

POWERFUL WOMEN: DOES EXPOSURE REDUCE BIAS? *

Lori Beaman

Raghabendra Chattopadhyay

Esther Duflo

Rohini Pande

Petia Topalova

Abstract

We exploit random assignment of gender quotas for leadership positions across Indian village councils to show that prior exposure to a female leader is associated with electoral gains for women. After ten years of quotas, women are more likely to stand for, and win, elected positions in councils required to have a female chief councilor in the previous two elections. We provide experimental and survey evidence on one channel of influence - changes in voter attitudes. Prior exposure to a female chief councilor improves perceptions of female leader effectiveness and weakens stereotypes about gender roles in public and domestic spheres.

*The views expressed in this paper are those of the authors and do not implicate the International Monetary Fund, its management, or Executive Board. We thank Ash Center Harvard, Nike Foundation, MIT, YCIAS Yale, and UNICEF for funding, Alexandra Cirone, Catherine Lee and Kudzai Takavarasha for research assistance, and Prasad Chakraborty and the SRG team for outstanding fieldwork. We also thank Abhijit Banerjee, Tim Besley, Shawn Cole, Dominic Leggett, Elhanan Helpman, Larry Katz, four anonymous referees and numerous seminar participants for comments.

I. Introduction

In 2008, women accounted for 18.4 percent of parliamentarians worldwide, and a woman headed the government in only thirteen countries (UNIFEM 2008). These gender disparities do not reflect legal restrictions – women can vote, support candidates, and run for office in almost every country. Rather, many suggest that in both rich and poor countries, women’s access to public office is, at least partly, restricted by voter and party bias that favors male politicians.

The belief that the gender of policy-makers may have important consequences for policy decisions, and the possibility that such bias, if present, may be malleable has led policy-makers to emphasize the importance of supporting early cohorts of female politicians. The assumption is that once voters learn that women can lead effectively, gender bias in politics will diminish. This possibility has led more than a hundred countries to introduce affirmative action policies for women in public office, either by law or through voluntary actions of political parties, over the last two decades (Krook 2005; Dahlerup 2006).¹ While these policies have significantly increased female representation in politics (Jones 2004) and have often altered subsequent policy-making (Chattopadhyay and Duflo 2004; Powley 2006), little is known about their impact on women’s electoral prospects and voter attitudes towards female leaders.

While in many settings exposure to members of another group creates “empathy,”² whether mandated exposure to female leaders can successfully alter social norms or perceptions of women’s ability to lead remains debated. A first reason for doubting their effectiveness is that voters may dislike quotas which restrict their choices, and therefore women leaders (Thernstrom and Thernstrom 1997). Another is that voters may perceive gender quotas as violating social norms and potentially reducing the value of traditionally male activities (Goldin 2002). As a result, quotas may precipitate a backlash against female leaders and strengthen taste-based discrimination (on this, also see Boisjoly et al [2006]; Rudman and Fairchild [2004]).

An important counterargument is that mandated exposure to women leaders informs voters on women’s ability to lead. Consider the case where voters are risk-averse and no women were

¹Details of quotas by country are available at <http://www.quotaproject.org>. A different, possibly complimentary, rationale for affirmative action is if policy preferences differ by gender and female turnout is relatively low, female politicians are required to ensure that women’s preferences are represented in government.

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

initially elected leaders (possibly because of some small initial taste discrimination). Since voters gain relatively more information on the male leaders they elect, they will perceive the choice of female leaders as risky and continue to favor male leaders causing biased perceptions about women’s effectiveness as leaders to persist (Aigner and Cain 1977). In such settings mandated exposure can reduce statistical discrimination and improve perceptions of female leaders’ effectiveness (unless women make incompetent leaders and being exposed to them causes voters to update negatively).³

In this paper we exploit random variation in mandated exposure to female leaders across village councils in India to provide direct evidence on these issues. Our data comes from the Indian state of West Bengal. Each village council in this state consists of a set of elected councilors. These councilors elect a chief councilor, or Pradhan. In every council election since 1998, a third of councilor positions in each council and a third of Pradhan positions across councils in a district have been randomly “reserved” for women. Only women can run for election to a reserved position. We take advantage of this randomization to examine the causal impact of mandated exposure on electoral outcomes, as well as villager attitudes and perceptions of women leaders.

We start by documenting significant electoral gains in the May 2008 election for women candidates contesting unreserved councilor and Pradhan positions in councils where the Pradhan position was reserved for a woman in the previous two electoral cycles (i.e. for ten years). Electoral data on unreserved councilor positions across all village councils in a West Bengal district shows that, relative to councils which never had a reserved Pradhan, almost twice as many women stood for, and won, these positions in councils where the Pradhan position had been reserved for women in the previous two elections. Data from a larger set of six West Bengal districts shows similar gains for women contesting Pradhan elections in councils where the Pradhan position is currently unreserved. In the May 2008 election the share of female Pradhans was 11% in councils where the Pradhan position had never been reserved and 18.5% in councils that were continuously reserved for a female Pradhan between 1998 and 2008.

An initial increase in female leadership can enhance the electoral prospects of subsequent generations of female leaders through multiple channels. In this paper we examine in detail one channel

³One reason why quotas could lead to less competent female leaders is if, as in Coate and Loury (1993), they work less in anticipation of voter perceptions that quotas lead to less competent politicians being elected.

(which is not exclusive of others): changes in voter attitudes towards female leaders. We use survey data to show that repeated exposure improves voter evaluation of female Pradhans. Relative to Pradhans in councils where the Pradhan position has never been reserved, female Pradhans in councils reserved for a female Pradhan for the first time receive worse evaluations. However, this is not true for women elected Pradhan in councils reserved for a female Pradhan the second time. This improvement in leader evaluation provides an explanation for the election results: subsequent to the improved ratings of female leaders in the second electoral cycle, more women contest and win village council elections in the third electoral cycle.

These results are consistent with an initial voter bias against female leaders which decreases with exposure. However, they are also consistent with changes in the selection, or behavior, of female leaders over time. While we do not find *prima facie* evidence that observable differences in the characteristics or actions of female leaders explain the results, we cannot rule out unobservable differences. Thus, our next step is to use experimental data to directly measure bias against female leaders and test whether exposure to a female leader changes voter attitudes.

Villagers were asked to evaluate the effectiveness of hypothetical leaders as described through vignettes and recorded speeches. The only variation across respondents was that leader gender was experimentally manipulated. In every village, half the respondents received “male” politician and the rest “female” politician vignettes and speeches. All other aspects of the speech and vignette were identical. We find that exposure to a female leader radically altered male villagers’ perceptions of female leader effectiveness. Men living in villages that had never been reserved judged the hypothetical leader as significantly more effective when the leader’s gender was experimentally manipulated to be male (rather than female). The evaluation gap disappears in currently or previously reserved villages. Among female villagers we observe a smaller (insignificant) bias which does not appear to be affected by the reservation status of the village.

We also conducted a series of computer-based Implicit Association Tests (IATs). The IAT is an experimental method, widely used in social psychology, which relies on the idea that respondents who more easily pair two concepts in a rapid categorization task more strongly associate those concepts (Nosek, Greenwald, and Banaji 2007). We first measured gender-occupation stereotypes by an IAT which examined the strength of association of male and female names with leadership

and domestic tasks. Exposure to female leaders (through reservation) increased the likelihood that male villagers associated women with leadership activities (as opposed to domestic activities).

We then used an IAT designed to measure taste, i.e. the association of men and women leaders with concepts of good and bad. In contrast to the gender-occupation IAT, we found strong same gender preference that was unaffected by reservation. Further, in survey responses both genders state an explicit distaste for female leaders (relative to male leaders). This explicit distaste is unaffected by reservation; if anything, we observe a backlash effect among men.

Overall, our results suggest that while deep preferences and social norms remain difficult to erode, beliefs on effectiveness are much more malleable, and they play a role in the voting decision. In the setting we study, we see an improvement in voter perceptions of female leaders, followed by electoral gains for women. This suggests that the use of political affirmative action, which causes voters to acquire information on the ability of traditionally disadvantaged groups, can durably influence political outcomes.

The rest of this paper is structured as follows. Section II describes the institutional context of political reservation, and Section III how it has affected electoral outcomes. Section IV analyzes how villagers' evaluation of their own leader changes with reservation, and Section V provides experimental evidence on the change in voters' attitudes. Section VI concludes.

II. Institutional Context

India has had universal franchise since Independence and many prominent elected female leaders. However, the average share of women in national and state legislatures is only 10 percent. An important policy response to this gender disparity in female political leadership was implemented in the 1993 constitutional amendment. This amendment mandated a three-tier elected local self-government institution (the Panchayat) and instituted gender quotas in all three tiers. Gender quotas have significantly increased local female leadership across Indian villages, and the number of elected female village leaders nationwide is now close to 40 percent. Below we describe the institution of political reservation as implemented in our study state of West Bengal and some summary statistics on how it altered the profile of leaders.

II.A. Village Councils and Political Reservation

In each Indian state, the Panchayat consists of a system of village (Gram Panchayat), block (Panchayat Samiti), and district (Zilla Parishad) councils. Council elections occur every five years. We focus on the village council, which is responsible for the provision of village infrastructure (such as public buildings, water, and roads) and for identifying government program beneficiaries. The main source of financing for its activities is the State government.

Unlike many states, which introduced Panchayats after the 1993 constitutional amendment, West Bengal has had an active elected Panchayat since 1978. A West Bengal Gram Panchayat (GP) has, on average, 10,000 voters spread across multiple villages. Each GP is divided into electoral wards, with a councilor directly elected from each ward by plurality rule. Councilors elect (from among the set of councilors) a chief village councilor, Pradhan. The 1993 amendment required that at each election one-third of ward councilor positions in each council and one-third of Pradhan positions in every district be reserved for women: only women can contest, and be elected, in a reserved position.

GP council decision making is by majority voting. The Pradhan is the only full-time council member and exercises significant control over the final council decisions (Besley, Pande, and Rao 2007). For this reason, we focus on the role of Pradhan reservation in affecting future electoral outcomes for women and public opinion towards female leaders.⁴

Pradhan reservation, for women and two disadvantaged minorities (Scheduled Castes (SC) and Scheduled Tribes (ST)), was introduced in a 1998 modification of the West Bengal Panchayat Constitution Rule (Government of West Bengal 1998). The rule requires that prior to an election GPs in a district be randomly assigned across three lists: Reserved for SC, Reserved for ST, and Unreserved.⁵ These lists are redone at every election to ensure that no GP is reserved for SC or ST in two consecutive elections. Within a list GPs are ordered by serial number. In 1998, every third GP starting with number one on each list was required to be reserved for a woman, and in 2003 every third GP starting with number two on each list. We reconstructed the 2003 reservation

⁴Following a referee suggestion we examined, but did not find, any impact of ward reservations on voters' attitudes.

⁵The randomization is on the basis of GP serial number rank and tables provided in the electoral law. The extent of SC and ST reservation in a district is proportional to their population share.

list using the GP serial numbers and electoral law tables and found that the rule held, with no exception (the same was true in 1998, see Chattopadhyay and Duflo [2004]).

