

# Web Appendix :Are Female Leaders Good for Education?

## Evidence from India.

Irma Clots Figueras\*

Department of Economics. Universidad Carlos III de Madrid.

May 2011

## 1 Data appendix

### 1.0.1 Electoral data:

Collected from different volumes of the Statistical Reports on the General Elections to the Legislative Assemblies. The election commission of India publishes one report for every election in each state. There is data at the constituency level for the 16 main states in India for elections held during 1967-2001.

*-Proportion of seats in the district won by women:* defined as the total number of seats in which a woman won the election in the district divided by the total number of seats in the district. Three years averages for each district are then computed.

*-Proportion of seats won by women in a close election against a man:* defined as the number of women in the district who won by less than 3.5% of votes against a man over the total number of seats in the district. Three year averages for each district are then computed.

*-Proportion of seats in which a man and a woman contested in a close election:* defined as the number of men and women in the district who won by less than 3.5% of votes against a candidate of the other gender over the total number of seats in the district. Three year averages for each district are then computed.

*-Proportion of seats won by each political party:* number of seats won by the political party divided by total seats in the district. Three year averages for each district are then computed.

---

\*Correspondence: Department of Economics. Universidad Carlos III de Madrid. C/Madrid 126, 28903, Getafe (Madrid). Email: iclots@eco.uc3m.es / irmaclots@gmail.com.

Congress parties include Indian National Congress, Indian National Congress Socialist Parties, and Indian National Congress. Hard Left parties include the Communist Party of India and the Communist Party of India Marxist Parties. Soft Left parties include Praja Socialist Party and Socialist Party. Janata parties include Janata, Lok Dal, and Janata Dal parties. Hindu parties include the Bharatiya Janata Party. Regional parties include Telegu Desam, Asom Gana Parishad, Jammu & Kashmir National Congress, Shiv Sena, Uktal Congress, Shiromani Akali Dal, and other state specific parties.

### 1.0.2 NSS Data:

The 55th Round of the National Sample Survey Organization Data. Household Schedule 10: Employment and Unemployment. The survey was conducted in India between July 1999 and June 2000. I use the questions asked to individual members of each household.

-*Primary education attainment*: There is a question that classifies individuals according to whether they are illiterate, literate through attending non-formal education courses or adult education centers, literate through the Total Literacy Campaign or other programmes, literate below primary education, individuals who achieved primary education and individuals who achieved middle, secondary, higher secondary or graduate education. I then create a variable that is equal to one if the individual obtained primary or a higher level of education.

-*School attendance*: There is a question that classifies individuals according to whether they are attending school or not, and if they are not, they ask them whether they have ever attended school.

-*Gender*: Question about gender. A dummy variable is created that is equal to one if the individual is a woman and zero otherwise.

-*Caste*: Question about the individual's social group. Dummy variable that is equal to one if the respondent belongs to the SC/STs and zero otherwise.

-*Religion*: Question about the individual's religion. Two dummy variables that are equal to one if the respondent is Muslim or Hindu, respectively, and zero otherwise.

-*Migration*: Question about the last usual residence. There are various possibilities: same district (urban/rural), same state but another district (rural/urban), another state(urban/rural) and another country. There is another question about the period in years since the individual left the last usual residence.

-*Workers*: Individuals older than 15 who are employed according to "Usual Activity Status"

defined by the NSS. A variable is created that is equal to one if the individual falls in this category and is zero otherwise.

-*Non-agricultural workers*: Classified according to NIC code of the "Usual Activity Status". Workers (as before) who did not work in the primary sector. I create a dummy equal to one if the individual works in the secondary or tertiary sectors and zero if the individual works in the primary sector.

-*Wages*: Wages received for the work done during the reference week as reported by the respondent, including wages in kind and in cash.

### **1.0.3 Demographics and Census:**

Data from 1961-1991 were obtained from the Indian district database created by Vanneman and Barnes. Data from the Indian Census 2001 comes from the webpage "Education For All in India".

-*Data on male and female literacy rates*: literate males (and females) older than 5 over total population of males (and females) older than 5 in the district.

-*Data on SC/ST population*: number of SC/ST individuals over the total population in the district.

-*Data on female population*: number of women over total population in the district.

-*Data on urban population*: number of individuals living in urban areas over total population in the district.

-*Data on schools*: proportion of villages that had a school.

## **1.1 Checks on the Identification Strategy**

This section shows some facts that support the validity of the identification strategy used. Three issues are addressed. First of all, evidence is provided that the outcome of a close election cannot be predicted by observables, which supports the idea that the outcome of a close election can be considered as good as random. In addition, it is shown that districts and constituencies in which female candidates won in close elections against men are similar in observables to those in which male candidates won in close elections against women. Finally, it is shown that districts that had close elections between men and women are not systematically different from other districts in India.

