harvard.

Government of West Bengal (1998). *The West Bengal Panchayat (Constitution) Rules, 1975*. Department of Panchayats & Rural Devellopment.

Greenwald, A., M. Banaji, and B. Nosek (2003). Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. *Journal of Personality and Social Psychology 85*(2), 197–216.

Greenwald, A., D. McGhee, and J. Schwartz (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology 75*(6), 1464–1480.

Huddy, L. and N. Terkildsen (1993). Gender stereotypes and the perception of male and female candidates. *American Journal of Political Science 37*(1), 119–147.

Ingelhart, R. and P. Norris (2003). *Rising Tide: Gender Equality and Cultural Change Around the World*. Cambridge: Cambridge University.

Jones, M. (2004). Quota legislation and the election of women: Learning from the costa rican experience. *Journal of Politics 66*(4), 1203–1223.

Keller, S. and T. Mirer (1974). The Simple Act of Voting. *American Political Science Review 68*, 572–91.

Kling, J., J. Liebman, and L. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica*.

Krook, M. (2005). *Politicizing Representation: Campaigns for Candidate Gender Quotas Worldwide*. New York: Ph.D. diss Columbia University.

Matland, R. E. (1994). 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*, 273–292.

Mendelberg, T. (2001). *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton: Princeton University Press.

Mullainathan, S. and E. Washington (forthcoming). Sticking with your vote: Cognitive dissonance and political attitudes. *American Economic Journal: Applied Economics*.

Nosek, B., A. Greenwald, and M. Banaji (2007). The implicit association test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic Processes in Social Thinking and Behavior*.

Powley, E. (2007). Rwanda: The impact of women legislators on policy outcomes affecting children and families. *Background Paper, State of the World's Children*.

Rehavi, M. (2008). Sex and politics: Do female legislators affect state spending? *mimeo, Berkeley*.

Rooth, D.-O. (2007). Implicit discrimination in hiring: Real world evidence. *IZA Discussion Paper 2764*.

Rudman, L. and K. Fairchild (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology 87*(2), 157–176.

Rudman, L. and S. Kilianski (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin 26*(11), 1315–1328.

Thernstrom, S. and A. Thernstrom (1997). *America In Black And White: One Nation, Indivisible*. New York, NY: Simon and Schuster.

UNICEF (2007). *The State of The World's Children: Women and Children: The Double Dividend of Gender Equality*. New York: UN.

Weisberg, H. A Multidimensional Conceptualization of Party Identification.

Zaller, J. (1992). *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.
0.8

A Appendix

Note There are four vignette versions - Male (Tapan Das), Female (Sandhya Das), Invest in Irrigation and Invest in Water. There are six speech versions - three male voice recordings and three female voice recordings

Vignette *READ OUT: We will read a short description of the Pradhan of village CHANDI in district South 24 Parganas. We will ask you some 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.*

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]*.

Speech *READ OUT: 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 you 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.*

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

Water and sanitation quantity variable includes: 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 variable includes: handpumps are perennial, provide clean water, no stagnant water, have drainage and sanitation – no stagnant water, drainage facility. The Irrigation variable is 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 variable includes: 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 variable includes whether there is a bus stand and if bus stand has shelter. The Schools and other education facilities quantity variable includes: 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 facility variable includes: SSK, Anganwadi, primary schools, middle schools, libraries and secondary schools. The Schools and other education facilities quality variable includes: whether all primary schools have drinking water, latrines, blackboards and reading and math test scores. Health Quantity variable includes: 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. Health Quality variable includes: facility having tap or hand-pump water and an indicator for having a labor room. Fair Price Shop Quality measure includes: whether prices displayed, no bad behavior of shop keeper, and no complaint against shop.

