

ONLINE APPENDIX

Health and the Political Agency of Women

Sonia Bhalotra and Irma Clots-Figueras

Appendix Figures and Tables

Figure A1: Fraction of seats in the district won by a female politician.

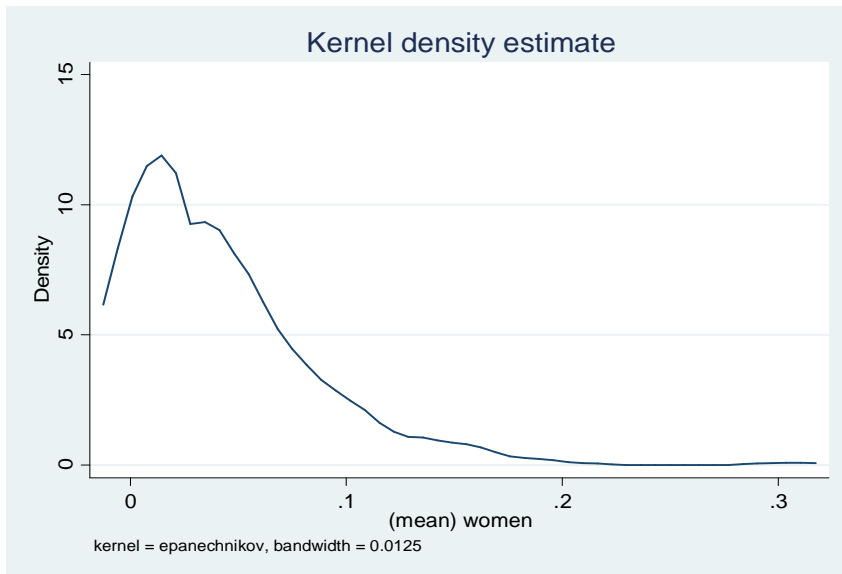


Figure A2: Predicted neonatal mortality risk against the share of women politicians in the district of birth (using district averages)

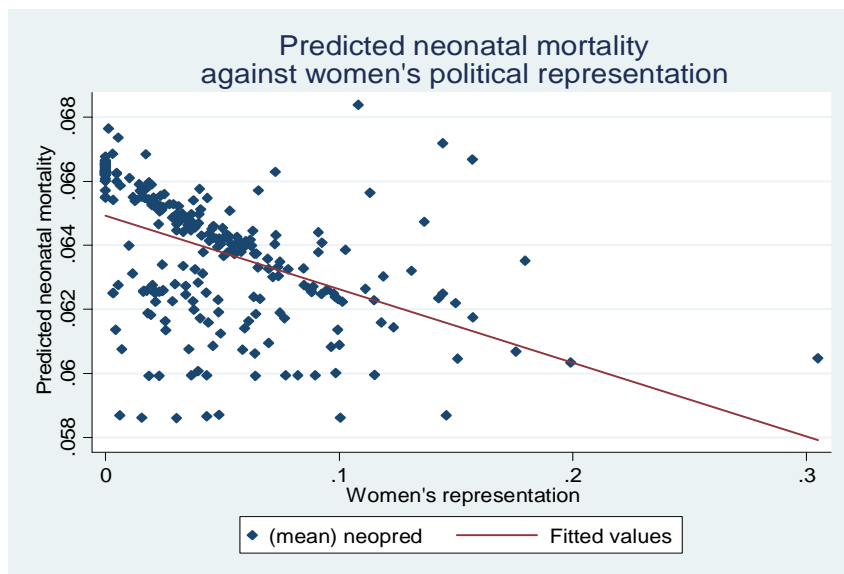


Table A1: Probability that a Woman Wins in a Close Election against a Man

Dependent variable: proportion of women who won in close elections 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**Constituency and Candidate Characteristics: Close Elections between Women and Men**

Unit of observation: candidate. Sample 1967-2001.