This assignment rule implies that Pradhan reservation for women is random with implicit stratification by SC/ST and administrative block (since GP serial number starts with a block identifier). In addition, a GP may be reserved for women twice in a row; for instance, if it was ranked first on a list in 1998 but second in 2003.

In Table I we use 1991 census data (i.e. from before reservation was introduced) for the 495 villages for which we collected survey and experimental data to confirm that the randomization procedure resulted in a balanced sample. We assign villages to one of four categories: Never Reserved, Only Reserved in 1998, Only Reserved in 2003, and Reserved in 1998 and 2003. Village characteristics as of 1991 are not jointly correlated with the reservation assignment of GPs (see p-values in columns (5) and (6) of Table I). There are statistically significant differences by reservation status for only three of the thirty variables: sex ratio under 6, presence of hand pumps, and presence of a permanent approach road.

Below, as background to our analysis, we describe how reservation increased female representation in leadership positions.

II.B. Political Reservation and Female Representation

With the exception of our analysis of Pradhan electoral outcomes, our study uses data from the 165 GPs located in Birbhum district, West Bengal. Birbhum, which is situated 200 km from the capital Kolkata, is a largely rural and fairly poor district.

In Figure I we use data on the reservation status and Pradhan electoral outcomes for Birbhum GPs in 1998 and 2003 to describe how GP reservation altered the extent of female leadership. Roughly 44% of all GPs (74 GPs) were never reserved, 21% were reserved once (36 and 35 GPs in 1998 and 2003, respectively) and 12% (20 GPs) were reserved in both elections. Turning to trends in female leadership, Pradhans in GPs reserved for women are always female.⁶ While few women are elected Pradhans in never-reserved GPs, this number increased from 7% in 1998, to 15% in

⁶Due to political disturbances one reserved GP did not have a Pradhan at the time of our survey.

2003, and by an additional 2% in 2008 (not shown in Figure 1). The increase between 1998 and 2003 is similar across never-reserved GPs and those only reserved in 1998 (16% in never-reserved and 14% in those only reserved in 1998).

Electoral data also show that reservation did not lead to a significant incumbency advantage for female leaders. Similar to other Indian elections, we observe significant incumbency disadvantage in both reserved and unreserved GPs. In 2003, only 5.6% of incumbent Pradhans in both reserved and unreserved positions were reelected. In 2008, the fraction of reelected incumbents was even lower at 2% and 3% in reserved and unreserved GPs, respectively.

We conducted household surveys among all 1998 and 2003 Pradhans in Birbhum. In Table II we use these data to examine whether, and how, male and female Pradhans differ. Column (6) of Table II shows that, relative to male Pradhans, female Pradhans are younger, less educated, less likely to be married and from poorer families. Male Pradhans also have more political experience, both as ward councilor and Pradhan. However, contrary to the popular claim that women Pradhans simply function as “shadows” for their husband (who are disqualified from running), very few women Pradhans are spouses of former male Pradhans.

To summarize, reservation randomly exposed villagers to female leaders who differ from their male counterparts along multiple dimensions. The extent of exposure was significantly higher in twice reserved GPs. Against this background, we now examine whether reservation-induced exposure to a Pradhan affected the subsequent electoral prospects of women in unreserved Pradhan and ward councilor positions.

III. Does Reservation Improve Women’s Future Electoral Prospects?

An important aim of the paper is to examine whether the electoral impact, if any, of political reservation is related to changes in voter attitudes towards female leaders. We use survey and experimental data collected in the 165 GPs in Birbhum district to measure voter attitudes, and we examine ward level electoral outcomes in this district as well. In addition, to be able to draw statistically significant conclusions on the impact of reservation on the impact of Pradhan level

elections, we expanded our sample to six districts (including Birbhum).⁷

As a precursor to the regression analysis, Figure II shows the main findings for the 2008 Pradhan and ward councilor elections. We distinguish among three GP categories (based on Pradhan reservation status): Never Reserved, Reserved once, and Reserved 1998 and 2003. For each category Figure II shows the share of female Pradhans elected in unreserved GPs (across the six districts), and the share of women who contested, and won, a non-reserved ward councilor seat (in Birbhum). In the Pradhan sample, we see that women were elected Pradhan in roughly 10% of the GPs which had no prior history of reservation. The share increases to about 13% for GPs reserved once, and to 17% for GPs reserved twice.

The fraction of unreserved ward council seats contested, and won, by women in never-reserved GPs in Birbhum was just 5%. While the number remains similarly low among GPs which had been reserved only once (in either 1998 or 2003) we see a doubling of the fraction female ward seat contestants and winners in GPs which had been reserved in the last two elections.

III.A. Regression Results: Empirical Strategy

We now investigate the robustness of the trends observed in Figure II in a regression framework. We expand our sample to include data from both the 2003 and 2008 GP elections (reservation was introduced in 1998). As described earlier, our Pradhan regressions use data from unreserved GPs in six districts and our ward councilor regressions data from unreserved ward council seats across the 165 GPs in Birbhum.

Randomization of reservation assignment allows us to study its reduced form effect by comparing the means of outcomes of interest across GPs with different reservation status. All reserved GPs, but relatively few unreserved and previously reserved GPs, have a female Pradhan (Figure I). Hence, the reduced form effect we estimate is close to what would be obtained by instrumenting for Pradhan's gender by the reservation status of the GP.

⁷Only twenty GPs in Birbhum were reserved for women in both 1998 and 2003, and 35 were reserved in 2003 (Figure I). Since not all of these are unreserved in 2008, the sample becomes too small to establish conclusive findings (the point estimate for Birbhum only are consistent with what we obtain in the full sample). The sample includes the districts of Nadia, Howrah, Hooghly, Birbhum, South 24 Paraganas and Burdwan. As election data is only kept in the districts, we expanded our sample by sending surveyors to districts close to Kolkata and Birbhum (Birbhum is roughly 200 kms from Kolkata) for logistic reasons, and obtained complete data for six districts.

For both ward councilor and Pradhan elections, we first estimate the following regression in the 2003 election data:

$$(1) \quad y_{igj} = \beta_1 R_{g1} + \alpha_j + \epsilon_{ig}$$

where y_{igj} is a dummy for whether the elected representative (and candidate in the case of ward councilor elections) i in GP g and block or district j is a woman, R_{g1} is an indicator for whether the GP was reserved in 1998. α_j denotes district dummies in Pradhan election regressions and block dummies in ward councilor election regressions. We report robust standard errors in Pradhan election regressions. In ward councilor regressions we cluster standard errors by GP, since each GP has multiple ward councilors.

In the 2008 data, we estimate the following regression:

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

R_{g1} and R_{g2} are indicator variables for the GP being reserved only in the first and second electoral cycle respectively (i.e. only in 1998 and only in 2003). R_{g2and1} is an indicator for the GP being reserved twice (in 1998 and 2003). Otherwise, the specification is identical to the previous regression. Finally, we estimate a regression where we stack the 2003 and 2008 data:

$$(3) \quad y_{igj} = \beta_2 R_{go} + \beta_{2and1} R_{g2and1} + \lambda_t + \alpha_j + \epsilon_{ig}$$

R_{go} is an indicator for GPs which were only reserved once (either in 1998 or in 2003). λ_t is an indicator for the 2003 election. Since there are multiple observations per GP for the Pradhan and ward councilor regressions we always cluster the standard errors by GP.

III.B. Results

We start by examining electoral outcomes for Pradhan elections in 2003 and 2008. Column (1) of Table III examines the 2003 election. Roughly 9% of the unreserved GPs elected female Pradhans, and there is no discernible impact of prior reservation status. In column (2) we consider the 2008 election. Once again, being reserved for the first time in the previous (2003) election does not affect female electoral success; however, being reserved in the previous two elections does. There is also some evidence that reservation one cycle ago (in the 1998 election) matters. Finally, in column (3) we combine the data from 2003 and 2008 and run the stacked regression. We continue to see a strong effect of being twice reserved. In contrast, the overall effect of being reserved only once (pooled across 1998 and 2003) is now insignificant. In separate regressions (not reported here), we have estimated the regression for 2008 elections excluding GPs where the incumbent was reelected Pradhan (lack of data on 1998 Pradhan names for all districts means we cannot control for 1998 winner status). We find a noisier and somewhat smaller (5.4%), but statistically indistinguishable, effect. This suggests that at least some of the longer term impact of reservation is due to spillovers on newcomers.

The influence of reservation on Pradhan electoral outcomes is important since Pradhans are responsible for most local policy decisions. That said, electoral results for ward councilors are also of independent interest since, unlike Pradhans, they are directly elected by villagers.

We, therefore, next examine the impact of prior Pradhan reservation on whether women contest, and win, in unreserved ward councilor elections in Birbhum. Columns (4) and (7) of Table III show no impact of 1998 GP reservation on the fraction of women who ran for, and won, ward councilor positions in the 2003 election. Columns (5) and (8) consider the impact of previous GP reservation in 2008. In non-reserved wards belonging to GPs that had not been reserved for a female Pradhan in 1998 or 2003, 4.9% of both the candidates and the elected ward councilors were women (33% of the wards are reserved, so roughly 38% of all ward councilors are women). As in the 2003 election, these numbers are unchanged for wards in GPs that were reserved only once. However, we once again see a striking difference for wards in GPs where the position of Pradhan was reserved both

in 1998 and 2003. Specifically, relative to wards in never-reserved GPs, the proportion of female candidates in unreserved wards in twice-reserved GPs increased by 3.7 percentage points and the proportion of women elected more than doubled (10.7% versus 4.9%). These results (not reported here) are robust to excluding ward seats where the 1998 and/or 2003 Pradhans contested and were reelected.

Finally, columns (6) and (9) report results where we stack the 2003 and 2008 data. The conclusions are similar, and we gain some precision: after one cycle, there are no more women candidates or elected in unreserved seats. After two cycles, there are about twice as many.⁸

To check that our ward level results are not sensitive to the sample size of GPs (in particular, there are only 20 GPs that were reserved both in 1998 and in 2003) we computed Fischer exact p-values using randomization inference for the main results in columns (5)-(6) and (8)-(9) (Rosenbaum 1996; Imbens and Rosenbaum 2005).⁹ We can reject the null of no effect at the 5% level for the coefficient on Reserved in 1998 and 2003 in columns (5) and (6) and at the 10% level in columns (8) and (9). These results are consistent with the t-statistics based on clustered standard errors and are available from the authors.

Taken together, these results suggest that reservation significantly improved women’s electoral prospects. They are also consistent with Bhavnani (forthcoming), who finds that previous ward-level reservation for women in Mumbai improved their future representation in unreserved wards.

IV. Voter Beliefs as a Pathway to Electoral Success

Reservations may have facilitated the entry of women in politics through multiple channels. First, female Pradhans may act as important role models and mentors. Their presence may have directly motivated other women to participate in politics and made the public sphere less intimidating for women. Chattopadhyay and Duflo (2004) show that women are more likely to attend, and speak during, village meetings in reserved GPs. Second, female Pradhans may have also helped

⁸We have also estimated regressions where we include an indicator for 2008 Pradhan reservation status and the interactions between that variable and previous reservation status. Women are more likely to contest and win unreserved ward seats if the GP is reserved for a female Pradhan. The effect at the ward level also persists: women are differentially more likely to compete for ward councilor positions in GPs twice previously reserved.