Figure 1: Political Reservation and Female Leadership
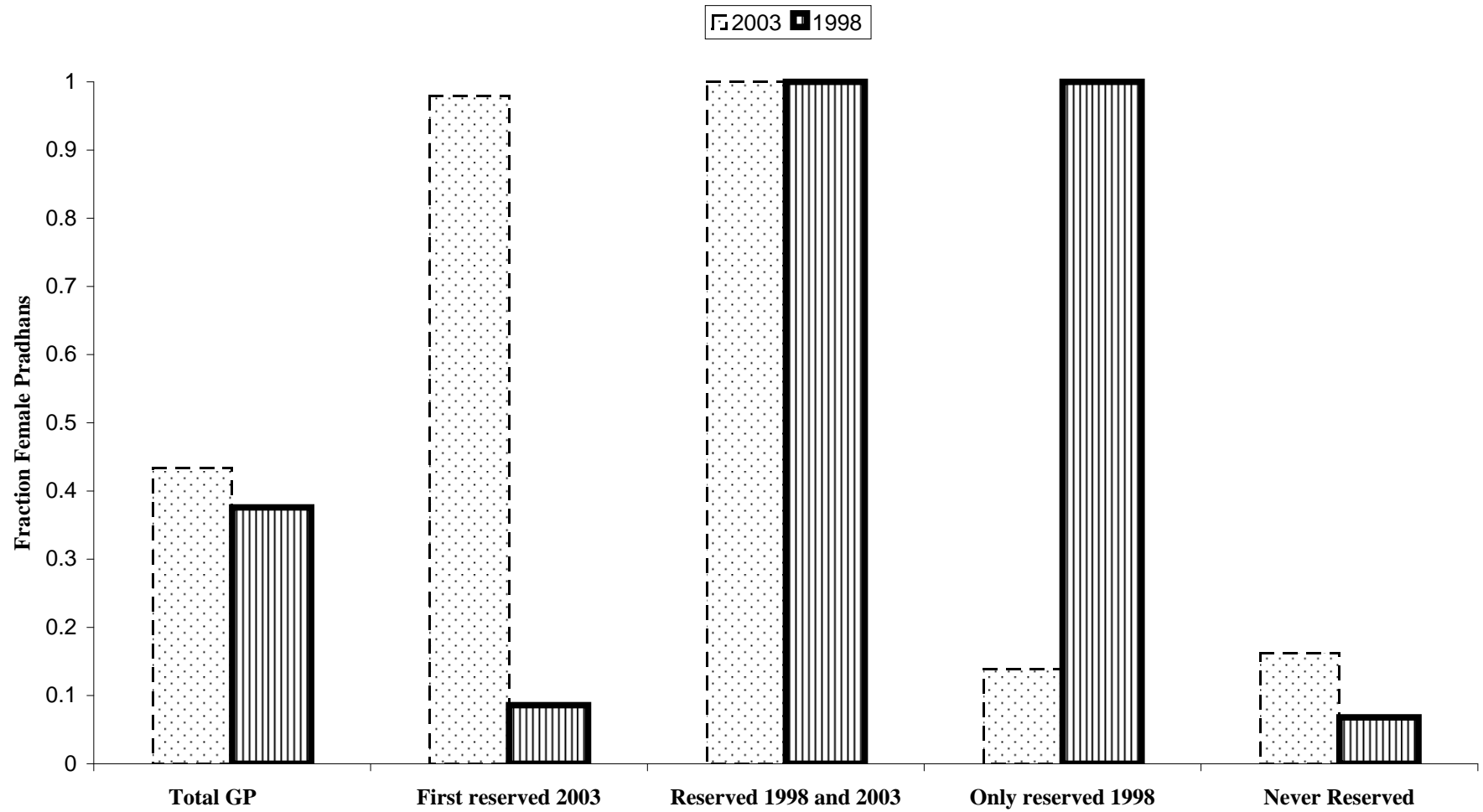
Note: here are 165 GPs total, 35 First Reserved 2003, 20 Reserved 1998 and 2003, 36 Only Reserved 1998 and 74 Never Reserved.

Table 1. Village-level Randomization check

Dependent Variable	First Reserved 2003	Reserved 1998 and 2003	Only Reserved 1998	Never Reserved	Diff: (1) and (2)	Diff: (3) and (4)
	(1)	(2)	(3)	(4)	(5)	(6)
Total Population	1,318	1,197	1,267	1,362	-102.279	-78.373
	(1318)	(1295)	(1418)	(1555)	(178.537)	(189.410)
Household Size	5.436	5.222	5.362	5.497	0.137	-0.059
	(0.612)	(0.449)	(0.571)	(2.454)	(0.093)	(0.135)
Literacy	0.387	0.351	0.391	0.373	0.014	0.021
	(0.119)	(0.129)	(0.121)	(0.137)	(0.032)	(0.017)
Fraction of Women Literate	0.287	0.265	0.296	0.275	-0.003	0.023
	(0.125)	(0.123)	(0.113)	(0.134)	(0.033)	(0.016)
Sex Ratio Under 6	1.093	1.034	1.046	1.036	0.206	0.015
	(0.590)	(0.276)	(0.218)	(0.226)	(0.113)	(0.025)
Share of SC / ST Population	0.432	0.502	0.450	0.471	0.004	-0.001
	(0.281)	(0.256)	(0.235)	(0.265)	(0.056)	(0.033)
Percentage of Irrigated Land	0.599	0.485	0.607	0.500	-0.038	0.054
	(0.321)	(0.349)	(0.323)	(0.352)	(0.054)	(0.049)
Village has a Bus or Train Stop	0.280	0.433	0.290	0.259	-0.130	0.061
	(0.451)	(0.500)	(0.456)	(0.439)	(0.097)	(0.054)
Pucca Road to the Village	0.110	0.300	0.176	0.194	-0.225	-0.006
	(0.314)	(0.462)	(0.383)	(0.397)	(0.082)	(0.052)
Number of Health Facilities	0.170	0.217	0.118	0.218	-0.036	-0.015
	(0.403)	(0.490)	(0.324)	(0.809)	(0.090)	(0.069)
Village has Tube Well	0.910	0.933	0.912	0.977	-0.022	-0.018
	(0.288)	(0.252)	(0.285)	(0.151)	(0.020)	(0.031)
Village has Hand Pump	0.124	0.000	0.100	0.057	0.000	0.013
	(0.331)	(0.000)	(0.302)	(0.233)	(0.000)	(0.012)
Village has Well	0.440	0.533	0.402	0.528	0.043	-0.066
	(0.499)	(0.503)	(0.493)	(0.500)	(0.082)	(0.074)
Village has Community Tap	0.065	0.050	0.042	0.010	0.022	-0.018
	(0.248)	(0.220)	(0.202)	(0.098)	(0.020)	(0.017)
Total Number of Schools	1.190	1.233	1.265	1.160	-0.167	0.149
	(0.895)	(0.810)	(0.889)	(0.910)	(0.160)	(0.111)
Overall Effect: F stat					0.300	0.600
Overall Effect: p value					*0.586*	*0.438*
N	105	60	108	222		

Notes:

1 "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.