### 1.1.1 Randomness of Close Election Outcomes

If there are political or demographic characteristics that predict the probability that women win in close elections in the district, the outcome of the close elections and, thus, the gender of the winners cannot be considered random. To estimate the probability that women won in close elections in a district, I calculate the proportion of close elections won by women by district in each electoral year. This proportion is then regressed on the fraction of seats contested by the different party groupings in close elections, the proportion of urban population, the proportion of female and SC/ST population, male and female literacy rates, the number of times that women have won elections in the past in that district and the proportion of reserved seats. Results are shown in Table A1, and they confirm that none of the coefficients turns out to be significant, suggesting that the outcome of a close election can indeed be considered as good as random.

### 1.1.2 Comparing on Observables

If the winner's gender in a close election between a man and a woman can be considered as good as random, we expect that districts in which more women won in close elections should be very similar to districts in which more men won in close elections.

Table A2 provides information on the differences in district characteristics according to the number of women who won against men and the number of men who won against women in the district. Districts are classified into two groups, those in which more men won and those in which more women won.

I then compute the differences in district characteristics between these two groups. This is done considering the elections in which the winner has led over the runner-up by margins of 3.5% of votes. I use information at the district level on the proportion of urban and SC/ST population, male and female literacy rates, the number of seats, the fraction of seats reserved for SC/STs, the proportion of villages with educational institutions and hospitals and the proportion of seats won by female and male candidates in elections that are not close. All columns show that districts in which more men won in close elections with this or a smaller margin and districts where more women won in close elections with this or a smaller margin are very similar in all these variables.

In summary, districts in which more women won in close elections are very similar to districts in which more men won in close elections.

One should also observe that constituency and individual characteristics of women and men winning in close elections are the same. I also analyze some of these characteristics that could compromise the comparability between close elections in which men won and close elections in which women won.

First of all, there might be concerns that two different constituencies in which a woman contested in a close election against a man might not be similar if in one of them there were many other women candidates, apart from the winner or the runner-up, contesting for the same seat. This would be a case in which political parties perceive the constituency as one in which there is “preference for female politicians” and tend to field female candidates there. If the number of female candidates contesting for the same seat as the two close candidates is significantly different for constituencies in which a man won in a close election against a woman and constituencies in which a woman won in a close election against a man, these two types of constituencies might have different characteristics. I have data on all the female candidates contesting in a particular constituency, apart from the winner and the runner-up. As shown in the top panel of Table A3, the number of other female candidates contesting against women who won in close elections against a man is very small and not significantly different from that for men who won in close elections against a woman.

It might also be that one of the candidates in a close election is in this situation because he or she is the incumbent for that seat in that particular constituency. This would make constituencies in which women and men won in close elections against a candidate of the other gender different in observables if men (or women) are those who tend to be the incumbent. Moreover, if there is incumbency advantage (or disadvantage) in these elections, more women (or men) would win in this type of elections and one could question the extent to which the outcome of a close election can be considered as good as random. It should also be taken into account that the policies applied by candidates who were the incumbent and who won the elections again might be different from those of candidates who occupy the seat for the first time, because they will have more experience as legislators. To address this concern I use the fact that I have information on the candidate’s names, thus, I can know whether a particular candidate was already in power during the previous electoral year in the same constituency where he or she is now contesting. I then create a dummy variable that is equal to one if the individual was the incumbent for that seat. However, as is shown in the second panel of Table A3, the percentage of winners in close elections who were the incumbent is statistically the same

for female and male legislators who won in close elections.

Another concern that needs to be addressed is that maybe there are some constituencies in which there have been more close elections between men and women in the past than in others. If this happens more often in constituencies where women won the close election than in constituencies in which men won, then these two types of constituencies would not be comparable, since in the one where there have been more close elections there would probably be more “preference for female politicians”. In the third panel of Table A3, I test whether constituencies in which a man or a woman won in a close election are different in terms of how many times the particular constituencies have had close elections between men and women. However, results show that the number of previous close elections is the same, whether a woman or a man won. Thus, women won in close elections in situations in which the electoral preferences for female politicians are similar to situations in which men won in close elections.