⁹For 2000 draws, the treatment “status” is randomly assigned to all GPs, in proportion to what is observed in the data. The distribution of the “effect” size from each draw is used to create the p-value. We also did the analysis dropping each GP one by one and found similar results.

create and strengthen political networks that benefit women politicians. We find evidence that parties are strategic and field more women candidates in GPs which are reserved for women Pradhans.¹⁰ If parties recognize the need to find suitable female Pradhan candidates in a third of all GPs in every election, then they may encourage elected female Pradhans to identify and mentor women candidates. This process may take some time, which would explain why significant results emerge only after two electoral cycles. Third, women leaders take different policy decisions (see Chattopadhyay and Duflo [2004] and our results in Table V); voters may discover a preference for these types of decisions over time.

In the rest of this paper, we explore one possible channel in detail (more than one channel may, of course, be at play); namely, that exposure to a female Pradhan may change voter attitudes towards female leaders. We start by using survey data to document variation in voters' opinion of their Pradhans across different reservation categories. In Section IV.D. we use experimental data (for the same villagers) and examine whether these differences in voter opinion are paralleled by differences in gender bias in hypothetical leader evaluations.

IV.A. Data

Between June 2006 and November 2007 we surveyed a random sample of 495 villages spread across the 165 GPs in Birbhum district. In each village, we used a "Participatory Resource Appraisal (PRA)" exercise to obtain data on available public goods and the incidence of new investments and repair since the previous election. A facilities audit also provided information on the quality of public goods. We also administered household surveys to a random sample of fifteen households per village, and individual modules to a prime-aged male and female in each of these households. Our final sample contains 6,642 male and 6,568 female respondents.

The survey asked respondents to evaluate their Pradhan along several dimensions. It also asked respondents detailed questions on public good provision and their satisfaction with level of provision. Finally, the survey elicited experimental data on villager evaluation of hypothetical leaders. On many issues of interest we ask respondents multiple questions. Within a family of outcomes, we

¹⁰In regression available from the authors, we found that in GPs where the position of Pradhan was reserved for a woman, more women run and were elected, even if the GP had never been reserved before. These results are significant, although the sample of GPs that were never reserved becomes quite small.

expect the coefficients on the variables of interest to go in the same direction. To avoid drawing inferences based on selected outcomes, we report effects which average across outcomes within a family (Kling, Liebman, and Katz 2007). Specifically, for each outcome we construct a normalized transformation by subtracting the mean for never-reserved GPs and dividing by the standard deviation in the never-reserved sample. We obtain an average effect for the family of outcomes by estimating the effect for the average across these normalized outcomes.¹¹

IV.B. Voter’s Evaluation of their Pradhan

Each respondent was asked four questions on the effectiveness of their actual Pradhan, including “How would you rank the effectiveness of the current Pradhan?” The responses are based on a scale of one to ten. In Table IV we report regressions where the dependent variable is the normalized average across these four questions, where each question is first normalized by the mean and standard deviation of responses in never-reserved GPs and then averaged. Our regressions are of the form specified in Equation (2), and include demographic controls (age, education, caste, religion, household size, and proxies for household wealth); investigator gender and survey year indicators; and all the village controls listed in Table I. Excluding control variables produce very similar results.

Leader evaluations by villagers in GPs reserved for the first time in 2003 were significantly worse than in never-reserved GPs. The first row of column (1) shows that male respondents evaluate their Pradhans 0.2 standard deviations below Pradhans in unreserved GPs (Pradhans in never reserved GPs received, on average, a ranking of 5.1 on a scale of ten). Column (5) shows a negative, but smaller and insignificant, effect for female villagers.

In contrast, the second row in Table IV shows that both male and female villagers’ evaluation of female Pradhans in twice-reserved GPs is statistically indistinguishable from that of Pradhans in never-reserved GPs. For men, we can reject equality between the coefficient of the “Only Reserved 2003” and “Reserved in 1998 and 2003” at the one percent level. For women, the coefficients do

¹¹As in Kling, Liebman, and Katz 2007, we obtain similar results if we obtain the average effect for the family of outcomes by using a seemingly unrelated regression (SUR) system (where we account for correlation across outcomes) and then averaging across outcomes.

not differ; this, in part, reflects the fact that we cannot reject the hypothesis that women rate their female leaders at par with male leaders. Finally, the third row shows that the current (mostly male) leaders in GPs that were previously reserved in 1998 and leaders from never-reserved GPs are similarly evaluated.¹²

The timing of these results is consistent with the electoral results. Villagers have a very negative opinion of their female leader during the first electoral cycle after reservation was introduced. If this causes parties and candidates to anticipate negative electoral outcomes for women, then few women would contest a seat or be designated Pradhan immediately after. However, with improvements in voter opinion of female candidates during the second electoral cycle with reservation more female candidates enter as candidates in the third electoral cycle.

Why are female Pradhan ratings lower in GPs reserved for the first time than in GPs reserved for the second time? In Section II.B. we saw that political reservation led to the election of women who differed from male Pradhans on multiple dimensions. It may be that the aggregate difference masked significant variation across first and second generation female Pradhans. Specifically, it could be that, relative to Pradhans in GPs only reserved in 2003, female Pradhans in twice-reserved GPs are more similar to male Pradhans and this underlies the change in voter evaluations. Motivated by this observation we examine two plausible explanations: Relative to first generation female Pradhans, second generation female Pradhans either have different characteristics or act differently.

IV.C. Pradhan Selection

Women elected as Pradhan from GPs reserved for the first and second time differ on very few observable characteristics. Column (5) of Table II shows that women leaders in first and second time reserved GPs differ (significant at the 10% level or less) on only two of the eleven measures (one demographic and one political). However, these differences are potentially important: jointly, the observable characteristics of women elected in the GPs first reserved in 2003 significantly differ from those elected in GPs reserved in both 1998 and 2003 (p-value .01).

On demographics, the only significant difference relates to marital status. Relative to first time

¹²A question-by-question analysis yields identical results, and these are available in Table A.1 of the online Appendix.

female Pradhans, those elected in GPs reserved for the second time are more likely to be married (as are male Pradhans), though they are still no more likely to have a spouse previously in the council. In terms of political experience and characteristics, the only difference is that women in GPs reserved for the second time are more likely to be from the ruling party (Communist Party of India (Marxist)) than either women in GPs reserved for the first time or Pradhans in unreserved GPs. We do not have a good explanation for this result.

In order to assess whether these differences can explain the perception results, columns (2) and (6) of Table IV report the evaluation regression where we control for all eleven Pradhan characteristics. For male respondents the coefficient on Only Reserved 2003 declines slightly, from -0.2 to -0.14 but remains significant as does the difference between first and second time reserved. In the case of female respondents, the coefficient remains small and insignificant. These results demonstrate that observable differences between male and female Pradhans do not drive the evaluation gap. However, we cannot rule out the possibility that the evaluation gap reflects unobserved differences in the selection of women across GPs reserved for the first and second time. For example, political parties may learn over time how to select female candidates who appeal to male voters, or women could become more willing to run for offices thereby increasing the average quality in this group. The differences could even reflect differences in perceptions of why women were elected Pradhan. For example, villagers may not understand the double reservation principle, and believe that, the second time around, the Pradhan was chosen because she was the best candidate, not because the Pradhan had to be a woman.

IV.D. Pradhan's Actions

One way to shed more light on Pradhans' effectiveness is to examine their performance along observable dimensions. In Table V we present evidence that, *prima facie*, male Pradhans do not outperform female Pradhans.

We start by using data from the PRA survey to construct the average quantity of public goods in the village, where we average across investments in the following sectors: water and sanitation, irrigation, roads, transport, schools and other educational facilities, and health. Across all public

goods, the average number of repairs or new construction since 2003 is 0.19 standard deviations higher in villages where the Pradhan position is reserved for the first time (column (1)).¹³ The effect in twice-reserved GPs is smaller and insignificant (0.04, with standard error of 0.06) although not statistically distinguishable from the effect in GPs only reserved in 2003. If anything, although Pradhans in twice-reserved GPs are less active than those in GPs reserved only once.

This result could occur because female Pradhans in first-time reserved GPs invest in lower quality public goods. The facilities audit allows us to measure the average quality of public goods across the water, roads, transport, school and other educational facilities, health, and fair price shop sectors. Columns (2)-(4) show similar quality of public good provision and also, on average, very similar levels of self-reported villager satisfaction with public goods across GPs in different reservation categories.¹⁴

Since public goods are mainly financed by State Government funds, the contrast between villagers' negative evaluation of female Pradhans and the fact that female Pradhans invest more with no discernible reduction in quality is unlikely to reflect resentment with a "big government" approach. Villagers may, however, have to pay for these goods through means such as voluntary contributions and bribes. In column (5) we see that on average, individuals in currently reserved GPs are less likely to have paid a bribe for obtaining a BPL card or drinking water connection. This is true for both GPs reserved for the first and second time.

The evidence suggests that women leaders, especially first time leaders, provide more public goods of equal quality at a lower effective price. However, the bundle of public goods chosen by female leaders may be less valued by male villagers. In the same setting, Chattopadhyay and Duflo (2004) showed that women leaders invest more in goods preferred by women.¹⁵ In column (6) we use their measure of female preferences (constructed using the difference in male and female villagers' formal complaints) and find that female leaders invest more in women-preferred goods. This provides a potential explanation for why male villagers rate first time female leaders negatively.

¹³The Online Appendix describes these data and Online Appendix Table A.2 shows very similar results for goodwise regressions. Our findings are also broadly consistent with Munshi and Rosenzweig (2008) who using nationally representative data find that women leaders in non-caste dominant GPs demonstrate higher competence as measured by public good provision.

¹⁴The only exception when we examine good-specific satisfaction is that men are unsatisfied with female leaders allocation of the limited supply of BPL cards. In contrast, Duflo and Topalova (2004), using nation-wide Indian data, found that villagers were, on average, less satisfied with public good provision when women were in charge.

¹⁵Our data shows that they continue to invest in different types of goods, see online Appendix Table A.2.

However, since the policy choices of female Pradhans in first and second-reserved GPs are equally pro-woman, this cannot explain the differential evaluation of Pradhans across first and twice-reserved GPs.

To investigate whether the difference in actions can explain voter opinions, columns (3) and (7) in Table IV report evaluation regressions where we control for the quantity and quality of public goods, using the indices from columns (1) and (2) of Table V. Columns (4) and (8) show that our evaluation regression results are robust to simultaneously controlling for Pradhan characteristics and actions.

While *prima facie*, the results in Tables IV and V suggest that Pradhan characteristics and actions do not account for the evolution of voters' opinion of female Pradhans, it is possible that we have failed to capture relevant aspects of public good provision such as the targeting of transfers or jobs. Alternatively, the evaluation gap may reflect the fact that first-time women leaders are simply worse at getting credit for their work (Deaux and Emswiler 1974). Another possibility is that female Pradhans are not just less likely to take bribes but are also less willing (or able) to bribe influential villagers (for instance, by allocating them BPL cards or otherwise favoring them), and these villagers respond by adversely influencing public opinion.¹⁶ In other words, women may make good leaders but bad politicians. In villages reserved for the second time, women may have acquired better skills as politicians.