2 Columns (1)-(4) report means with standard deviations in parenthesis. Columns (5)-(6) report tests of differences of means across columns (1) and (2) with standard errors in parentheses. Tests are based on regressions with block FE and standard errors clustered by GP.

3 N is the maximum number of observations and applies to all demographic variables. Infrastructure variables have between 454 and 478 total observations across GPs. Source of data is 1991 Census of India.

Table 2. Knowledge and Participation in Local Politics

	Knows the Name of Current Pradhan		Knows the Name of Previous Pradhan		Approached the Pradhan		Has Heard of Village Meeting		Attended Village Meeting Last Year	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A										
Ever Reserved	-0.082	-0.053	-0.074	-0.046	-0.033	-0.006	-0.028	-0.015	-0.002	-0.010
	(0.026)	(0.024)	(0.025)	(0.021)	(0.018)	(0.010)	(0.018)	(0.022)	(0.018)	(0.008)
Panel B										
First Reserved 2003	-0.125	-0.071	-0.009	-0.034	-0.044	-0.013	-0.042	-0.044	-0.006	-0.015
	(0.033)	(0.033)	(0.027)	(0.027)	(0.021)	(0.013)	(0.023)	(0.026)	(0.025)	(0.011)
Reserved 1998 and 2003	-0.110	-0.024	-0.138	-0.023	-0.064	0.001	0.037	0.042	-0.006	-0.005
	(0.043)	(0.037)	(0.048)	(0.033)	(0.026)	(0.014)	(0.030)	(0.033)	(0.027)	(0.011)
Only Reserved 1998	-0.020	-0.054	-0.097	-0.074	0.000	-0.004	-0.059	-0.024	0.004	-0.007
	(0.034)	(0.032)	(0.034)	(0.025)	(0.026)	(0.013)	(0.027)	(0.029)	(0.024)	(0.010)
Test: 2003 = both 1998 and 2003 = 1998	0.016	0.590	0.009	0.205	0.101	0.684	0.026	0.060	0.916	0.684
Mean of Never Reserved Sample	0.669	0.332	0.591	0.234	0.534	0.134	0.759	0.560	0.341	0.066
	(0.009)	(0.009)	(0.009)	(0.008)	0.009	0.006	(0.008)	(0.009)	0.009	0.004
N	6717	6780	6717	6780	6716	6779	6717	6780	6717	6780

Notes:

1 Each dependent variable is an indicator variable which equals one if the respondent answers in the affirmative. 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.

2 All regressions include (i) block fixed effects (ii) individual controls: age, age squared, household size, religion, caste, education, wealth (pca), landholdings, (iii) an indicator for the gender of the enumerator (iv) Census village controls: under 6 sex ratio, having a village hand pump and having a permanent road (v) a survey period indicator, which equals one if responses were collected during the first round (2006) of household visits. Standard errors are clustered by GP.

Table 3. Explicit and Implicit Preferences for Female Leaders

	Feeling Ladder				IAT (D-measure of bias against females)			
	Male versus Female Villager		Male versus Female Pradhan		Male/Female Names and Good/Bad		Male/Female Politician and Good/Bad	
	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Ever Reserved	0.098	-0.025	0.216	0.046	-0.006	0.0003	-0.010	-0.0084
	(0.077)	(0.075)	(0.110)	(0.107)	(0.032)	(0.043)	(0.034)	(0.037)
Panel B								
First Reserved 2003	0.142	0.014	0.237	0.054	-0.047	0.013	-0.007	0.016
	(0.100)	(0.100)	(0.156)	(0.144)	(0.043)	(0.051)	(0.050)	(0.049)
Reserved 1998 and 2003	-0.105	-0.136	0.092	-0.009	0.034	0.045	0.008	-0.001
	(0.125)	(0.107)	(0.157)	(0.145)	(0.042)	(0.070)	(0.052)	(0.054)
Only Reserved 1998	0.187	0.007	0.275	0.071	0.011	-0.059	-0.027	-0.040
	(0.105)	(0.098)	(0.150)	(0.143)	(0.046)	(0.051)	(0.045)	(0.051)
Test: 2003 = both 1998 and 2003 = 1998	0.136	0.439	0.622	0.898	0.274	0.230	0.830	0.624
Mean of Never Reserved Sample	0.691	0.181	1.446	0.560	0.134	-0.157	0.093	-0.079
	(0.052)	(0.047)	(0.067)	(0.064)	(0.025)	(0.026)	(0.027)	(0.025)
N	3511	3672	3511	3671	510	408	554	510

Notes:

1 The dependent variables in Columns (1)-(4) are the difference in the responses to the question on a scale of 1-10 " How do you feel towards a male y" and" How do you feel towards a female y", where y is either Villager or Pradhan. Dependent variables in columns (5)-(8) are the IAT D measure, defined as the difference in average response latencies between the two stereotypical and non-stereotypical blocks divided by the standard deviation of latencies in the two blocks. A positive D-measure indicates a stronger association of female and bad (i.e. a relative preference for men).

