Politician Identity and Development – Evaluating the IV Strategy built on a Close Elections RD Design

By MITUL SURANA AND ERROL D'SOUZA*

November 25, 2018

Abstract

We re-examine the instrumental variable methodology being widely used to obtain causal estimates of the impact of politician identity on development outcomes. The instrument, which is built upon an electoral regression discontinuity design, is constructed by aggregating the outcomes of close elections. In this article, we re-evaluate the validity of the instrument and the conditions under which it is applicable. Using an example of the application of the instrument we show that the current practice in the literature of non-randomly setting the instrument to zero when it is not defined renders it inappropriate for the purpose of establishing causality.

JEL codes: C36, D72, O10, O20

---

* Surana: Department of Economics, Indian Institute of Management Ahmedabad (email: mituls@iima.ac.in); D'Souza: Department of Economics, Indian Institute of Management Ahmedabad (email: errol@iima.ac.in). We thank Maitreesh Ghatak, Pritha Dev, Ankur Sarin and Abhiman Das for comments and suggestions. All errors are our own.

I. Introduction

Does the identity of those who represent us matter for policies and socio-economic outcomes? One approach following Downs (1957) takes the identity of representatives to be irrelevant if they have incentives to include any electorally salient issues into their policy agendas. In these spatial theories of electoral behaviour parties choose policy positions to minimize the distance between themselves and voters. If some group has significant preferences that are not being represented, parties will be induced to secure that space so as to maximize vote share. In contrast others such as Besley and Coate (1997) argue that candidates' personal preferences matter and voters vote for the candidates closest to their policy views. If the expected payoff of being elected is greater than the cost of running for office individuals will stand for election. Once elected, the candidates will implement policies of their own choosing and not respond to voter preferences. In that case politician identity is relevant for policy making and outcomes that affect affiliated groups of citizens.

The empirical literature on politician identity has utilized the close elections regression discontinuity (RD) design for purposes of causal inference. The application of a close elections RD design is straightforward when the relevant outcome variable is also available at the same geographic level as the election variables, i.e. at the electoral constituency level. However, development outcomes are not readily available at the electoral constituency level.[1] Instead, they are typically available at the larger administrative district level. Therefore, recent work has employed a novel instrumentation strategy in conjunction with close elections RD design to identify causal effects of election results on development outcomes. In particular, this includes studies which estimate the impact of politician identity, such as gender (Clots-

[1] Some works have used night lights data, which is available at the constituency level, as a measure of economic development to circumvent this problem. See for example Prakash et al. (2015) and Baskaran et al. (2018).

Figueras, 2011, 2012; Bhalotra and Clots-Figueras, 2014; Lee, 2018), party affiliation (Nellis, Weaver, and Rosenzweig, 2016; Nellis and Siddiqui, 2018), education (Lahoti and Sahoo, 2017), religion (Bhalotra, et al., 2014; Bhalotra, et al., 2018), and ruling party/coalition affiliation (Tandon, 2015; Bhavnani and Jensenius, 2016; Beg, forthcoming), on a range of development outcomes.

As development outcomes are available at the level of the administrative district, which typically contains multiple electoral constituencies, these studies aggregate the outcomes of elections to the administrative district level to generate their independent variables.[2] Further, to identify causal effects they use aggregated measures of outcomes in close elections as instruments for these variables. In this article we evaluate the validity of the instrument thus constructed and the conditions under which it is applicable. We show that the instrument is valid only for those districts in which there was at least one close election. Replacing the instrument to zero where it is missing, as has been a practice in the aforementioned studies, violates the instrument exogeneity assumption and thereby prevents any causal inference. This article, therefore, questions the causal interpretation of the estimates which rely on an incorrect application of the instrument.

The rest of this article is organised as follows. Section II describes the instrumentation methodology in question using one example of its application. The example that we choose for illustrating the method is the one studied by Bhalotra and Clots-Figueras (2014) where they estimate if the gender identity of politicians influences health outcomes in India. The choice of the example for illustration is based on the fact that the data and code for this paper are readily

[2] Bhavnani and Jensenius (2016) use data on development outcomes at electoral constituency level. While they are not aggregating at the level of geography, they do so over time.

available.[3] Section III evaluates the validity of the instrument and highlights the problems in the way the instrument is employed. It then re-examines the internal validity of the key findings of Bhalotra and Clots-Figueras (2014) in light of the issues raised with respect to the empirical foundations of the instrumentation strategy. Section IV evaluates the results of some of the other works that have used a similar instrumentation strategy. Section V concludes.