For all these reasons, actual leaders' approval ratings are insufficient to establish that male villagers are initially biased against female leaders, and that exposure to a female leader reduces this bias. We, therefore, turn to experimental measures which elicit villagers' opinion of women as leaders.

V. Does Reservation Reduce Bias against Female Leaders?

Bias against female leaders may reflect taste or statistical discrimination. The two could also reinforce each other if taste discrimination prevents the initial election of female candidates and this, in turn, prevents voters from gathering information on their leadership. Lack of information on women's competence may prevent risk averse voters from subsequently electing women, creating a

¹⁶We thank one referee for this suggestion.

vicious circle where women are never elected.¹⁷ In such an environment, gender quotas, by reducing statistical discrimination, may enhance women’s long-run electoral prospects even if taste-based bias persists. To investigate this possibility we use survey data that captures aspects of both taste and statistical discrimination.

V.A. Data

Hypothetical Leader Effectiveness Our individual survey modules included speech and vignette experiments which fall under the “Goldberg paradigm” (Goldberg 1968). Such experiments, in which the gender of the protagonist is randomly varied, have been widely used in the U.S. to isolate bias in the perceived effectiveness of women as leaders (Huddy and Terkildsen 1993; Matland 1994; Eagly and Karau 2002).

Each respondent heard a short tape-recorded leader speech. In the speech, which was adapted from an actual village meeting, the leader responds to a villager complaint about a broken tubewell by requesting villagers to contribute money and effort for local public goods (the online Appendix provides the transcript). Respondents were randomly assigned one of six speech recordings (three per gender) and told that this was recorded during a village meeting in another district. After hearing the speech, the respondent evaluated the leader’s perceived performance and overall effectiveness along seven themes, including whether the leader addressed villagers’ concerns correctly, and whether he/she would be good at collecting resources from villagers.

Each respondent also heard a randomly selected vignette in which a situation of resource scarcity was described and the leader chose to invest in either a drinking water or an irrigation project. Vignettes varied along two dimensions: the leader’s gender and choice. Variation in the leader’s choice was introduced because, relative to men, women invest more in drinking water and are more likely to cite drinking water as an issue of concern. In contrast, men state a greater concern with irrigation (Chattopadhyay and Duflo 2004). Randomizing the choice insures, on average, a “gender neutral” decision.¹⁸

¹⁷The working paper version of this paper, <http://www.nber.org/papers/w14198>, provides a model that formalizes this argument.

¹⁸In practice, both genders considered investment in water was the best decision, so that Pradhan’s approval was much higher among male and female respondents who received the vignette where drinking water was chosen.

While leader gender was randomly varied across respondents, a respondent was exposed to the same leader gender in the speech and vignette. In our analysis, we combine villagers' responses across the speech and vignette and ask whether, holding actions constant, villagers, on average, rank female leaders below male leaders. The disaggregated results are available in the on-line appendix to this paper.

Measure of Gender-Occupation Stereotype We used an activity-based Implicit Association Test to examine whether exposure to a female leader alters the extent of gender stereotyping of occupations by villagers. We also conducted two taste-based IATs to measure general feelings towards women (described below). To the best of our knowledge, we are the first to conduct IATs in a developing country. Below, we first describe the general features of an IAT and then our activity-based IAT.

An IAT is a computerized test that aims to measure attitudes of which respondents may not be explicitly cognizant. It uses a double-categorization task to measure the strength of respondent association between two concepts. To account for limited computer familiarity and high illiteracy, our IATs used audio or pictorial prompts. While we are unaware of other studies which conduct IATs with an illiterate population, the pioneers of the IAT argue that words, pictures or sounds can be used to represent concepts, making it possible to administer IATs to “the blind, young children, and others who are unable to read” (Carney et al. 2007). English versions of our three IATs, along with a full list of prompts, are available electronically.¹⁹

During an IAT test block the respondent observes a computer screen and categorizes a sequence of stimulus-based categorization tasks. Appendix Figure 1 shows a screenshot from the leadership-domestic IAT. Here the respondent sees two pictures (a setting and a person) on either side of the computer screen. In the “stereotypical” block, the male picture and leadership setting are grouped together on one side (say, the right side) and female picture and domestic setting on the other (left) side. The “non-stereotypical” block reverses this association between gender and domestic and leadership words (so now male pictures and domestic setting are on the right, and female pictures and leadership setting are on the left).

While viewing the screen the respondent hears or sees a sequence of stimuli, such as a spoken

¹⁹The IATs are posted at <http://www.povertyactionlab.org/projects/project.php?pid=102>. Our IATs resemble those used for children (Baron and Banaji 2006).

word or picture. She then uses the computer button to assign each stimulus to the correct side of the screen. For each stimulus, there is an unambiguously correct response which is made clear during two prior practice blocks.²⁰ The time a respondent takes to accomplish each categorization task is recorded in milliseconds. The presumption is that respondents with a stronger association between two concepts find the sorting task easier and complete it faster. The relative strength of association can, therefore, be detected by comparing response time across the stereotypical and non-stereotypical block. The normalized difference in mean response times between the “non-stereotypical” and “stereotypical” test blocks is the D-measure of IAT bias, with higher values indicating stronger implicit stereotype (Greenwald, Banaji, and Nosek 2003).

The complete IAT procedure for one subject consists of two practice rounds, with single cues on each side, followed by the two test blocks, with the stereotypical and non-stereotypical cues on each side. The order in which test blocks are administered is randomized. Following standard practice the IAT was automatically stopped if participants were too slow or made too many mistakes (Nosek, Banaji, and Greenwald 2002). The rationale is that the IAT relies on rapid responses and, therefore, on the respondent’s understanding of the task at hand. Too many mistakes, or too slow a response time, suggests that the test captured the respondent’s difficulty with the test rather than her attitudes. Our cutoff, which was based on extensive piloting and was more generous than those typically used, was to stop the test for respondents with an average response time of over 6 seconds or less than 60% average correct responses in the first block. To ensure similar treatment across blocks, at the analysis stage we dropped all respondents with an error rate of above 65% in either test block.²¹

We administered the IATs to adults aged 15 to 45 in a random subsample of 5 households per village. Each respondent was randomly assigned one of the three IATs. Overall, we had 4,378 respondents from 1,968 households.²² The screening criteria reduced the number of usable observations to 2,816 (across the three IATs). Both criteria were more likely to exclude older and female respondents.²³ Finally, we dropped stimulus responses, but not the entire test block, for

²⁰A mistake in this example would mean that the respondent classified a female name on the male side, or a leadership word on the domestic side.

²¹The results are not sensitive to this threshold.

²²We exclude IAT data from the one pilot village in each of seventy-seven GPs.

²³Relative to included participants, the average D-measure of respondents excluded at the analysis stage was 0.12 standard deviation higher for men and 0.14 standard deviation lower for women. Our results are robust to not

which the respondent took more than 10 seconds to categorize a prompt.²⁴

Over the last ten years, IATs have been widely used to measure implicit bias in various subfields in psychology (Nosek, Greenwald, and Banaji 2007), neuroscience, market research, and recently, economics (Bertrand, Chugh, and Mullainathan 2005; Rooth 2007). However, their use is not uncontroversial. A first class of issues concerns IATs' internal validity.²⁵ IAT current procedures have been developed to avoid such bias, and our IATs respect those procedures. A second category of issues concerns "construct validity." These include whether IATs measure the strength of associations between concepts or some other reasons such as salience (Rothermund and Wentura 2004), or whether IATs capture cultural, rather than individual, association of categories (Karpinski and Hilton 2001; Olson and Fazio 2004; Arkes and Tetlock 2004). Finally, IATs reflect instant decisions, and may not matter for deliberative decisions such as voting. Compared to the internal validity concerns, these criticisms are more fundamental and have led to a lively debate on the legitimate use of IAT. Nevertheless, the IAT has been shown to be good predictor of individual behavior in several applications, including voting behavior (Nosek, Banaji, and Greenwald 2002; Arcuri et al 2008).²⁶

Finally, an important innovation of our analysis is to examine the malleability of IAT bias in response to exposure to female leaders. We are only aware of one other field study which investigates the malleability of implicit beliefs. This study found that students in American colleges with a higher proportion of female professors were more likely to associate women with professional activities (Dasgupta and Asgari 2004). However, endogenous selection into college makes a causal interpretation of these results difficult.

Turning to our choice of IAT, we use an activity-based IAT to assess whether villagers exposed to reservation are less likely to associate women with domestic activities and men with leadership activities. The IAT examines the association between male and female names and domestic (e.g. taking rest) and leadership (e.g. public speaking) activities. To avoid biasing the results towards

conducting the additional screening during analysis; these results are available from the authors.

²⁴We assumed the respondent was temporarily inattentive; standard practice in psychology literature is to top-code answers longer than 3 seconds at 3 seconds.

²⁵IATs could, for example, be affected by stimuli familiarity, the specific choice of stimuli, the order in which they appear, previous experience with the IAT, or the association of categories with a particular side.

²⁶A list is maintained on Anthony Greenwald's web page http://faculty.washington.edu/agg/iat_validity.htm.

associating women with domestic activity, we chose gender neutral domestic activities which were equally likely to be performed by men and women. This IAT does not capture a value judgement. Rather, it is informative of whether the respondent considers both men and women as potential leaders.

Measures of Voter Taste We use taste-based IATs and survey responses to construct measures of voters' explicit and implicit taste for male and female leaders in general.

Taste-based IATs have been widely used to measure group prejudice (Banaji 2001).²⁷ Our first taste IAT assesses the associational strength between male and female names and positive (e.g. nice) and negative (e.g. nasty) 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 good attributes on one side of the screen, and female names or leader pictures and bad attributes on the other side.

Households who received IATs were also explicitly 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 villager.²⁸ This question is adapted from the “Feeling Thermometer,” which has been widely used in the political science literature to produce rank-orderings of parties and candidates (Kelley and Mirer 1974) and to measure partisan affiliation (Weisberg 1980; Alvarez 1990). IATs and explicit measures, such as the above, are complements insofar as explicit measures such as the feeling thermometer capture socially acceptable responses, rather than a respondent's true opinion (Greenwald, McGhee, and Schwartz 1998).

V.B. Results

Hypothetical Leader Effectiveness

We start by examining villagers' evaluation of the hypothetical Pradhan actions as described in the tape-recorded speech and vignette. Unlike actual Pradhans, here, by construction, there

²⁷See the Project Implicit website, <http://implicit.harvard.edu/implicit/demo/background/bibliotopic.html>, for an extensive list.

²⁸We did not refer to a particular female leader, either by name or function (Pradhan); the question refers to any female leader.

are no observable or unobservable performance differences between the male and female leaders. If respondents choose to *infer* differences in Pradhan performance based on Pradhan gender, then this is evidence of statistical discrimination. Our objective is to evaluate whether such discrimination, if any, was affected by the reservation policy.