Finally, if elections in which men and women won in close elections are similar, they should have the same electoral turnout, otherwise, one type of constituency would be more active in electoral terms than the other. And, more importantly, the distribution of votes between the first two candidates and the rest should be the same. This is the case because if in one case the total votes were distributed among many candidates, these could not be considered as close elections between the winner and the runner-up. The last two panels of Table A3 show that women who won in close elections won by the same number of votes as men who won in close elections, and in constituencies where the total number of votes was the same. Because constituencies in India were designed to have the same population, this means that turnout was the same, and the distribution of votes between the first candidate and the rest was the same as well. This further corroborates that constituencies in which a man or a woman won in a close election are perfectly comparable and thus, the gender of the winner can, indeed, be considered as good random. These two panels also eliminate concerns that, if in a constituency there were three candidates with almost the same number of votes, one could not consider the election between the winner and the runner up as a close election. In fact, the winners in close elections tend to receive approximately 40% of votes, which means that the runner-up will receive a minimum of 36.5% of votes. This leaves the other candidates with 23.5% of votes, which is a very large difference compared to the winner. Thus, even if there was another strong candidate in the constituency, he or she did not have any chance of winning the election<sup>1</sup>.

---

<sup>1</sup>As was proven before, there are no concerns regarding the gender of these other candidates.

### 1.1.3 External Validity

Overall, between 1967 and 2001, 141 out of 297 districts never had a close election between a man and a woman, which is slightly less than half the districts in my sample. However, it could be argued that close elections between men and women take place in districts that are different, or more progressive, than the average district in India. Even if there is a significant number of individuals affected by close elections, if districts that never had close elections are very different from those that did, results obtained in this paper would not be representative for the entire country of India. Table A4 shows that districts that never had close elections and those that did are similar in observables. For districts that never had close elections and districts that did, it shows descriptive statistics for population characteristics, the proportion of reserved seats, the total number of seats, and public goods such as hospitals and educational institutions weighted by the population in the years when elections took place.

Finally, there might be concerns that the probability of contesting a close election between a woman and a man is different for each political party. If this were the case, close elections would not reflect the overall situation in the parliament because only a few parties would be involved. Table A5 shows how the distribution of seats among the different party groupings is the same for close elections between men and women as for the rest. Thus, party composition seems not to be a concern, because the party composition in close elections reflects that of the overall parliaments in the states.

### 1.1.4 Covariate Balance

In Table A7 I check whether pre-determined covariates also display jumps at the discontinuity point. First I use information on other individual characteristics, such as religion, whether the individual is SC/ST or not and gender. These are not really pre-determined, but if the population composition in the district remains constant over time, they can be considered as such. I also use information on district-level variables from the census, such as the proportion of the population that is female, SC/ST or lives in urban areas, together with adult male and female literacy rates when the individual was young. I first run the 2SLS specification using these controls as dependent variables. I do this for the urban sample, the rural sample and the whole sample. The coefficients are not significant, except for the adult female literacy rate when the individual was young. Then I run the RD specification with two second order polynomials,

---

one at each side of the discontinuity using the covariates as dependent variables. All coefficients are very small and only three are significant when I control for district and cohort fixed effects.

## 1.2 Education in India

In India, although within the States and Union Territories there are differences in the number of years constituting primary, middle and secondary education, there is a uniform structure of school education.

The primary stage consists of classes I-V, in Andhra Pradesh, Bihar, Haryana, Jammu & Kashmir, Madhya Pradesh, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal.<sup>2</sup> On the other hand, it consists of classes I-IV in Assam, Gujarat, Karnataka, Kerala, and Maharashtra.

The middle stage consists of classes VI-VIII in Bihar, Haryana, Jammu & Kashmir, Madhya Pradesh, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal. Classes V-VII in Assam, Gujarat, Karnataka, Kerala, and Maharashtra, and classes VI-VII in Andhra Pradesh and Orissa.

The secondary stage consists of classes IX-X in Bihar, Haryana, Jammu & Kashmir, Madhya Pradesh, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal. And classes VIII-X in Assam, Gujarat, Karnataka, Kerala, Maharashtra, Andhra Pradesh, and Orissa.

The minimum age for admission in the first class of the primary stage is 5 or 6 years of age, depending on the State or Union Territory. The majority of States and Union Territories have established free education; however, in some States education is not free for classes IX and above<sup>3</sup>.

---

<sup>2</sup> Among others. Only the 16 main states in India are considered in this study.

<sup>3</sup> The highest annual fee is Rs. 360 in Meghalaya, while the lowest is Rs. 48 in Assam. Mean annual household income is around Rs. 34551.

**Table A1: Probability that a Woman Wins in a Close Election against a Man**

Dependent variable: proportion of women who won in a close election against a man per district and electoral year. Sample 1967-2001.