II. Description of the Instrumentation Strategy

Suppose we want to study the impact of the election of female politicians to the state legislatures on health outcomes in India. The main challenge in obtaining an unbiased causal estimate of the effect of female legislators on health is that there could be other unobservable omitted variables, such as voter preferences, which could drive both health outcomes and the election of female candidates.

One way to address this challenge is to use a close elections RD design. The design has been widely used in political applications in recent years (Lee, 2008; Pettersson-Lidbom, 2008; Ferreira and Gyourko, 2009). An electoral RD design takes advantage of the fact that the assignment of a candidate to the treatment (winning the election) or control (losing the election) is near randomly determined in very close elections, i.e. where the margin of victory is close to zero. The RD design utilizes the discontinuity in treatment assignment at the threshold of zero vote margin and its validity is based on the assumption that other unobservable characteristics do not vary discontinuously around this threshold. While there have been concerns regarding whether the outcome of close US House elections can be considered as quasi-random (Snyder, 2005; Grimmer et al., 2011; Caughey and Sekhon, 2011), studies using data on Indian elections

[3] Rehavi (2007) was the first to use such a design to estimate the causal effect of female state legislators on state spending in the US. Clots-Figueras (2011) was the first to use this design in Indian context. As the dataset they use are not readily available, we prefer to use Bhalotra and Clot-Figueras (2014) to illustrate our arguments.

are supportive of the claim that the outcome of elections where the margin of victory is less than 5% can be considered to be randomly determined (Uppal, 2009; Eggers et al., 2015).

Therefore, to address the identification problem, one can take advantage of the fact that outcomes of very close elections between male and female candidates can be considered to be quasi-random. However, the conventional close election RD based set up cannot be directly applied in this case. This is because most of the household surveys in India, like the National Sample Survey (NSS), National Family Health Survey (NFHS), District Level Household Survey (DLHS), India Human Development Survey (IHDS), etc., do not provide details of individual residence at the electoral constituency level. Such information is only available up to the larger administrative district level. As a district contains multiple electoral constituencies, the political representation data has to be aggregated to the district level to match it to the health data.[4] Therefore, to estimate the causal effect of female legislators on health outcomes, Bhalotra and Clots-Figueras (2014) (BC henceforth) employ an instrumental variables strategy that is built upon the electoral RD design.

District-level variables for female political representation are constructed by aggregating the constituency level data. Consequently, the measure for female political representation is the fraction of constituencies in the district won by female candidates. As this variable is endogenous, they instrument it with the fraction of constituencies in the district won by female candidates in close elections against male candidates.[5] Since the outcome of each close election between a female and male candidate in a district is assumed to be random, the average of those election outcomes might also assumed to be random. However, this may not necessarily be the case. For instance, it is possible that some districts have more female candidates involved in

[4] The number of electoral constituencies in a district varies. The median district has around 9 constituencies.

[5] They define close elections as those in which "the winner and the runner-up are of opposite gender and the vote margin between them is less than 3.5%".

close races against male candidates and this may be correlated with voter preferences. In other words, the presence of close elections involving female candidates in a district is not random. Therefore, to satisfy the exclusion restriction, BC control for the fraction of constituencies in a district that had close elections between female and male candidates in both the first and second stage equations.

To summarize, conditional upon the fraction of seats in a district that are *contested* by female candidates in close elections ($TotalClose_{dt}$), the fraction of seats in the district *won* by female candidates in close elections ($FemaleClose_{dt}$) is assumed to be exogenous and hence considered a valid instrument for the fraction of seats in a district won by female candidates ($Female_{dt}$). Thus, the model that is estimated is as follows:

$$Y_{idt} = \alpha \ + \beta Female_{dt} + \lambda TotalClose_{dt} + \epsilon_{idt} \qquad (1)$$

$$Female_{dt} = \theta + \kappa FemaleClose_{dt} + \mu TotalClose_{dt} + u_{idt} \qquad (2)$$

where Y_{idt} is the neonatal survival outcome for individual child *i* born in district *d* in year *t*.[6] Equation 2 is the first stage and equation 1 is the second stage from two stage least squares estimation.