2 Reservation indicators are as defined in notes to Table 1 and 2. All columns include block fixed effects and individual controls listed in Table 2. Standard errors are clustered by GP.

3 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.

4 Different sample sizes across the columns reflects the fact that we administered the IAT and ladder questionaire to approximately 5 households within each village, but each respondent played one of three IATs. In addition, the IAT sample excludes data from pilot villages.

Table 4. Gender and Leadership: Implicit Associations

	Leadership/Domestic and Male/Female IAT (D-measure of bias against females)	
	Male	Female
	(1)	(2)
Panel A		
Ever Reserved	-0.077	0.0209
	(0.030)	(0.040)
Panel B		
First Reserved 2003	-0.086	0.122
	(0.040)	(0.053)
Reserved 1998 and 2003	-0.036	-0.106
	(0.047)	(0.072)
Only Reserved 1998	-0.094	-0.027
	(0.040)	(0.049)
Test: 2003 = both 1998 and 2003 = 1998	0.507	0.010
Mean of Never Reserved Sample	0.110	0.150
	(0.021)	(0.027)
N	477	357

Notes:

1 Dependent variables are the IAT D measure, defined in Table 3. A positive D-measure indicates a stereotype associating women with domestic activities and men

2 The regressions include the controls defined in Table 2, and standard errors are clustered by GP. The sample size reflects the sampling strategy explained in notes to Table 3.

3 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.

Table 5. Perception of Female Effectiveness as Leaders: Experimental Evidence (Speech and Vignettes)

	Average Coefficients											
	Average Effect		Pradhan is Effective		Cares about villagers' welfare		Approve of Pradhan		Perform Duties Well		Agree with Pradhan	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A												
Female Pradhan	-0.055	-0.035	-0.047	-0.029	-0.055	-0.003	-0.026	-0.056	-0.076	-0.043	-0.054	-0.044
	(0.027)	(0.031)	(0.032)	(0.034)	(0.033)	(0.032)	(0.029)	(0.033)	(0.036)	(0.042)	(0.039)	(0.027)
Female Pradhan * Ever Reserved	0.096	0.020	0.084	0.022	0.102	0.008	0.068	0.041	0.121	0.026	0.074	0.001
	(0.037)	(0.039)	(0.042)	(0.043)	(0.043)	(0.041)	(0.040)	(0.042)	(0.049)	(0.051)	(0.052)	(0.044)
Test: Female Pradhan + Female Pradhan *												
Ever Reserved	0.041	-0.014	0.037	-0.006	0.047	0.005	0.041	-0.015	0.045	-0.018	0.020	-0.044
	(0.024)	(0.023)	(0.027)	(0.026)	(0.027)	(0.025)	(0.027)	(0.026)	(0.032)	(0.029)	(0.035)	(0.035)
Panel B												
Female Pradhan * First Reserved 2003	0.118	-0.003	0.101	-0.011	0.128	-0.025	0.052	0.059	0.173	-0.026	0.045	0.053
	(0.047)	(0.048)	(0.052)	(0.052)	(0.056)	(0.050)	(0.049)	(0.055)	(0.063)	(0.064)	(0.071)	(0.058)
Female Pradhan * Reserved 1998 & 2003	0.097	0.050	0.108	0.045	0.111	0.056	0.089	0.042	0.096	0.055	0.085	0.047
	(0.061)	(0.060)	(0.064)	(0.070)	(0.069)	(0.069)	(0.062)	(0.064)	(0.083)	(0.073)	(0.079)	(0.071)
Female Pradhan * Only Reserved 1998	0.077	0.030	0.054	0.046	0.073	0.015	0.071	0.028	0.090	0.065	0.096	-0.078
	(0.048)	(0.046)	(0.056)	(0.053)	(0.054)	(0.050)	(0.056)	(0.051)	(0.062)	(0.058)	(0.067)	(0.066)
Test: FP* 2003 = FP* (both 1998 and 2003)												
= FP* 1998	0.757	0.671	0.666	0.556	0.663	0.491	0.852	0.869	0.486	0.348	0.809	0.204

Notes:

1 Columns (1) and (2) report a weighted average effect for all the dependent variables in speech and vignettes. Columns (3)-(12) report five separate average coefficients, where we average across the outcome in the speech and vignette regression (the dependent variables are standardized by subtracting the mean for the never reserved sample and dividing by the standard deviation). In all cases standard errors are computed by estimating the outcomes in a SUR framework.

2 All regressions include the controls defined in Table 2, 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 Tables 1 and 2 notes for definitions). Female Pradhan is an indicator for when the leader delivering the speech and vignette was female.

4 Test: FP* 2003 = FP* (both 1998 and 2003) = FP* 1998 reports the p-value from a Wald test of the equality of the coefficients on Female Pradhan interactions with First Reserved in 2003, Reserved 1998 and 2003 and Only Reserved 1998 respectively.