We examine whether villagers judge male and female Pradhans differently, and whether this difference varies with the reservation status of the GP. Let F_{ig} indicate whether respondent i received a “female” stimulus (i.e. heard the speech in a female voice, or was described the vignette with a female leader). We estimate:

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

and

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

where R_g indicates a GP which has ever been reserved for a female Pradhan, in 1998, 2003 or both, and the indicator variables R_k control for the main effect of different reservation categories, as defined in previous sections.

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.

Columns (1) and (2) of Table VI report the coefficients from regressions in which the dependent variable is the average of the normalized outcomes across the speech and vignette questions. In Panel A we observe a significant bias among men in never-reserved villages. The coefficient on female Pradhan suggests that they rate the effectiveness of a hypothetical female Pradhan 0.054 of a standard deviation below that of a male Pradhan. While the bias is lower and insignificant among women, we cannot reject the hypothesis of a similar bias across genders.

In contrast, the coefficient on the interaction between female Pradhan and ever reserved is a

strongly significant 0.091. Combining the two Panel A coefficients suggests that reservation erases, indeed reverses, this bias (though the resulting pro-female bias is only weakly significant). Panel B shows a similar impact of reservation across reservation categories: all coefficients are positive, with similarly-sized and statistically indistinguishable point estimates. The results are striking and suggest that exposure to at least one female leader due to reservation can erase statistical discrimination by male villagers. Moreover, this effect persists even after the woman has left office.

In comparison, our estimates, though noisy, suggest very limited updating among female villagers. One possible explanation is that female villagers are less involved in local politics: women are significantly less likely to know the Pradhan's name or to have ever been in direct contact with him/her (online Appendix Table A.4). If women are largely unaware of local politics, then it is unsurprising that reservation does not affect their evaluations. Even among men, the impact of reservation on statistical discrimination is concentrated among those who know the Pradhan's name (online Appendix Table A.5). Another possibility is that a female leader, by virtue of being a counterstereotypic figure, makes women's traditional roles appear lower status and therefore women react negatively to this. Finally, it could be that members of a group are simply less likely to update about their own group based on another group member's behavior. The results suggest the absence of confirmatory bias in this setting: since, compared to women, men started with a worst opinion of women's relative ability to lead, they would have updated less than women.

Gender-Occupation Stereotypes

Next we use an occupation-based IAT to examine whether exposure to female leaders reduces villagers' propensity to associate leadership activities with men and domestic activities with women.

In Columns (1) and (2) of Table VII, Panel A we examine the impact of having lived in a village which was ever reserved. The regression specification examines the difference between ever and never reserved GPs, where we control for respondent demographics, 1991 village control variables, block fixed effects, and investigator and survey year indicator. The Panel B regression examines whether this effect varies across reservation categories (regression specification is as in equation (2) with the additional controls listed above).

Both genders associate leadership activities more strongly with men in never-reserved GPs: they 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 significantly reduces this association among male respondents.

This effect is, however, absent among women. If anything, Panel B shows a strengthening of the stereotype associating women and domestic activities among women exposed to a female Pradhan for the first time in 2003.

The vignettes and speech and the IAT suggest that statistical discrimination against women and the stereotyping of men with leadership activities disappear relatively fast, i.e. within the first reservation cycle: two years after being first required to elect a woman, when presented with the same information on a leader's action, villagers judge a male and a female leader as equally able. In contrast, villagers who have never been exposed to a female leader judge a hypothetical female leader as less effective. This can explain why the evaluation of actual female leaders improves between the first and second reservation cycle. The first woman to be elected suffers from the same statistical discrimination as the female Pradhan in the speech and vignette in never-reserved GPs: presented with similar information on her effectiveness (e.g., her first decisions) they rate her as less effective than they would rate a man. Progressively (maybe as they learn that the result of the same action taken by a man or a woman is the same), they begin to update their opinion of a woman leader (including the current one). However, despite favorable updating, their overall opinion of their current Pradhan continues to be influenced by their low initial prior of the leader at the time of her election. Hence, it remains lower than that of male Pradhans. Meanwhile, similar to the women in the speech and vignettes, the initial negative prior is absent for the second cohort of female Pradhans and they are judged using standards similar to men.

Preference for Male Leaders

We now investigate whether exposure alters villagers' implicit and explicit taste for female leaders. As a reminder, we measure implicit preference for male leaders by the D-measure in the taste IAT. To benchmark respondents' ranking of leaders, we also report IAT results for respondent preferences for male versus female villagers. We measure an explicit preference for male leaders by the difference between a villager's general rating of male and female Pradhans (on a 1-10 scale). Our regression specifications are as in the previous section.

The results are in Table VII. For both taste IATs, the mean D-measures in the never-reserved

samples suggest a strong same gender preference among male and female villagers (columns (3)-(6)). Men are significantly more likely to associate good with male names or pictures of male leaders, while women associate female names (and images of female leaders) with positive attributes. Of particular interest is the coefficient in the first row of column (5) which suggests that, unlike the occupation IAT, men exposed to reservation do not update favorably towards women in the leader taste IAT. The point estimate is small and insignificant (0.014, with a standard error of 0.037). We can reject at 95% confidence interval a decrease of -0.076 standard deviation on the D-measure. Note that the D-measure is 0.093 in the unreserved sample. Thus contrary to what we found for the effectiveness perception and gender-occupation measures, we can reject the hypothesis that exposure removed male citizens' implicit preferences for male leaders.

Next we examine villagers' explicit feelings. In columns (7) and (8) we see that villagers are not shy about admitting explicit preferences for males. Male villagers in never-reserved villages rate male leaders 1.44 points higher than female leaders. The difference, while smaller (0.56), remains significant among female villagers. The distaste for female leaders is not ameliorated by exposure. On the contrary, the coefficient of "Ever Reserved" (Panel A, column (7)) is positive for both genders. For male villagers, the effect is also significantly positive, suggesting that their relative explicit preference for male leaders (compared to female leaders) is strengthened in GPs which have experienced a gender quota.

These results suggest that affirmative action does not, in the short to medium-run, alter voter taste for female leaders. As in most of the literature on IATs, we find some differences between implicit and explicit measures (Hoffman et al 2005). However, in contrast to rich countries where explicit bias tends to be muted even when respondents exhibit strong implicit bias (Bertrand, Chugh, and Mullainathan 2005), we find a strong explicit bias against women leaders (which is worsened by initial reservation). A possible explanation is deep rooted social norms that members of disadvantaged groups, including women, should not lead. The finding that men are *more* biased in once-reserved GPs, but only as measured by explicit attitudes, supports a "backlash" hypothesis. That is, men protest against a quota system which forces them to elect women by stating a dislike for female leaders, even though their implicit preferences are unaffected. If the backlash reflects such a protest, and is driven by a general distaste for seeing members of disadvantaged groups in

leadership positions, then this backlash should be stronger in GPs which are simultaneously reserved for multiple categories - women and either Scheduled Castes or Scheduled Tribes. Regressions which control for both types of reservation show that the backlash effect can be entirely attributed to GPs which are simultaneously reserved for SC and Women: that is, voters stated dislike for women is particularly strong in GPs reserved for SC (results not reported). Finally, we underestimate the backlash to the extent that knowledge of the reservation policy induces people who have not yet had female Pradhan reservation to become more negative towards women.²⁹

Social norms that militate against female leadership, combined with same-gender preferences, could also explain the absence of consistent results across explicit and implicit attitudes for female villagers. Specifically, women’s explicit attitudes may reflect that they acknowledge that ranking female leaders below male leaders is prescribed behavior, even though they like female leaders (and this is captured by their implicit beliefs).³⁰

VI. Conclusion

Indian villagers prefer male leaders and have negatively biased priors on the effectiveness of female leaders. On a scale of one to ten, male villagers rate their feeling towards female leaders more than one point below that towards male leaders; moreover, in never-reserved villages hypothetical leaders with identical performance are evaluated as less effective if they are female. While reservation does not make male villagers more sympathetic to the idea of female leaders, our IAT and speech and vignette results suggest that it makes them more likely to associate women with leadership and improves their evaluation of female leader effectiveness.

The timing of effects are consistent across the actual Pradhan evaluation and the experimental results. The speech and vignette results suggest that voters immediately update their opinion on the effectiveness of female leaders in general (i.e. after being exposed to reservation once). The

²⁹In Appendix Table A.6 we report multiple robustness checks. These include showing that evaluation of new Pradhans is not worse (suggesting our results are not driven by women being new Pradhans. We also show that the improvement in perceptions of female Pradhans in twice reserved does appear to reflect getting used to reservation. First time women Pradhans in GPs where the position was previously reserved for SCs receive negative evaluations.

³⁰We should note a final caveat for our IAT findings. It is possible that the very similar results across the two taste IATs reflect the fact that the second IAT may have failed to sufficiently capture a “leadership” dimension. Specifically, it may be that in both cases respondents only perceived the gender of the character as salient (and the respondents did not associate pictures of female leaders with leadership).

second generation of female Pradhans benefits from this decline in bias: this may explain why they are liked as much as male Pradhans. In contrast, the evaluation of the first Pradhan elected during reservation continues to be colored by voters' initial prior against women.

The timing of the impact on electoral outcomes follows naturally from the timing of the impact on citizens' explicit opinions: if, after one round of reservation, parties observe that the general distaste for female leaders has not improved (as we see in Table VII), and that voters are negatively disposed towards their current leaders (Table IV), then they have no incentive to field women candidates. Analogously, individual women may not consider running in an environment where female leaders are not well regarded. However, once female Pradhans' approval ratings improve, fielding women becomes a good idea. Thus, if reservation improves electoral outcomes for women, the effect should be stronger after two cycles of reservations than after one cycle. This is what we observe in the electoral data. In the 2008 Panchayat elections, the third since the reservation policy was implemented, more women were elected ward councilors and Pradhans in GPs that had been reserved for women in the last two elections.³¹ The results provide striking evidence that while ten years of exposure to women leaders may not have changed voters' stated preference for male leaders, by giving voters a chance to learn about the effectiveness of women leaders, they have effectively improved women's access to political office.

These results, of course, leave open the possibility that other channels of influence are also at work. Another important caveat is that our study occurred in a state which has seen a single party in power for the last thirty years, where village councils have worked well, and where gender discrimination is considered less extreme than in some other Indian states. In other settings, taste discrimination against women leaders may play a larger role in determining electoral outcomes. An encouraging sign for the external validity of our findings is the study of Bhavnani (forthcoming) who finds similar results in Mumbai municipal wards. Conversely, the setting for this study is India, a poor country with relatively high levels of gender discrimination. It is possible that impact of quotas in rich countries, which tend to have less gender discrimination, would look different.

That said, these results provide some of the first evidence on the potential role for public policy

³¹Even though we do not see impact of reservations on the attitude of the average woman, we have argued that the most likely explanation is their limited exposure to politics. Women who are capable of standing for election must, however, be the most politically aware.

in mitigating voter discrimination. While the political underrepresentation of women is widely documented, there is almost no credible evidence on whether public policy can influence voters' belief systems and prejudice. Much of the evidence from outside the political sphere suggests that policies that limit awareness of a candidate's gender may be an optimal short-run response to discrimination (Goldin and Rouse 2000). Our results suggest that, in political settings where such strategies are infeasible, political affirmative action can play an important medium-run role. While the first generation of women leaders may encounter significant prejudice, their experience can pave the way for others to go further.