	1
Proportion of seats contesting close elections Congress	-1.412 (2.607)
Proportion of seats contesting close elections Regional Parties	-3.332 (4.882)
Proportion of seats contesting close elections Hindu	-1.247 (2.706)
Proportion of seats contesting close elections Janata	-1.81 (2.075)
Proportion of seats contesting close elections Others	-0.433 (2.389)
Proportion of seats contesting close elections Independent	-1.546 (2.303)
Dummy=1 if the district never had close elections before	0.241 (0.635)
Proportion of urban population	12.587 (12.85)
Number of times that a woman has won an election in the district in the past	-0.006 (0.051)
Proportion of SC/ST population	18.497 (19.496)
Proportion of population that is female	-15.535 (27.662)
Male literacy rate	-1.923 (10.817)
Female literacy rate	-0.494 (7.249)
Proportion of seats reserved for SC/ST's	-2.931 (5.313)
Observations	164
Adjusted R-squared	-0.059

Robust standard errors clustered at the district level. District and year fixed effects are included in the regression.

**Table A2: District Characteristics: Close Elections between Women and Men**

Unit of observation: district- electoral year. Sample 1967-2001

---

---

Differences in the proportion of urban population	-0.0102
(Districts in which more men than women won compared to districts in which more women than men won)	0.0179)
Differences in male literacy rate	-0.0292
(Districts in which more men than women won compared to districts in which more women than men won)	0.0255)
Differences in female literacy rate	-0.0282
(Districts in which more men than women won compared to districts in which more women than men won)	(0.0296)
Differences in the proportion of villages with educational institutions	0.0147
(Districts in which more men than women won compared to districts in which more women than men won)	(0.0342)
Differences in the proportion of villages with hospitals	0.0055
(Districts in which more men than women won compared to districts in which more women than men won)	(0.0107)
Differences in the proportion of SC/ST reserved seats	-0.0038
(Districts in which more men than women won compared to districts in which more women than men won)	(0.0272)
Differences in the proportion of seats won by women who won in elections that are not close	-0.0013
(Districts in which more men than women won compared to districts in which more women than men won)	(0.0083)
Differences in the proportion of seats won by men who won in elections that are not close	-0.0123
(Districts in which more men than women won compared to districts in which more women than men won)	(0.0103)
Number of districts with election-years	201

---

---

**Table A3****Constituency and candidate characteristics: Close Elections between Women and Men**

Unit of observation: candidate. Sample 1967-2001.

Group	Obs	Mean	Std. Err.
<b>Other female candidates in the constituency</b>			
Man won in close election	120	0.1083	0.0370
Woman won in close election	110	0.2000	0.0480
Difference		-0.0917	0.0601
<b>Winner was the incumbent</b>			
Man won in close election	120	0.2167	0.0378
Woman won in close election	110	0.2182	0.0396
Difference		-0.0015	0.0547
<b>Number of close elections in the past</b>			
Man won in close election	120	1.0750	0.0241
Woman won in close election	110	1.0727	0.0249
Difference		0.0023	0.0347
<b>Votes received by the winner</b>			
Man won in close election	120	31894.1700	1328.4220
Woman won in close election	110	33596.4500	1330.2330
Difference		-1702.2880	1883.4150
<b>Total votes in the constituency</b>			
Man won in close election	120	80188.3300	2769.9040
Woman won in close election	110	80947.2700	2655.8640
Difference		-758.9394	3851.7720

**Table A4****Comparison: Districts with and without Close Elections**

(District in an electoral year, sample 1967-2001)

		Close elections	No close elections
Urban population (prop)	mean	0.2149	0.1947
	sd	0.0052	0.0039
	observations	968	1124
Male literacy rate	mean	0.5241	0.5454
	sd	0.0054	0.0049
	observations	946	1098
Female literacy rate	mean	0.2878	0.2865
	sd	0.0059	0.0054
	observations	946	1098
SC/ST population (prop)	mean	0.2618	0.2443
	sd	0.0047	0.0038
	observations	968	1124
SC/ST seats proportion	mean	0.2564	0.2178
	sd	0.0055	0.0054
	observations	1223	1323
Seats total	mean	10.8397	7.9426
	sd	0.1382	0.1259
	observations	1223	1323
Any educational institution	mean	0.8212	0.7932
	sd	0.0079	0.0084
	observations	316	348
Hospitals	mean	0.0290	0.0262
	sd	0.0025	0.0025
	observations	736	812

**Table A5****Proportion of Seats Won by Parties**

Party	Close Elections	No close elections
	Percent	Percent
Congress	40.43	41.16
Hard Left	7.83	8.17
Hindu	11.74	11.44
Independents	6.96	5.81
Janata	9.57	13.98
Regional	12.61	10.38
Soft Left	3.91	2.31
Others	6.96	6.75
Total	100	100