Table 6. Female Leadership: Experimental Evidence by Exposure Level

	Average Effect (speech and vignettes)					
	OLS			IV		
	All	Male	Female	All	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Female Pradhan	-0.042	-0.032	-0.045	0.025	0.038	0.026
	(0.031)	(0.051)	(0.037)	(0.047)	(0.104)	(0.054)
Female Pradhan * Ever Reserved	0.039	0.059	0.032	-0.042	-0.039	-0.020
	(0.038)	(0.066)	(0.047)	(0.056)	(0.127)	(0.068)
Female Pradhan * Ever Reserved * Knows Current or Previous Pradhan	0.042	0.050	-0.025	0.188	0.182	0.108
	(0.044)	(0.076)	(0.073)	(0.087)	(0.165)	(0.160)
Ever Reserved	0.017	-0.036	0.035	0.064	0.000	0.049
	(0.040)	(0.057)	(0.043)	(0.055)	(0.105)	(0.061)
Ever Reserved * Knows Current or Previous Pradhan	-0.023	0.000	0.037	-0.100	-0.039	0.006
	(0.041)	(0.055)	(0.057)	(0.082)	(0.139)	(0.120)
Female Pradhan * Knows Current or Previous Pradhan	-0.003	-0.029	0.027	-0.122	-0.121	-0.160
	(0.036)	(0.058)	(0.055)	(0.069)	(0.132)	(0.113)
Knows Current or Previous Pradhan	0.076	0.057	0.061	0.183	0.168	0.150
	(0.032)	(0.042)	(0.043)	(0.074)	(0.121)	(0.095)

Notes:

1 Coefficients are from an OLS or IV regression. The dependent variable is the average of the speech and vignette variables listed in Appendix Table 1. We estimate the average effect in a SUR framework as explained in Table 5.

2 All regressions include individual controls as defined in Table 2, and standard errors are clustered by GP. The regressions also include Ever Reserved (see Table 2 notes for definitions). "Knows Current or Previous Pradhan" is an indicator variable if the respondent knows the name of the current or previous pradhan. Female Pradhan is an indicator for when the leader delivering the speech and vignette was female.

3 The excluded instruments for "knows current or previous pradhan" include the indicator variables: knows MLA, reads newspaper, has heard of Gram Sabha, lives in the same village as pradhan and same caste as pradhan.

Table 7. Actual Pradhan Evaluation

	Average Effect		Pradhan is effective		Pradhan did a good job					
					looking after village needs		looking after your needs		making BPL lists	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
First Reserved 2003	-0.202	-0.072	-0.205	-0.090	-0.206	-0.109	-0.214	-0.045	-0.206	-0.087
	(0.057)	(0.059)	(0.064)	(0.070)	(0.062)	(0.063)	(0.060)	(0.064)	(0.067)	(0.061)
Reserved 1998 and 2003	0.014	-0.009	-0.029	-0.027	-0.033	-0.042	0.003	0.024	0.129	0.017
	(0.071)	(0.049)	(0.078)	(0.059)	(0.079)	(0.059)	(0.067)	(0.047)	(0.083)	(0.053)
Only Reserved 1998 (currently unreserved)	-0.009	0.029	0.006	-0.003	0.001	-0.011	-0.007	0.079	-0.070	0.048
	(0.059)	(0.054)	(0.068)	(0.066)	(0.068)	(0.062)	(0.060)	(0.054)	(0.063)	(0.054)
Test: 2003 = both 1998 and 2003 = 1998	0.008	0.270	0.025	0.511	0.022	0.372	0.005	0.187	0.003	0.086
Test: 2003 = both 1998 and 2003	0.009	0.329	0.052	0.419	0.053	0.377	0.007	0.300	0.001	0.140
N	6642	6568	6530	6315	6590	6363	6471	6323	6246	5893

Notes:

1 Average evaluation coefficients are from of a seemingly unrelated regression (SUR) including the four dependent variables listed in columns (3)-(10). All dependent variables are normalized by the mean and standard deviation of the never reserved sample.

2 All regressions include controls as defined in Table 2, and standard errors are clustered by GP.

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.

Table 8. Pradhan Performance: Public Goods, Bribes and Satisfaction

	Average Public Good Provision		Average Bribes	Average Satisfaction		Alignment with Female Preferences
	Quantity	Quality		Male	Female	
	(1)	(2)	(3)	(5)	(6)	(7)
First Reserved 2003	0.192	-0.043	-0.077	0.038	-0.004	0.521
	(0.070)	(0.046)	(0.032)	(0.040)	(0.043)	(0.279)
Reserved 1998 and 2003	0.039	-0.030	-0.058	-0.072	-0.040	0.659
	(0.061)	(0.052)	(0.030)	(0.052)	(0.045)	(0.358)
Only Reserved 1998 (currently unreserved)	0.097	-0.069	-0.033	-0.012	0.029	0.563
	(0.082)	(0.037)	(0.037)	(0.047)	(0.043)	(0.243)
Test: 2003 = both 1998 and 2003 = 1998	0.127	0.763	0.482	0.174	0.375	0.942
Test: 2003 = both 1998 and 2003	0.343	0.847	0.596	0.065	0.494	0.730

Notes:

 1 Average coefficients are from seemingly unrelated regressions (SUR) and the individual outcomes are defined in Appendix table A2 (for column (1)) and in Data Appendix for the rest. All outcome variables are normalized by the mean and standard deviation of the never reserved sample.

 2 The sample in Columns (1), (2) and (7) regressions are 495 villages, while columns (3)-(6) regressions use the household surveys. Household survey regressions include the set of controls defined in Table 2.

 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. We report the coefficients from the interaction of reservation status and the average difference between fraction of requests for goods in category i from women and from men. Categories analyzed are: drinking water, road improvement, education, irrigation, and other.