NORTHWESTERN UNIVERSITY

INDIAN INSTITUTE OF MANAGEMENT, CALCUTTA

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, NATIONAL BUREAU OF ECONOMIC
RESEARCH, AND CENTER FOR ECONOMIC POLICY RESEARCH

HARVARD UNIVERSITY AND CENTER FOR ECONOMIC POLICY RESEARCH

INTERNATIONAL MONETARY FUND

References

- Aigner, Dennis and Glenn Cain, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 30 (1977), 175–187.
- Alvarez, Michael, "The Puzzle of Party Identification - Dimensionality of an Important Concept," *American Politics Research*, 18 (1990), 476–91.
- Arcuri, Luciano, Luigi Castelli, Silvia Galdi, Cristina Zogmaister, and Alessandro Amadori, "Predicting the Vote: Implicit Attitudes as Predictors of the Future Behavior of the Decided and Undecided Voters," *Political Psychology*, 29 (2008), 369–387.
- Arkes, Hal Richard and Philip Tetlock, "Attributions of Implicit Prejudice, or Would Jesse Jackson 'Fail' the Implicit Association Test?" *Psychological Inquiry*, 15 (2004), 257–278.
- Banaji, Mahzarin, "Implicit Attitudes Can Be Measured," in Henry Roediger and James Nairne, eds., "The Nature of Remembering: Essays in Honor of Robert G. Crowder," (Washington, DC: American Psychological Association 2001), 117–150.
- Baron, Anthony and Mahzarin Banaji, "The Development of Implicit Attitudes: Evidence of Race Evaluations from Ages 6 and 10 and Adulthood," *Psychological Science*, 17 (2006), 53–58.
- Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan, "Implicit Discrimination," *American Economic Review*, 95 (2005), 94–98.
- Besley, Timothy, Rohini Pande, and Vijayendra Rao, "Just Rewards? Local Politics and Public Resource Allocation in South India," (2007), mimeo, Harvard.
- Bhavnani, Rikhil, "Do Electoral Quotas Work After They Are Withdrawn? Evidence from a Natural Experiment in India," *American Political Science Review*, forthcoming.
- Boisjoly, Johanne, Greg Duncan, Michael Kremer, Dan Levy, and Jacque Eccles, "Empathy or Antipathy? The Consequences of Racially and Socially Diverse Peers on Attitudes and Behaviors," *American Economic Review*, 96 (2006), 1890–1905.
- Carney, D.R., B.A. Nosek, A.G. Greenwald, and M.R. Banaji, "The Implicit Association Test (IAT)," in R. Baumeister and K. Vohs, eds., "Encyclopedia of Social Psychology,"

- (Thousand Oaks: CA: Sage Publications 2007), 463–464.
- Chattopadhyay, Raghavendra and Esther Duflo, “Women as Policy Makers: Evidence from a Randomized Policy Experiment in India,” *Econometrica*, 72 (2004), 1409–1443.
- Coate, Stephen and Glenn Loury, “Will Affirmative-Action Policies Eliminate Negative Stereotypes?” *American Economic Review*, 83 (1993), 1220–1240.
- Dahlerup, Drude, *Women, Quotas and Politics* (New York: Routledge 2006).
- Dasgupta, Nilanjana and Shaki Asgari, “Seeing is Believing: Exposure to Counterstereotypic Women Leaders and its Effect on the Malleability of Automatic Gender Stereotyping,” *Journal of Experimental Social Psychology*, 40 (2004), 642–658.
- Deaux, K. and T. Emswiller, “Explanations of Successful Performance on Sex-linked Tasks: What is Skill for the Male is Luck for the Female,” *Journal of Personality and Social Psychology*, 29 (1974), 80–85.
- Duflo, Esther and Petia Topalova, “Unappreciated Service: Performance, Perceptions, and Women Leaders in India,” (2004), mimeo.
- Eagly, A.H. and S.J. Karau, “Role Congruity Theory of Prejudice toward Female Leaders,” *Psychological Review*, 109 (2002), 573–598.
- Goldberg, Philip, “Are Women Prejudiced Against Women?” *Transaction*, 5 (1968), 28–30.
- Goldin, Claudia, “A Pollution Theory of Discrimination: Male and Female Differences in Occupations and Earnings,” (2002), NBER Working Paper no. 8985.
- Goldin, Claudia and Cecilia Rouse, “Orchestrating Impartiality: The Impact of ‘Blind’ Auditions on Female Musicians,” *American Economic Review*, 90 (2000), 715–741.
- Government of West Bengal, *The West Bengal Panchayat (Constitution) Rules 1975*, Department of Panchayats & Rural Development (1998).
- Greenwald, A.G., D. McGhee, and J.L.K. Schwartz, “Measuring Individual Differences in Implicit Cognition: The Implicit Association Test,” *Journal of Personality and Social Psychology*, 74 (1998), 1464–1480.

- Greenwald, Anthony, Mahzarin Banaji, and Brian Nosek, "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm," *Journal of Personality and Social Psychology*, 85 (2003), 197–216.
- Hofmann, Wilhelm, Bertram Gawronski, Tobias Gschwendner, Huy Le, and Manfred Schmitt, "A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures," *Personality and Social Psychology Bulletin*, 31 (2005), 1369–1385.
- Huddy, Leonie and Nayda Terkildsen, "Gender Stereotypes and the Perception of Male and Female Candidates," *American Journal of Political Science*, 37 (1993), 119–147.
- Imbens, Guido and Paul Rosenbaum, "Robust, Accurate Confidence Intervals with a Weak Instrument: Quarter of Birth and Education." *Journal of the Royal Statistical Society*, 168 (2005), 109–126.
- Jones, Mark, "Quota Legislation and the Election of Women: Learning from the Costa Rican Experience," *Journal of Politics*, 66 (2004), 1203–1223.
- Karpinski, Andrew and James Hilton, "Attitudes and the Implicit Association Test," *Journal of Personality and Social Psychology*, 81 (2001), 774–788.
- Kelley, Stanley and Thad Mirer, "The Simple Act of Voting," *American Political Science Review*, 68 (1974), 572–91.
- Kling, Jeffrey, Jeffrey Liebman, and Lawrence Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75 (2007), 83–119.
- Krook, Mona, *Politicizing Representation: Campaigns for Candidate Gender Quotas Worldwide* (New York: Ph.D. diss Columbia University 2005).
- Matland, Richard E., "Putting Scandinavian Equality to the Test: An Experimental Evaluation of Gender Stereotyping of Political Candidates in a Sample of Norwegian Voters." *British Journal of Political Science*, 24 (1994), 273–292.
- Munshi, Kaivan and Mark Rosenzweig, "The Efficacy of Parochial Politics: Caste, Commitment, and Competence in Indian Local Governments," (2008), mimeo.
- Nosek, Brian, Mahzarin Banaji, and Anthony Greenwald, "Harvesting Implicit Group Atti-

- tudes and Beliefs from a Demonstration Web Site,” *Group Dynamics: Theory, Research and Practice*, 6 (2002), 101–115.
- Nosek, Brian, Anthony Greenwald, and Mahzarin Banaji, “The Implicit Association Test at Age 7: A Methodological and Conceptual Review,” in J. A. Bargh, ed., “Automatic Processes in Social Thinking and Behavior,” (New York: Psychology Press 2007).
- Olson, Michael and Russel Fazio, “Reducing the Influence of Extrapersonal Associations on the Implicit Association Test: Personalizing the IAT,” *Journal of Personality and Social Psychology*, 86 (2004), 653–667.
- Powley, Elizabeth, “Rwanda: The Impact of Women Legislators on Policy Outcomes Affecting Children and Families,” (2006), Background Paper, State of the World’s Children.
- Rooth, Dan-Olof, “Implicit Discrimination in Hiring: Real World Evidence,” (2007), IZA Discussion Paper 2764.
- Rosenbaum, Paul, “Observational Studies and Nonrandomized Experiments,” in Subir Ghosh and C.R. Rao, eds., “Handbook of Statistics,” volume 13 (Amsterdam: Elsevier 1996), 181–197.
- Rothermund, Klaus and Dirk Wentura, “Underlying Processes in the Implicit Association Test(IAT): Dissociating Salience from Associations,” *Journal of Experimental Psychology: General*, 133 (2004), 139–165.
- Rudman, Laurie and Kimberly Fairchild, “Reactions to Counterstereotypic Behavior: The Role of Backlash in Cultural Stereotype Maintenance,” *Journal of Personality and Social Psychology*, 87 (2004), 157–176.
- Thernstrom, Stephan and Abigail Thernstrom, *America In Black and White: One Nation, Indivisible* (New York, NY: Simon and Schuster 1997).
- UNIFEM, *Progress of World’s Women 2008*, United Nations (2008).
- Weisberg, Herbert, “A Multidimensional Conceptualization of Party Identification,” *Political Behavior*, 2 (1980), 33–96.

Table I
Village-level Randomization Check

Census variable	Only reserved	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 (1318)	1,197 (1295)	1,267 (1418)	1,362 (1555)	-102.279 (178.537)	-78.373 (189.410)
Fraction SC / ST population	0.432 (0.281)	0.502 (0.256)	0.450 (0.235)	0.471 (0.265)	0.004 (0.056)	-0.001 (0.033)
Average household size	5.436 (0.612)	5.222 (0.449)	5.362 (0.571)	5.497 (2.454)	0.137 (0.093)	-0.059 (0.135)
Sex ratio under 6	1.093 (0.590)	1.034 (0.276)	1.046 (0.218)	1.036 (0.226)	0.206 (0.113)	0.015 (0.025)
Literacy	0.387 (0.119)	0.351 (0.129)	0.391 (0.121)	0.373 (0.137)	0.014 (0.032)	0.021 (0.017)
Fraction women literate	0.287 (0.125)	0.265 (0.123)	0.296 (0.113)	0.275 (0.134)	-0.003 (0.033)	0.023 (0.016)
Fraction irrigated land	0.599 (0.321)	0.485 (0.349)	0.607 (0.323)	0.500 (0.352)	-0.038 (0.054)	0.054 (0.049)
Village has a bus or train stop	0.280 (0.451)	0.433 (0.500)	0.290 (0.456)	0.259 (0.439)	-0.130 (0.097)	0.061 (0.054)
Village has permanent approach road	0.110 (0.314)	0.300 (0.462)	0.176 (0.383)	0.194 (0.397)	-0.225 (0.082)	-0.006 (0.052)
Village has tube well	0.910 (0.288)	0.933 (0.252)	0.912 (0.285)	0.977 (0.151)	-0.022 (0.020)	-0.018 (0.031)
Village has hand pump	0.124 (0.331)	0.000 (0.000)	0.100 (0.302)	0.057 (0.233)	0.000 (0.000)	0.013 (0.012)
Village has well	0.440 (0.499)	0.533 (0.503)	0.402 (0.493)	0.528 (0.500)	0.043 (0.082)	-0.066 (0.074)
Village has community tap	0.065 (0.248)	0.050 (0.220)	0.042 (0.202)	0.010 (0.098)	0.022 (0.020)	-0.018 (0.017)
Number of schools	1.190 (0.895)	1.233 (0.810)	1.265 (0.889)	1.160 (0.910)	-0.167 (0.160)	0.149 (0.111)
Number of health facilities	0.170 (0.403)	0.217 (0.490)	0.118 (0.324)	0.218 (0.809)	-0.036 (0.090)	-0.015 (0.069)
Overall Effect: F stat					0.300	0.600
Overall Effect: p value					0.586	0.438
N	105	60	108	222		

Notes:

- 1 The census variables are from the 1991 Census of India. N refers to the maximum number of observations. Infrastructure variables have between 454 and 478 total observations across GPs.
- 2 "First Reserved 2003," "Reserved 1998 and 2003," "Only Reserved 1998," and "Never Reserved" are indicator variables for GPs reserved for a female Pradhan for the first time in 2003, in both 1998 and 2003, only in 1998, and not reserved in either election, respectively.
- 3 Columns (1)-(4) report means with standard deviations in parenthesis. Columns (5)-(6) report tests of differences of means across columns (1) and (2) and columns (3) and (4), respectively. Standard errors are in parentheses. Tests are based on regressions with block fixed effects and standard errors are clustered by GP.