Group	Man won in close election	Woman won in close election	Difference
Other female candidates in the constituency			
Observations	120	110	
Mean	0.2833	0.3636	-0.0803
Std Error	0.0582	0.0616	0.0847
Winner was the incumbent			
Observations	120	110	
Mean	0.2167	0.2182	-0.0015
Std Error	0.0378	0.0396	0.0547
Number of close elections in the past			
Observations	120	110	
Mean	1.0750	1.0727	0.0023
Std Error	0.0241	0.0249	0.0347
Votes received by the winner			
Observations	120	110	
Mean	31894.1700	33596.4500	-1702.2880
Std Error	1328.4220	1330.2330	1883.4150
Total votes in the constituency			
Observations	120	110	
Mean	80188.3300	80947.2700	-758.9394
Std Error	2769.9040	2655.8640	3851.7720

Note: Although the mean and median differences in the number of other female candidates are not significantly different between constituencies in which a man narrowly wins against a woman vs those in which a woman narrowly wins against a man, we investigated the distribution of this variable. We can also reject the null that the two distributions are different: the Mann-Whitney-Wilkinson test which is robust to small numbers of observations has a p- value of 0.2718. However, in the subset of cases where the number of other female candidates is two, women appear more likely to win. We therefore re-estimated the model (a) controlling for other female candidates in the first and second stage of the IV estimation and (b) dropping districts in which there is a constituency where the number of other female candidates is 2. The estimates are robust to these variations, see Table A7.

Table A3**Comparison: Districts with and without Close Elections**

(District in an electoral year, sample 1967-2001)		Close elections	No close elections	P Val. Diff
Urban population (prop)	mean	0.2149	0.1947	0.0017
	sd	0.0052	0.0039	
	observations	968	1124	
Male literacy rate	mean	0.5241	0.5454	0.0034
	sd	0.0054	0.0049	
	observations	946	1098	
Female literacy rate	mean	0.2878	0.2865	0.8719
	sd	0.0059	0.0054	
	observations	946	1098	
SC/ST population (prop)	mean	0.2618	0.2443	0.0031
	sd	0.0047	0.0038	
	observations	968	1124	
SC/ST seats proportion	mean	0.2564	0.2178	0.0000
	sd	0.0055	0.0054	
	observations	1223	1323	
Seats total	mean	10.8397	7.9426	0.0000
	sd	0.1382	0.1259	
	observations	1223	1323	
Any educational institution	mean	0.8212	0.7932	0.0154
	sd	0.0079	0.0084	
	observations	316	348	
Hospitals	mean	0.0290	0.0262	0.4327
	sd	0.0025	0.0025	
	observations	736	812	

Table A4**Proportion of Seats Won by Parties**

Party	Close Elections	No close elections	P val. Diff
	Percent	Percent	
Congress	40.43	41.16	0.8249
Hard Left	7.83	8.17	0.8483
Hindu	11.74	11.44	0.8857
Independents	6.96	5.81	0.4609
Janata	9.57	13.98	0.0542
Regional	12.61	10.38	0.2706
Soft Left	3.91	2.31	0.1078
Others	6.96	6.75	0.9017
Total	100	100	

Sample 1967-2001

Table A5: 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 A6: Full set of covariates for baseline regression (column 6, Table 2)

VARIABLES	neonatal mortality
Fraction of seats in district won by female politician	-0.2061*** (0.078)
Fraction of seats in the district that had close elections between women and men	0.0258 (0.037)
Female literacy rate	-0.1407 (0.119)
Male literacy rate	0.0327 (0.138)
Proportion of SC/STs	-0.1514 (0.120)
Proportion of the population that is urban	-0.0202 (0.120)
Proportion of the population that is female	0.3528 (0.858)
Female dummy	-0.0096*** (0.002)
Birth order 2	-0.0312*** (0.003)
Birth order 3	-0.0573*** (0.006)
Birth order 4	-0.0817*** (0.008)
Birth order 5	-0.0999*** (0.011)
Birth order 6	-0.1142*** (0.014)
Birth order 7	-0.1186*** (0.016)
Birth order 8	-0.1292*** (0.020)
Birth order 9	-0.1454*** (0.023)
Birth order 10	-0.1887*** (0.029)
Birth order 11	-0.1089*** (0.040)
Birth order 12	-0.2282*** (0.079)
Birth order 13	-0.0064 (0.230)
Birth order 14	-0.2419*** (0.086)
Multiple birth	0.3011*** (0.019)
Observations	71,498
Number of mothers	18,754

Table A7: Other Female Candidates

	1	2
Neonatal mortality		
Fraction of seats in the district won by a female politician	-0.2106*** (0.079)	-0.2024** (0.081)
Controlling for number of other women candidates	x	
Dropping districts with constituencies that had 2 other female candidates		x
Mother FE	x	x
Cohort FE	x	x
Controls	x	x
Observations	71,498	70,027

See Notes to Table 2. We instrument with close election outcomes as in the baseline specifications for mortality and inputs to health and survival.