Table 9. Pradhan Characteristics

	Demographics						Political Characteristics					
	Age	Years of Education	Married	SC/ST	Wealth Index: Quartile 1	Wealth Index: Quartile 4	Average Effect	First Time as Pradhan in 2003 Term	GP Experience	Other Political Experience	Affiliated with CPI(M)	Spouse Ever Elected to Panchayat
	(1)	(2)	(3)	(4)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
First Reserved 2003	-5.780	-2.329	-0.184	0.047	0.170	-0.113	-0.089	0.266	-0.431	-0.195	-0.029	0.006
	(1.890)	(0.719)	(0.089)	(0.107)	(0.100)	(0.073)	(0.073)	(0.083)	(0.162)	(0.091)	(0.087)	(0.009)
Reserved 1998 and 2003	-6.692	-1.209	0.018	-0.023	0.089	0.019	0.056	0.036	-0.085	-0.081	0.198	0.047
	(2.305)	(0.792)	(0.084)	(0.138)	(0.121)	(0.118)	(0.085)	(0.128)	(0.215)	(0.130)	(0.091)	(0.054)
Only Reserved 1998 (currently unreserved)	0.718	0.550	-0.026	0.025	-0.073	0.032	-0.050	0.008	0.110	-0.104	-0.061	0.032
	(2.112)	(0.641)	(0.064)	(0.109)	(0.086)	(0.090)	(0.107)	(0.108)	(0.270)	(0.109)	(0.092)	(0.029)
Mean of Never Reserved Sample	41.194	10.278	0.889	0.514	0.208	0.208	-0.029	0.611	1.611	0.417	0.681	0
	(1.146)	(0.377)	(0.037)	(0.059)	(0.048)	(0.048)	(0.042)	(0.058)	(0.109)	(0.059)	(0.055)	(0)
Test: 2003 = both 1998 and 2003	0.721	0.229	0.085	0.659	0.574	0.278	0.147	0.084	0.129	0.416	0.049	0.438

Notes:

1 Wealth Index is based on a principal components analysis using the number of household assets. GP experience is the number of times Pradhan was elected to the GP council. Other Political Experience is an indicator for whether the Pradhan has or holds other political office, including: booth member, member of higher panchayat, MP, MLA, youth party president, district / block level posts of party, Mahila Samiti. All regressions include are block fixed effects and standard errors are clustered at the GP level.

2 Reservation types are defined in the notes of Table 1.

Table 10. Robustness

	Average effect	
	Male	Female
	(1)	(2)
I. Evaluation of New Pradhans		
First Time as Pradhan in 2003 Term	0.016	0.007
	(0.072)	(0.073)
II. Evaluation of SC Pradhans		
GP Reserved for SC in 2003	-0.100	-0.030
	(0.067)	(0.064)
GP Reserved for SC in 1998	-0.082	-0.052
	(0.058)	(0.053)
GP Reserved for SC in 2003 * Reserved for Woman in 1998	0.125	0.006
	(0.132)	(0.119)
III. Evaluation of Female Pradhans and SC Reservation		
First Reserved 2003 for Woman	-0.212	-0.070
	(0.074)	(0.075)
Reserved 1998 and 2003 for Woman	0.042	0.003
	(0.087)	(0.063)
Only Reserved 1998 for Woman	0.007	-0.028
	(0.086)	(0.076)
Reserved for Female in 2003 * Previously Reserved for SC	0.040	-0.007
	(0.132)	(0.131)
IV. SC Reservation on Female Leadership: Speech and Vignettes Results		
Panel A		
Female Pradhan	-0.046	-0.040
	(0.035)	(0.034)
Female Pradhan * Ever Reserved SC	-0.014	0.020
	(0.039)	(0.039)
Panel B		
Female Pradhan * Reserved for SC in 2003	0.004	0.017
	(0.042)	(0.045)
Female Pradhan * Reserved for SC in 1998	-0.028	0.030
	(0.047)	(0.048)

Notes:

1 The regressions in Panel I-III use the average evaluation coefficient for Pradhan (explained in notes to Table 7) as the dependent variable. In Panel I the explanatory variable of interest is whether the Pradhan was elected for the first time in 2003 and the sample is restricted to the set of never reserved Pradhans. In Panel II the explanatory variables of interest are whether the GP was reserved for SC in the 2003 and 1998 elections and the interaction of being reserved for a woman pradhan in 1998 and being reserved for an SC pradhan in 2003. In Panel III, the explanatory variables of interest are the women reservation status indicators and the interaction of "Previously Reserved for SC" indicator with the "First Reserved 2003" indicator. The regressions also include the interaction of "Only Reserved 1998" and "Reserved 1998 and 2003" with ""Previously Reserved for SC" indicator.

2 The regressions in Panel IV use the average perception of Pradhan effectiveness in the speech and vignettes (explained in notes to Table 5) as the dependent variable. The explanatory variables of interest are the indicators for female pradhan in the speech/vignette and the interaction of this indicator with the reservation for SC status of the GP. In addition, the regressions include the indicator variables for ever reserved for SC in Panel A, and reserved for SC in 2003 and 1998 in Panel B.