To get a consistent estimate of β, the instrument must satisfy the following exogeneity assumption:

$$E(\epsilon|FemaleClose, TotalClose) = E(\epsilon|TotalClose) \qquad (3)$$

III. Problems with the Application of the Instrument

The identifying assumption of the instrumentation strategy is based on the fact that

[6] Their estimations also include mother and cohort fixed effects, individual and district specific controls, state-specific linear trends, and third degree polynomial controls for victory margin for every male-female election in the district. Since our focus here is on the instrument, we do not show these in the specification. However, we include these in our estimations when we examine their results.

conditional upon being in a close election, the outcome of such a contest is random, i.e. $FemaleClose$ is randomly determined only when $TotalClose \neq 0$. Thus,

$$E(\epsilon|FemaleClose, TotalClose \neq 0) = E(\epsilon|TotalClose \neq 0)$$

However, when there are no close elections in a district, i.e. when $TotalClose = 0$, the instrument is not defined. In such cases, BC set the instrumental variable in the estimation sample to zero ($FemaleClose = 0$). Setting the instrument to zero (or any other constant) when there are no close elections cannot be considered a random outcome unless the absence of close elections itself is randomly determined, i.e.,

$$if\ we\ set\ FemaleClose = 0, when\ TotalClose = 0$$

$$then, E(\epsilon|FemaleClose, TotalClose) = E(\epsilon|TotalClose)$$

$$only\ if, E(\epsilon|TotalClose = 0) = E(\epsilon)$$

But, it is not reasonable to assume that the absence of close elections between male and female candidates is randomly determined. This is because districts that have no close elections could be different from districts that have close elections along unobservable factors such as voter preferences. In fact, the very reason that BC provide for controlling for the fraction of seats that had close elections is that "although the gender of the winner in a close election may be considered random, the existence of close elections between men and women may not".

To give an example as to why setting the instrument to zero is misguided, let us consider, for instance, that voters in certain districts do not care much about health outcomes and have a very strong preference for male candidates such that they do not want a female even in a close contest with a male. In such circumstances, it is voter preferences that determine that $Female$ and $TotalClose$ are both zero and neonatal mortality is high. Therefore, if we set the instrument, $FemaleClose$, to zero (or any other constant), it would be correlated with the error term, thus violating the exogeneity assumption.

A more accurate application of the instrumentation strategy will need to take into account that the instrument is valid only for those districts where there exists at least one constituency where a female politician contested in a close election against a male politician. Hence, any causal inference with respect to the effect of politician gender under this empirical strategy can only be made based on the sub-sample of such districts where there exist a close election. It is important to note that an estimate obtained from such a sub-sample will limit our ability to make inferences about districts which didn't have close elections. But this is the limitation of all RD based designs where the internal validity of the estimate is achieved at the cost of its external validity.

BC's electoral dataset consists of elections during the period 1967-1998 in 16 major states of India across 246 districts with a total of 1,826 district-election year cells in the estimation sample (Panel A of Table 1).[7] The proportion of districts that had at least one close election between a male and a female candidate is 0.0745. In other words, 92.55% of the district-election years did not experience any close election between a male and a female candidate.

[Insert Table 1 here]

Panel B shows that out of the 1,826 district-election year cells in BC's estimation sample, 1,690 have no close elections. Within these 1,690 district-election years which do not experience any close election, 1,300 (77%) do not elect any female legislator. It is highly likely that unobservable factors are determining both absence of close elections as well as absence of female legislators and therefore using an instrument which is a deterministic function of the former for these observations would make the instrument endogenous.

Now, we re-examine BC's key findings. We first reproduce their results from their preferred specification which includes cohort and mother fixed effects, individual and district specific controls, state-specific linear trends and third order polynomials in victory margin (column 1

[7] Their dataset and code are available at: https://www.aeaweb.org/articles?id=10.1257/pol.6.2.164

of Table 2). They find that a 1 percentage point increase in the share of female legislators in the district 'causes' a statistically significant reduction in neonatal mortality by 0.21 percentage points. However, this estimate is based on the sample in which the instrument is not defined for around 93% of the observations and hence replaced by zero. Therefore, as discussed above, one cannot infer causality from this estimate.