Table A8: Further robustness checks: migration and politician caste

VARIABLES	1	2	3	4
	Neonatal mortality			
Fraction of seats in the district won by a woman	0.0592 (0.2364)	-0.1089* (0.0640)	-0.1703** (0.0728)	
Fraction of seats in the district won by low caste (SC/ST) women				-0.1109 (0.085)
Fraction of seats in the district won by higher caste women				-0.1340* (0.077)
Robustness	<i>migrant placebo</i>	<i>No restriction</i>	<i>Child born current loc</i>	<i>gender*caste</i>
Cohort FE	X	X	X	X
Mother FE	X	X	X	X
Controls	X	X	X	X
Margins	3rd ord pol.	3rd ord pol.	3rd ord pol.	3rd ord pol.
Observations	9921	129417	102748	71,498
Number of seqid	2951	33779	27633	18,754

See Notes to Table 2. Using the specification in column 6 of Table 2, column 1 runs the placebo on migration described in the text, effectively matching children to the "wrong" set of politicians. Column 2 removes the baseline restriction which resulted in using only the sample of mothers who had never migrated. Column 3 uses the weaker restriction which is that we include only children who were born in their current place of residence- so we definitely have them correctly matched to politician gender but we exclude any siblings they may have for whom the match is uncertain. In column 4, we divide the female representation variable by caste.

Table A9: Further falsification exercise: Mother's height

	(1)	(2)	(3)	(4)
Mother's height (mean 151.5 sd 5.77)				
Fraction of seats in the district won by women (year before oldest child was born)	1.2731 (3.133)	1.1662 (2.991)		
Fraction of seats in the district won by women (year before youngest child was born)			-0.2743 (3.132)	-0.6892 (3.008)
Mother group FE		x		x
District FE	x		x	
Cohort FE	x	x	x	x
Controls	x	x	x	x
Observations	21,129	21,129	20,795	20,795

See Notes to Table 2. We instrument with close election outcomes as in the baseline specifications for mortality and inputs to health and survival. Now mother groups are defined using the same variables as before but height.

TABLE A10: Post-neonatal mortality

	OLS	2SLS	2SLS	2SLS
	1	2	3	4
Fraction of seats in district won by female politician	0.0155 (0.014)	0.0483 (0.059)	0.0219 (0.074)	0.0104 (0.074)
District FE	X	X		
Cohort FE	X	X	X	X
Mother FE			X	X
Controls				X
Margins		3rd ord pol.	3rd ord pol.	3rd ord pol.
Full sample				
Observations	63,876	63,876	63,876	63,876
Number of mothers			17,647	17,647

See Notes to Table 2. The baseline model is col. 6 of Table 2, the difference being that the dependent variable neonatal mortality is replaced with post-neonatal mortality. The absence of any effects may be because (a) postneonatal mortality is necessarily conditional upon (endogenous) survival through the neonatal period, (b) attribution of neonatal mortality to state policy interventions that improve antenatal and early postnatal care is clearer than is the case for post-neonatal mortality, which is more sensitive to household-level variables including nutritional investments in children and the hygiene, sanitation and pollution levels in the household, or (c) neonatal mortality is more sensitive to investments in mothers (as opposed to children) than post-neonatal mortality.

Table A11: Mortality rates for older children

	(1)	(2)	(3)
Mortality rates (age)	0-4	5-9	10-14
Fraction of seats in the district won by a female	-0.0654 (0.124)	-0.0258 (0.055)	-0.0574 (0.043)
Controls	x	x	x
Year Dummies	x	x	x
Mother FE	x	x	x
Observations	60,812	44,860	30,754
Number of mothers	16,001	12,075	8,755

See Notes to Table 2. We instrument with close election outcomes as in the baseline specifications for mortality and inputs to health and survival.