3 All regressions include individual controls as defined in Table 2, and standard errors are clustered by GP.

4 Reservation types are defined in the notes of Table 1.

Table 11: Pradhan Perceptions

	Satisfied with Current Life	Gender Caused Problem to do Duties	Does not plan to re-run for Pradhan
	(1)	(2)	(3)
Reserved for the firt time (either 1998 or 2003)		0.341	0.124
		(0.138)	(0.056)
First Reserved 2003	-0.470		
	(0.217)		
Reserved 1998 and 2003	0.379		-0.033
	(0.262)		(0.102)
Only Reserved 1998	-0.148		
	(0.216)		
Survey 2000		-0.147	-0.026
		(0.117)	(0.049)
Test: 2003 = both 1998 and 2003	0.006		
Mean of Never Reserved Sample	3.708		0.215
	(0.113)		(0.028)
N	161	108	327
Year of survey	2006-07	2000 (for Pradhans elected 1998), 2006-07 (for Pradhans elected 2003)	

Notes:

1 Satisfied with Current Life is on a scale of 1 to 5. Gender Caused Problem to do Duties and Does not plan to re-run for Pradhan are indicator variables.

2 In columns (2) and (3) we use surveys conducted in 2000 and 2006-07. The 2000 survey collected data on Pradhan's elected in 1998 and the 2006-07 interview includes Pradhans elected in 2003. Reserved for the first time is a dummy=1 if the GP was reserved for the first time in either 1998 and Pradhan was interviewd in 2000 or the GP was reserved for the first time in 2003 and the Pradhan was interviewed in 2006-07. The "Survey 2000" is equals one if the data comes from the 2000 survey. All other reservation variables are as defined in Table 1.

3 The regression in column (1) includes the sample of Pradhans elected in 2003, the regression in column (2) the Pradhans elected from women reserved GPs in 1998 and 2003 and column (3) inlcudes all Pradhans elected in either 1998 and 2003. (XXWHY IS

4 In column (1) the p-value is from a Wald test of the equality of the coefficients on First Reserved in 2003 and Reserved 1998 and 2003.

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

	Speech							Irrigation Vignette				Water Vignette			
	Pradhan is effective	Address-ed villager satisfact-orily	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	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)	(12)	(13)	(14)	(15)
I. Males															
Panel A															
Female Pradhan	-0.066	-0.111	-0.082	-0.041	-0.034	-0.070	-0.082	-0.061	-0.059	-0.145	-0.077	0.037	0.029	0.045	0.062
	(0.040)	(0.041)	(0.042)	(0.042)	(0.042)	(0.038)	(0.042)	(0.054)	(0.055)	(0.054)	(0.060)	(0.064)	(0.065)	(0.053)	(0.053)
Female Pradhan * Ever Reserved	0.098	0.114	0.129	0.099	0.054	0.144	0.129	0.155	0.170	0.124	0.096	-0.015	-0.015	0.019	0.037
	(0.053)	(0.055)	(0.055)	(0.056)	(0.056)	(0.052)	(0.053)	(0.072)	(0.075)	(0.072)	(0.077)	(0.080)	(0.082)	(0.071)	(0.076)
Panel B															
Female Pradhan * First Reserved 2003	0.156	0.139	0.199	0.191	0.053	0.181	0.180	0.140	0.169	0.074	0.035	-0.027	-0.032	0.035	0.046
	(0.068)	(0.067)	(0.068)	(0.071)	(0.067)	(0.070)	(0.071)	(0.095)	(0.096)	(0.085)	(0.092)	(0.101)	(0.109)	(0.100)	(0.108)
Female Pradhan * Reserved 1998 & 2003	0.096	0.064	0.092	0.064	0.060	0.122	0.135	0.173	0.219	0.168	0.142	0.052	0.032	-0.029	0.031
	(0.081)	(0.093)	(0.088)	(0.095)	(0.100)	(0.092)	(0.084)	(0.093)	(0.112)	(0.107)	(0.106)	(0.118)	(0.120)	(0.115)	(0.123)
Female Pradhan * Only Reserved 1998	0.043	0.117	0.083	0.035	0.054	0.125	0.084	0.165	0.149	0.153	0.130	-0.043	-0.027	0.022	0.025
	(0.070)	(0.073)	(0.071)	(0.075)	(0.075)	(0.066)	(0.064)	(0.100)	(0.104)	(0.103)	(0.107)	(0.097)	(0.093)	(0.081)	(0.095)
Test: FP* 2003 = FP* (both 1998 and 2003) = FP* 1998	0.349	0.747	0.281	0.167	0.998	0.756	0.429	0.953	0.870	0.637	0.549	0.724	0.863	0.884	0.985
N	6715	6716	6715	6714	6551	6715	6716	3173	3173	3172	3116	3544	3544	3543	3503
II. Females															
Panel A															
Female Pradhan	-0.051	-0.029	-0.055	0.008	-0.089	-0.081	-0.072	-0.058	-0.010	-0.098	-0.082	0.031	0.090	0.006	0.032
	(0.045)	(0.046)	(0.046)	(0.045)	(0.045)	(0.045)	(0.045)	(0.058)	(0.056)	(0.055)	(0.057)	(0.047)	(0.047)	(0.041)	(0.047)
Female Pradhan * Ever Reserved	0.010	0.030	0.019	-0.026	0.075	0.057	0.042	0.018	-0.049	-0.049	-0.036	0.016	0.002	0.003	0.008
	(0.055)	(0.057)	(0.057)	(0.056)	(0.058)	(0.057)	(0.056)	(0.072)	(0.071)	(0.075)	(0.076)	(0.060)	(0.061)	(0.061)	(0.062)
Panel B															
Female Pradhan * First Reserved 2003	-0.034	-0.026	-0.033	-0.084	0.071	0.004	-0.001	0.002	-0.070	-0.003	-0.020	0.012	0.016	0.064	0.090
	(0.068)	(0.074)	(0.069)	(0.066)	(0.078)	(0.071)	(0.075)	(0.083)	(0.082)	(0.093)	(0.088)	(0.076)	(0.086)	(0.078)	(0.086)
Female Pradhan * Reserved 1998 & 2003	0.025	0.079	0.065	0.004	0.075	0.083	0.060	0.051	-0.001	-0.029	0.093	0.035	0.044	0.035	-0.115
	(0.075)	(0.079)	(0.086)	(0.091)	(0.078)	(0.093)	(0.072)	(0.095)	(0.102)	(0.094)	(0.105)	(0.093)	(0.090)	(0.114)	(0.103)
Female Pradhan * Only Reserved 1998	0.049	0.062	0.050	0.019	0.087	0.099	0.079	-0.006	-0.084	-0.133	-0.149	0.013	-0.032	-0.079	-0.001
	(0.066)	(0.066)	(0.069)	(0.067)	(0.068)	(0.066)	(0.066)	(0.104)	(0.098)	(0.108)	(0.110)	(0.079)	(0.075)	(0.084)	(0.071)
Test: FP* 2003 = FP* (both 1998 and 2003) = FP* 1998	0.496	0.392	0.410	0.314	0.977	0.422	0.570	0.846	0.741	0.522	0.183	0.970	0.721	0.341	0.202
N	6771	6772	6771	6770	6372	6770	6770	2869	2869	2869	2740	3908	3908	3902	3797