Sample 1967-2001

**Table A6: Baseline results: complete regressions (2SLS)**

Dependent variable: primary education attainment (1=primary education or higher, 0=otherwise)

	2SLS	2SLS	2SLS
	All	Urban	Rural
	1	2	3
Proportion of seats won by women	0.0298 (0.180)	0.7284** (0.295)	-0.1098 (0.217)
Proportion of seats that had close elections between women and men	0.1436 (0.104)	-0.2081 (0.154)	0.2752** (0.137)
Proportion of the population that is urban	-0.3650* (0.211)	-0.0341 (0.189)	-0.1749 (0.276)
Proportion of the population SC/ST	0.5660*** (0.165)	0.4443* (0.244)	0.5909*** (0.163)
Proportion of the population female	-0.1307 (1.522)	-1.1342 (2.494)	0.7585 (1.737)
Male literacy rate in the district	0.2366 (0.239)	-0.4587 (0.301)	0.1962 (0.262)
Female literacy rate in the district	0.2435 (0.155)	0.1117 (0.232)	0.4267** (0.174)
Muslim dummy	-0.2357*** (0.020)	-0.2454*** (0.024)	-0.2159*** (0.024)
Hindu dummy	-0.0666*** (0.013)	-0.0536*** (0.015)	-0.0573*** (0.017)
SC/ST dummy	-0.1925*** (0.009)	-0.1884*** (0.016)	-0.1888*** (0.010)
Female dummy	-0.2055*** (0.007)	-0.0631*** (0.006)	-0.2388*** (0.008)
Rural dummy	-0.2120*** (0.011)		
Observations	105208	34604	70604

Robust standard errors clustered at the district level. District and year fixed effects are included in the regression.

**Table A7: Covariates**

Coefficient for the fraction of constituencies in the district that had close elections between women and men.

	Urban		Rural			Whole Sample			
	2SLS	RD	RD fe	2SLS	RD	RD fe	2SLS	RD	RD fe
Muslim dummy	-0.0083 (0.190)	0.0050 (0.053)	0.0510* (0.026)	0.1124 (0.095)	-0.0029 (0.036)	0.0177 (0.022)	0.1113 (0.096)	0.0243 (0.025)	0.0226 (0.019)
Hindu dummy	0.2328 (0.277)	0.0226 (0.058)	-0.0344 (0.029)	-0.1178 (0.128)	0.1010* (0.056)	-0.0364 (0.026)	-0.0932 (0.116)	0.0166 (0.036)	-0.0313 (0.021)
SC/ST dummy	-0.0316 (0.195)	-0.0235 (0.038)	-0.0236 (0.024)	-0.1975 (0.186)	0.0433 (0.056)	-0.0437 (0.030)	-0.2127 (0.160)	-0.0166 (0.029)	-0.0333 (0.025)
Female dummy	-0.2866 (0.420)	-0.0263 (0.036)	-0.0798 (0.062)	-0.0993 (0.279)	-0.0008 (0.023)	0.0354 (0.035)	-0.1819 (0.255)	0.0021 (0.015)	0.0210 (0.032)
Proportion of the population that is urban	-0.0371 (0.092)	0.0498 (0.069)	0.0048 (0.010)	0.0201 (0.036)	-0.0250 (0.034)	0.0063 (0.006)	0.0103 (0.045)	0.0239 (0.028)	0.0069 (0.007)
Proportion of the population SC/ST	0.0257 (0.074)	0.0758* (0.043)	0.0011 (0.010)	0.0252 (0.040)	0.1047*** (0.039)	0.0030 (0.005)	0.0244 (0.045)	0.0176 (0.021)	0.0028 (0.006)
Proportion of the population female	0.0059 (0.007)	0.0037 (0.005)	0.0031 (0.002)	-0.0030 (0.006)	0.0071 (0.004)	0.0008 (0.001)	-0.0015 (0.006)	0.0018 (0.003)	0.0014 (0.001)
Male literacy rate in the district	-0.0741 (0.057)	0.0222 (0.045)	-0.0013 (0.010)	-0.0342 (0.048)	-0.0374 (0.048)	0.0053 (0.007)	-0.0413 (0.049)	0.0255 (0.029)	0.0041 (0.007)
Female literacy rate in the district	-0.2550** (0.099)	0.0081 (0.055)	-0.0063 (0.014)	-0.1425** (0.068)	-0.0856 (0.053)	0.0067 (0.010)	-0.1650** (0.072)	0.0045 (0.034)	0.0052 (0.011)

Robust standard errors clustered at the district level.