Table II
Pradhan Characteristics

	Only reserved 2003 (1)	Reserved 1998 and 2003 (2)	Only reserved 1998 (3)	Never reserved (4)	Diff: (1) and (2) (5)	Diff: (1)+(2) vs. (3)+(4) (6)
Age	35.424 (8.143)	33.842 (6.801)	41.571 (12.425)	41.194 (9.728)	0.912 (2.550)	-6.339 (1.518)
Years of education	8.091 (4.126)	8.895 (3.017)	11.229 (2.691)	10.278 (3.199)	-1.119 (0.927)	-2.104 (0.548)
Married	0.727 (0.452)	0.895 (0.315)	0.857 (0.355)	0.889 (0.316)	-0.202 (0.117)	-0.103 (0.065)
SC/ST	0.588 (0.500)	0.450 (0.510)	0.543 (0.505)	0.514 (0.503)	0.070 (0.158)	0.013 (0.084)
Wealth index: quartile 1	0.382 (0.493)	0.300 (0.470)	0.143 (0.355)	0.208 (0.409)	0.081 (0.143)	0.164 (0.077)
Wealth index: quartile 4	0.088 (0.288)	0.250 (0.444)	0.257 (0.443)	0.208 (0.409)	-0.132 (0.121)	-0.075 (0.065)
First time as Pradhan in 2003 term	0.912 (0.288)	0.850 (0.366)	0.657 (0.482)	0.611 (0.491)	0.062 (0.116)	0.265 (0.068)
Number of times elected as ward councilor	1.176 (0.521)	1.450 (0.686)	1.686 (1.451)	1.611 (0.928)	-0.346 (0.227)	-0.340 (0.142)
Other political experience	0.212 (0.415)	0.368 (0.496)	0.286 (0.458)	0.417 (0.496)	-0.114 (0.140)	-0.121 (0.079)
Affiliated with CPI(M)	0.697 (0.467)	0.842 (0.375)	0.686 (0.471)	0.681 (0.470)	-0.227 (0.114)	0.072 (0.067)
Spouse ever elected to Panchayat	0 (0)	0.050 (0.224)	0.029 (0.169)	0 (0)	-0.041 (0.053)	0.011 (0.023)

Notes:

- 1 Wealth Index is based on a principal component analysis using the number of household assets. Other Political Experience is a normalized indicator for whether the Pradhan has or holds other political office, including: ward member, member of higher panchayat, state or national legislator, youth party president, district/block level posts of party, women's committee president.
- 2 Columns (1)-(4) report means with standard deviations in parenthesis. Column (5) reports a test of the difference in means across columns (1) and (2). Column (6) is the difference between currently reserved and unreserved GPs. Standard errors are in parentheses in columns (5)-(6). Tests are based on regressions with block fixed effects and standard errors are clustered by GP.

Table III
2003 and 2008 Electoral Outcomes

	Pradhans			Contestants			Winners		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Only reserved 1998	0.027 (0.023)	0.056 (0.031)		-0.003 (0.011)	-0.009 (0.011)		0.015 (0.020)	0.002 (0.019)	
Only reserved 2003		0.003 (0.026)			-0.007 (0.011)			0.000 (0.018)	
Only reserved once (either 2003 or 1998)			0.031 (0.022)			-0.004 (0.008)			0.012 (0.013)
Reserved 1998 and 2003		0.076 (0.041)	0.079 (0.041)		0.037 (0.014)	0.036 (0.014)		0.057 (0.032)	0.057 (0.030)
Test: Equality of reservation indicators [p values]		0.157	0.253		0.009	0.006		0.224	0.127
Year of election	2003	2008	2003 & 2008	2003	2008	2003 & 2008	2003	2008	2003 & 2008
Never reserved sample:									
Mean	0.092	0.109	0.099	0.076	0.049	0.066	0.083	0.049	0.071
Standard deviation	(0.290)	(0.312)	(0.299)	(0.265)	(0.216)	(0.248)	(0.276)	(0.217)	(0.257)
N	870	875	1745	3880	3431	7311	1425	1191	2616

Notes:

- Columns (1)-(3) show regressions based on Pradhan appointments in GPs not currently reserved for women Pradhans, and columns (4)-(9) are based on GP election results for Ward Councilor seats not currently reserved for women.
- Columns (1)-(3) use data from four districts in West Bengal: Birbhum, Burdwan, Hooghly, Howrah, Nadia, and South 24 Parangas, and the outcome variable is an indicator equal to one if the appointed Pradhan is female. Columns (4)-(9) use the election results from elections in Birbhum district for council member seats not reserved for women. In Columns (4)-(6), the outcome variable is an indicator equal to one if the contestant for a seat is female, while in Columns (7)-(9), the outcome variable is an indicator equal to one if the winner of a seat is female.
- Reserved Once indicates that the GP was reserved in only 1998 or 2003. Reserved 1998, 2003 and 2008 indicates that the GP was reserved in all 3 elections.
- The p-value is from a Wald test of the equality of the coefficients on First Reserved in 2003 and Reserved 1998 and 2003.
- In columns (1)-(3), regressions include district fixed effects, and standard errors are adjusted for heteroskedasticity. In columns (4)-(9), regressions include block fixed effects, and standard errors are clustered by GP.

Table IV
Evaluation of Actual Pradhan: Average Effect

	Male				Female			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Only reserved 2003	-0.197 (0.058)	-0.139 (0.063)	-0.210 (0.058)	-0.152 (0.063)	-0.075 (0.060)	-0.012 (0.061)	-0.076 (0.060)	-0.015 (0.062)
Reserved 1998 and 2003	0.014 (0.072)	0.013 (0.083)	0.013 (0.072)	0.011 (0.083)	-0.006 (0.050)	0.010 (0.057)	-0.004 (0.050)	0.011 (0.057)
Only reserved 1998	0.001 (0.056)	-0.010 (0.057)	0.003 (0.056)	-0.009 (0.056)	0.037 (0.054)	0.018 (0.052)	0.040 (0.054)	0.021 (0.052)
With pradhan characteristics controls	N	Y	N	Y	N	Y	N	Y
With pradhan action controls	N	N	Y	Y	N	N	Y	Y
Test: 2003 = both 1998 and 2003 = 1998 [p value]	0.008	0.124	0.004	0.080	0.216	0.904	0.191	0.866
Test: 2003 = both 1998 and 2003 [p value]	0.012	0.084	0.009	0.065	0.301	0.736	0.285	0.686
N	6642	6642	6642	6642	6568	6568	6568	6568

Notes:

- 1 The outcome variable averages across four questions: "Is Pradhan effective," and Did Pradhan: "Look after village needs"; "Look after your needs"; and "Make BPL lists well."
- 2 All regressions include: (i) Block fixed effects (ii) Individual controls: age, age squared, household size, religion, caste dummies (for scheduled caste, scheduled tribe and other backward caste), years of education, a wealth index (based on a principal component analysis using household assets) and dummy for land ownership (iii) Village controls: all variables in Table I (iv) Survey year and surveyor gender indicator. Standard errors are clustered by GP. Columns (2) and (5) include Pradhan characteristics from Table II, and columns (3) and (7) include the indices of public good quantity and quality (see Table V). Columns (4) and (8) include both Pradhan characteristics and the public good quantity and quality indices.
- 3 We report the p-values from Wald equality tests.

Table V
Pradhan Performance: Public Goods, Bribes and Satisfaction

	Average public good provision		Average satisfaction		Average bribes	Alignment with female preferences
	Quantity	Quality	Male	Female		
	(1)	(2)	(3)	(4)	(5)	(6)
Only reserved 2003	0.192 (0.070)	-0.043 (0.046)	0.037 (0.042)	-0.001 (0.039)	-0.094 (0.031)	0.521 (0.279)
Reserved 1998 and 2003	0.039 (0.061)	-0.030 (0.052)	-0.063 (0.052)	-0.042 (0.044)	-0.072 (0.029)	0.659 (0.358)
Only reserved 1998	0.097 (0.082)	-0.069 (0.037)	-0.008 (0.045)	0.025 (0.038)	-0.045 (0.038)	0.563 (0.243)
Test: 2003 = both 1998 and 2003 = 1998 [p value]	0.127	0.763	0.242	0.381	0.360	0.942
Test: 2003 = both 1998 and 2003 [p value]	0.343	0.847	0.095	0.435	0.493	0.730

Notes:

- 1 The outcome variables are: the average quantity across public goods (Column 1), the average across quality measures for public goods (Column 2), the average across satisfaction with various public goods respectively (Columns (3)-(4)), and averaged bribes (Column 5). Column (6) tests whether there is more investment in reserved GPs in goods mentioned more frequently by women, as measured by formal complaints to the GP in 2000. We report the coefficients from the alignment with female preferences measure in Chattopadhyay and Duflo (2004).
- 2 The sample in columns (1), (2) and (6) regressions are 495 villages, while columns (3)-(5) regressions use household surveys and include the controls defined in Table IV.

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

	Average effect	
	Male	Female
	(1)	(2)
Panel A		
Female Pradhan	-0.054 (0.027)	-0.035 (0.031)
Female Pradhan * ever reserved	0.091 (0.036)	0.024 (0.038)
Test: female Pradhan + female Pradhan * ever reserved	0.038 (0.023)	-0.011 (0.022)
Panel B		
Female Pradhan * only reserved 2003	0.112 (0.047)	-0.001 (0.048)
Female Pradhan * reserved 1998 & 2003	0.092 (0.062)	0.052 (0.060)
Female Pradhan * only reserved 1998	0.073 (0.046)	0.035 (0.045)
Test: FP* 2003 = FP* (both 1998 and 2003) = FP* 1998 [p value]	0.774	0.652

Notes:

- 1 The outcome variables are averages across all questions in speech and vignettes: "Is Pradhan effective?," "Cares about villagers' welfare?," in the speech and vignettes; "Did Pradhan address villagers satisfactorily?," "Will Pradhan allocate BPL cards well?," "Will Pradhan get resources by lobbying?," "Will Pradhan collect villagers' share well?" and "Will Village approves Pradhan's budget?" in the speech; and "Agree with Pradhan" and "Would vote for Pradhan" in the vignettes.
- 2 Female Pradhan is an indicator which is 1 if the leader speaking was female or the Pradhan in the vignettes was female. All regressions include the controls defined in Table IV, and standard errors are clustered by GP.