Notes:

1 Dependent variables in columns (1)-(4) and (6)-(7) are originally on a scale of 1 to 10, while dependent variables in columns (5) are indicator variables, then normalized by the mean and standard deviation of the never reserved sample. In the vignette experiments, Effectiveness of leader and cares about villager welfare are orginally 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 All regressions include individual controls as defined in Table 2, 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 Tables 1 and 2 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 Female Pradhan is a dummy variable indicating that the voice delivering the speech was female or the Pradhan depicted in the vignettes was female.

Table A3. Effect of Female Leadership on Public Goods Quantity

			Coefficients on:		
	N	GP reserved only in 2003	GP previously reserved in 1998	GP reserved in 2003 and 1998	Mean of Never Reserved
At Least One New Tubewell was Built	495	0.152	0.073	0.160	0.365
		(0.066)	(0.063)	(0.088)	(0.482)
At Least One Tubewell was Repaired	482	0.208	0.130	0.080	0.628
		(0.067)	(0.064)	(0.089)	(0.484)
At Least One Drainage/Sanitation Facility was Built	495	0.053	-0.113	0.052	0.428
		(0.067)	(0.059)	(0.091)	(0.496)
At Least One Drainage/Sanitation Facility was Repaired	396	0.150	-0.017	0.032	0.178
		(0.067)	(0.062)	(0.071)	(0.384)
At Least One Irrigation Pump was Built	495	0.137	0.005	-0.013	0.180
		(0.053)	(0.051)	(0.050)	(0.385)
At Least One Irrigation Pump was Repaired	319	0.110	-0.078	-0.005	0.417
		(0.092)	(0.086)	(0.123)	(0.495)
Number of metal roads built or repaired since 2003	495	0.274	0.046	0.079	0.118
		(0.117)	(0.070)	(0.065)	(0.448)
Number of transportation related infrastucture (Bus Stop, bus service, taxi)	495	0.074	0.250	0.303	1.302
		(0.175)	(0.160)	(0.225)	(1.201)
At Least One Educational Facility was Built	495	0.053	-0.030	0.026	0.117
		(0.042)	(0.036)	(0.055)	(0.322)
At Least One Educational Facility was Repaired	465	0.165	0.039	0.001	0.296
		(0.072)	(0.069)	(0.097)	(0.458)
At Least One Community Education Center	495	-0.007	0.030	-0.001	0.009
		(0.010)	(0.023)	(0.009)	(0.095)
There is a NGO Child center/Creche	495	-0.045	-0.039	-0.027	0.045
		(0.016)	(0.021)	(0.023)	(0.208)
Number of Health Facilities (PHC, Health SubCenter)	495	-0.025	0.027	-0.005	0.257
		(0.049)	(0.052)	(0.084)	(0.468)
At Least One Health Facility was Built	495	0.011	-0.004	-0.018	0.014
		(0.015)	(0.014)	(0.009)	(0.116)
At Least One Health Facility was Repaired (0 if no fac)	495	0.061	0.016	0.047	0.009
		(0.023)	(0.016)	(0.024)	(0.095)
Number of Trained Dais, Untrained Dais and Private Doctors	495	-0.069	-0.158	0.384	1.014
		(0.232)	(0.226)	(0.423)	(2.012)
Average Effect		0.192	0.039	0.097	
		(0.070)	(0.061)	(0.082)	

Notes:

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

2 Reservation status defined in Table 1.

3 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.