Table A12: Effect of Female Politicians on Births: Fertility and Gender of Birth

	1	2	3	4	5	6	7	8
Dependent variable:	birth	birth	birth	birth	female	female	female	female
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Fraction of seats in the district won by a female politician	0.0913 (0.057)	0.0800 (0.059)	-0.0058 (0.101)	0.0100 (0.082)	0.2997** (0.125)	0.3008** (0.127)	0.2533* (0.147)	0.2217 (0.143)
District FE	x	x			x	x		
Cohort FE	x	x	x	x	x	x	x	x
Mother FE			x	x			x	x
Controls		x		x		x		x
Observations	379,584	379,584	379,584	379,584	71,873	71,873	71,873	71,873
Number of seqid			25,171	25,171			18,842	18,842

See Notes to Table 2. As discussed in the text, we expand the data to include for every mother a panel of every year at which she was at risk of birth. We then estimate the probability of continuing fertility (1 if birth, 0 if not, in each year, for each woman) in columns 1-4 and the probability that the birth is female conditional upon birth in columns 5-8. Since columns 1-4 suggest that there is no endogeneity in birth, the conditioning on birth in columns 5-8 is unlikely to create any significant selectivity. The four columns in each set of regressions parallel the sequential specifications in Table 2. We instrument with close election outcomes as in the baseline specifications for mortality and inputs to health and survival.

Data Appendix

Electoral data are collected from different volumes of the Statistical Reports on General Elections to State Legislative Assemblies. The election commission of India publishes a report for every state election. We have data at the constituency level for the 16 main states for elections held during 1967-2001. For aggregation of constituency data to the district level we use *State Elections in India*, a publication of the Election Commission which lists the constituencies included in each district in each election year together with Constituency Delimitation orders. Some districts have divided, some have been newly created and others have disappeared during the sample period. We use the 1991 census district definition and include only those districts that were intact over time. We dropped from the political data 196 districts that had changed boundaries since 1967. We lose a further 21 districts in merging the political data to the electoral data. The included and excluded districts have similar infant and neonatal mortality rates, health-seeking behaviours, year of birth, education of the mother and caste. Results are available from the authors on request. Most terms are defined in the text. Amongst controls we include the proportion of seats in the district won by each political party. Following Besley and Burgess (2002) we construct the following party groups. 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 Alkali Dal, and other state specific parties.

Survival and Health Variables

Neonatal mortality refers to death in the first month of life. *Infant* mortality measures mortality in the first year of life. To allow for age-heaping in the data at one and twelve months, we define the mortality indicators as inclusive of the terminal date. The samples used for regressions are adjusted to allow every child full exposure to the relevant risk. For example, for analysis of neonatal mortality we drop children born less than a month before the date of the survey.

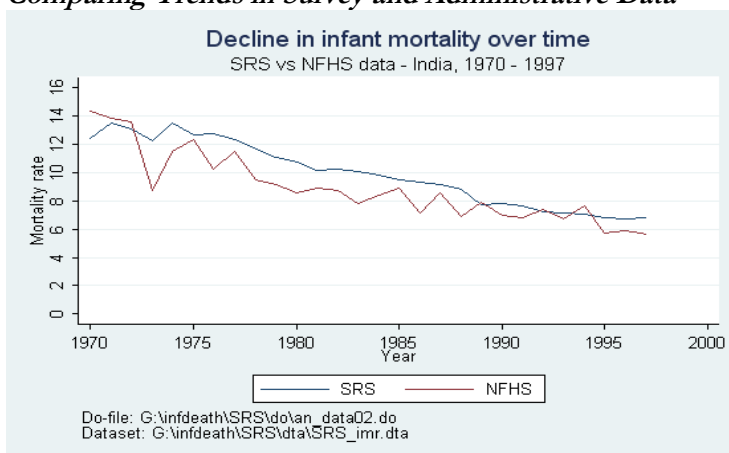
Place of delivery is classified as being either home or at a facility and facilities are further classified as government *vs* private. We construct three indicators corresponding to these place alternatives. *Breastfeeding* is very prevalent in India so we do not use an indicator for whether or not it occurs. The NFHS data contain detailed information on initiation of breastfeeding and its duration. Its duration is often interrupted by disease or death of the child or illness of the mother, so we do not use it. Instead, we define an indicator for whether or not the mother initiated breastfeeding in the first 24 hours following a birth. Indian and especially Hindu mothers often sacrifice the first milk, containing colostrum, to the earth as a matter of tradition. Colostrum contains nutrients and antibodies that are especially important in an environment where under-nutrition and disease are prevalent. *Antenatal care* is measured as the number of visits sought from a health worker.