Table VII
Explicit and Implicit Preferences for Female Leaders

	IAT (D-measure of bias against females)						Feeling ladder	
	Leadership/domestic and male/female		Male/female names and good/bad		Male/female politician and good/bad		Male versus female Pradhan	
	Male (1)	Female (2)	Male (3)	Female (4)	Male (5)	Female (6)	Male (7)	Female (8)
Panel A								
Ever reserved	-0.076 (0.032)	0.021 (0.041)	-0.004 (0.031)	-0.007 (0.043)	0.014 (0.037)	-0.023 (0.038)	0.208 (0.112)	0.099 (0.110)
Panel B								
Only reserved 2003	-0.090 (0.041)	0.112 (0.053)	-0.023 (0.045)	0.005 (0.051)	0.024 (0.051)	-0.004 (0.049)	0.271 (0.158)	0.088 (0.145)
Reserved 1998 and 2003	-0.023 (0.052)	-0.098 (0.075)	0.016 (0.041)	0.035 (0.074)	0.036 (0.057)	-0.011 (0.056)	0.063 (0.159)	0.053 (0.152)
Only reserved 1998	-0.098 (0.042)	-0.022 (0.051)	0.001 (0.045)	-0.061 (0.052)	-0.012 (0.048)	-0.050 (0.051)	0.240 (0.150)	0.139 (0.140)
Test: 2003 = both 1998 and 2003 = 1998 [p value]	0.402	0.021	0.756	0.316	0.704	0.709	0.560	0.875
Never reserved sample:								
Mean	0.110	0.150	0.134	-0.157	0.093	-0.079	1.446	0.560
Standard deviation	(0.340)	(0.384)	(0.425)	(0.418)	(0.452)	(0.441)	(2.655)	(2.572)
N	477	357	510	408	554	510	3511	3671

Notes:

- 1 The outcome variables are: the difference in average response latencies between the stereotypical and non-stereotypical block in the IAT divided by the standard deviation of latencies (IAT D-measure in Columns ((1)-(6)) and the differences in the ranking between male and female Pradhans on a scale of 1-10 (Columns (7)-(8)).
- 2 "Ever Reserved" is an indicator for whether a GP was reserved for a female Pradhan in either 1998, 2003 or in both elections.
- 3 The IAT and ladder questionnaire were administered to adults in a random subset of households per village, and each respondent was administered one of the three IATs.

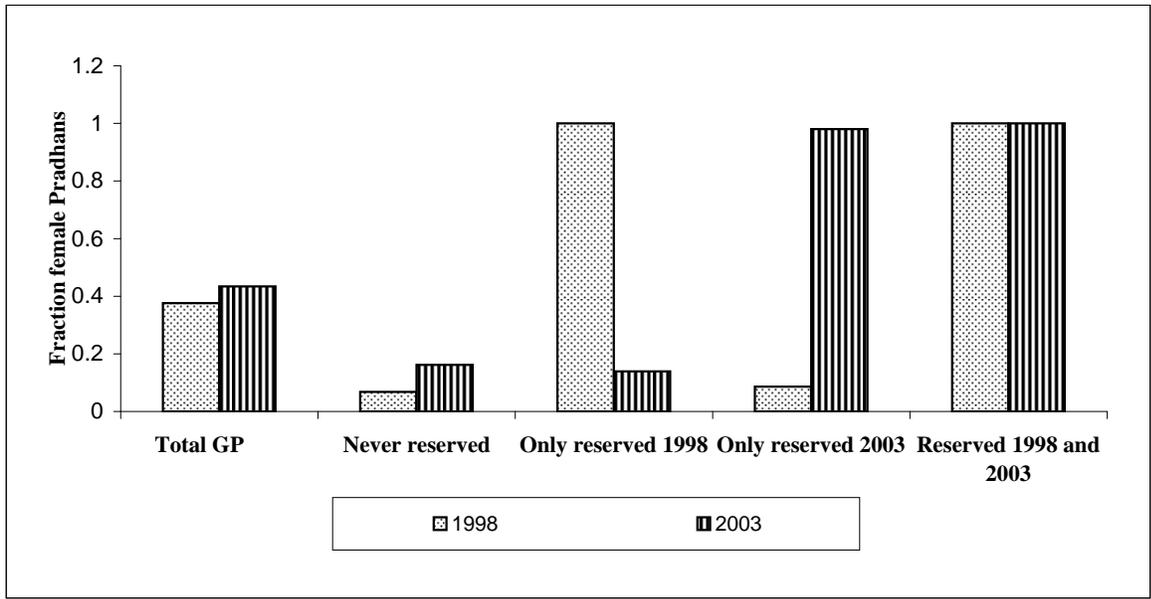


Figure I
Political Leadership and Female Leadership

Note: There are 165 GPs in Birbhum. Of these, 35 GPs were reserved for the first time in 2003 (First Reserved 2003), 20 in both 1998 and 2003 (Reserved 1998 and 2003), 36 GPs only in 1998 (Only reserved 1998), and 74 were never reserved.

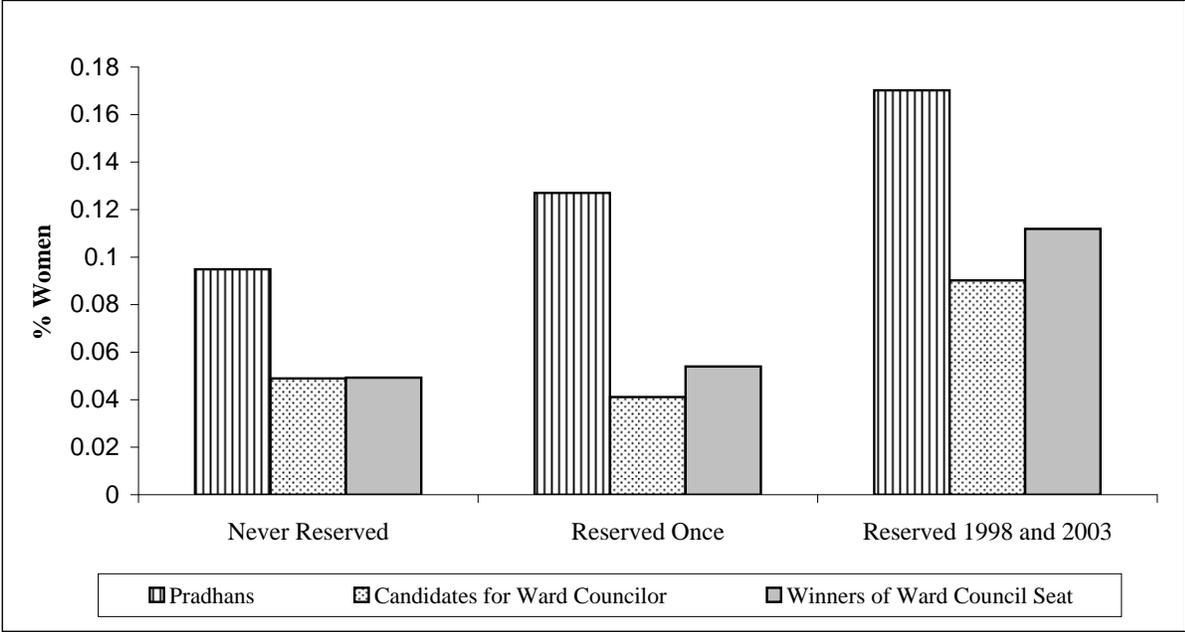


Figure II
2008 Ward Council and Pradhan Election Outcomes

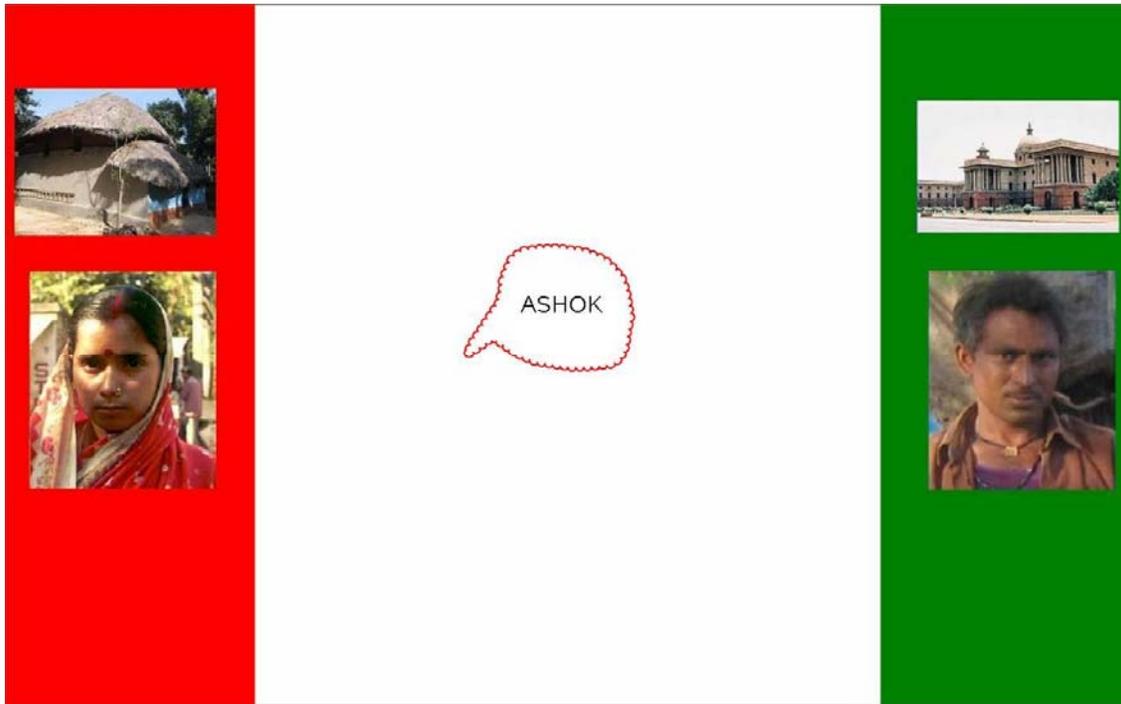


Figure A.1
Gender and Leadership IAT, Stereotypical Block

Notes:

1. The bubble in the middle of the screen represents an oral prompt: in this case, a male name. A respondent should push the right button in response to hearing “Ashok” to indicate it belongs to the male villager category.
2. All three IATs can be found at <http://www.povertyactionlab.org/projects/project.php?pid=102>



Figure A.2
Participant Performing IAT