Migration and other sample restrictions- Some births in the sample occurred when the mother was at a place other than her current residence. As we do not have the date of migration, we apply a stricter criterion than necessary, restricting the sample to mothers for whom all births occurred in the mother's current place of residence. If we ignore migration status, the potential sample contains 129,417 births to 33,779 mothers (column 2, Table A8). Conditioning upon the index child being born in the location of the mother at the survey date lowers this to 102,748 births of 27,633 mothers (column 3, Table A8). The restriction employed in Table 2 (column 6), which is that all births of a mother in the sample occurred in her current place of residence limits the sample further to 71,498 births of 18,754 mothers.

Issues with retrospective fertility data: The data used for the analysis of neonatal and infant mortality are wedge-shaped because of an upper limit on the age of women at the time of interview. In particular, older births in the data are born disproportionately of women who were young at the time of birth (Rindfuss et. al. 1982), and the sample of women, while representative at the date of survey, becomes increasingly unrepresentative in age at birth as we go further back

in time. Under the assumption that calendar time variation in the distribution of age at birth in the sample is uncorrelated with the probability that women politicians win in close elections against men, this will not bias our estimates.¹ In panel-B of Table 3 we included flexible controls for the age of the mother at birth and established robustness of the baseline results to this variation (columns 1, 2, 5). Another concern with retrospective data relates to recall bias, that is, that individuals are more likely to mis-report events that are further in the past. Following a tradition in demography (e.g. Guo 1993, Sastry 1997a, b), we discard information on children born in early years when the data are relatively thin and unrepresentative, in our sample, births before 1967. The trend in infant mortality rates in our data compares well with the trend in administrative data: see the figure below.² In relation to these potential concerns, it is important to underline that we do not identify the coefficient of interest from cohort variation. Also, our estimates for health inputs (Tables 6 and 7) do not use retrospective data and they show similar patterns to the estimates for mortality.

Comparing Trends in Survey and Administrative Data



Notes: The administrative data are data collected under the Sample Registration System by the Census office of the Government of India

The data exhibit some age heaping at 6-month intervals. To account for this, we code infant mortality to include death in the twelfth month (rather than follow the convention of defining it as death strictly before the twelfth month) and, similarly, we code neonatal mortality to include deaths on the 28th day. This makes no difference to our estimates.

To the extent that some women do not survive to the date of interview, we again have a select sample of births. This will not bias our estimates. In principle it will limit their external validity but maternal mortality, the most likely cause of attrition of the mother cohorts, is not a large percentage: it was estimated at 540 per 100,000 births in 2000 (World Bank Health, Nutrition and Population Data).

¹ We directly investigated whether age at birth is endogenous, i.e. correlated with the gender of politicians. To do this we estimated an equation for the age of the mother at birth as a function of the share of women politicians in the district of birth using the specification in col.6 of Table 2. The estimated coefficient on the share of women politicians is -0.047 (s.e. 0.13, p-value 0.72) which is statistically insignificant and close to zero: the coefficient implies that an additional woman leader in the median district lowers age at birth by 0.047 years relative to a mean of 22.9 years (s.d. 5.1 years). Also, that our final specification is conditional upon mother fixed effects makes it unlikely that there are omitted variables correlated with age at birth that may also be correlated with the variable of interest (the share of seats in the district won by a woman).

² Note that year-specific shocks are captured in the estimated equations by birth cohort dummies. Any remaining concerns that our data contain a trend in the under-reporting of neonatal and infant deaths is addressed by the fact that we include flexible district-level controls for underlying trends (mother's cohort*district fixed effects).