[Insert Table 2 here]

BC also perform a robustness exercise in which they restrict the sample to those district-year observations "in which there was at least one election between a man and a woman as these are the observations upon which the identification rests" and find that their estimate is similar (column 2 of Table 2). However, it is important to note that the identification is not based on those district-year observations in which there was at least one male-female election but rather based on those in which there was at least one '*close'* male-female election.

In column 3 of Table 2, we restrict the sample to district-years where there was at least one close election between a female and a male candidate. The sample size falls considerably, but then it is only for this sub-sample that the instrument is valid. We find that BC's results do not hold. The sign of the coefficient reverses, though it is not significant.

Table 3 shows the means of the dependent variable and the main independent variable for district-years with and without close elections. We see that neonatal mortality is much higher and female representation lower in districts with no close elections and this is what we believe is driving BC's results.

[Insert Table 3 here]

The given data does not suggest that female legislators have a significant 'causal' effect on neonatal mortality. While we reckon that there might not be sufficient power for us to claim that we fail to reject the null hypothesis that the gender of legislators has no effect on neonatal mortality, but our findings, at the very least, suggest that there isn't sufficient evidence to claim that there is a causal effect either. More importantly, our findings highlight that when the instrument is not defined for a subset of the sample, the manner in which the full sample instrument is constructed has important consequences for the estimates that are obtained and whether causal interpretation can be attributed to those estimates.

IV. Widespread Utilization of the Instrument

We have identified several other works that use a similar empirical strategy to examine the causal impact of politician identity (see Appendix Table A1).[8] 13 out of the 14 works that we have identified have set the instrument to zero when it is undefined. Some papers report the robustness of their estimates to the removal of such observations where the instrument is not defined. Clots-Figueras (2012) in her analysis of the impact of female leaders on educational outcomes in India uses an instrumental variable design based on close elections between male and female candidates as described above. Based on the full sample (where the instrument is set to zero when it is not defined), she finds that female state legislators have a significant positive effect on the probability that an individual obtains primary education in urban areas. The paper also reports the estimates for the restricted sample of district-election years with at least one close male-female election, but, the results are not statistically significant for this sub-sample (see the fifth set of results in Table 3 in Clots-Figueras, 2012).

Bhalotra, Clots-Figueras, and Iyer (2018) study the impact of Muslim legislators on sex selective abortion in India by using the fraction of constituencies in a district won by Muslim

[8] Though a majority of the papers that we have identified have used data from India, this instrumentation strategy has been used in different contexts. Rehavi (2007) was the first to use a similar design to estimate the causal effect of female state legislators on state spending in the US. Recent literature on the political economy of development in Pakistan has also used this instrumentation strategy to estimate the impact of secular-party incumbency on religious conflict (Nellis and Siddiqui, 2018) and to estimate the impact of ruling party affiliation on allocation of river waters (Beg, forthcoming).

politicians in close elections against non-Muslim politicians as an instrument for Muslim representation. They find that the election of Muslim legislators increases the probability of a girl birth. As part of the robustness tests, they also estimate the effect after restricting the sample to those district-election years where there was at least one close election between Muslim and non-Muslim candidates, but, the estimate is no longer statistically significant (see Table 2 column 5 in Bhalotra, Clots-Figueras, and Iyer, 2018).

While these papers have reported the estimates after excluding the observations where the instrument is not defined, it is important to note that the estimates do not remain significant. Further, even if the estimates obtained from such a robustness exercise were significant, this would not be a validation for using, for the main IV analysis, a sample where the instrument exogeneity is violated.

To demonstrate why it is important to base the main analysis on the sample where the instrument is valid, we re-examine the findings of Nellis, Weaver, and Rosenzweig (2016), who estimate the impact of the election of legislators from Congress Party on Hindu-Muslim riots in India. They use close elections between Congress party and non-Congress party candidates to account for the endogeneity of election outcomes. A standard regression discontinuity design could not be used as the information on riots is not available at the electoral constituency level. Therefore, the aggregate measure of Congress party representation at the district level, i.e. the fraction of constituencies in the district won by Congress candidates ($Congress_{dt}$), is instrumented with the fraction of constituencies in the district won by Congress candidates in close elections against non-Congress candidates ($CongressClose_{dt}$). The fraction of constituencies that had close elections between Congress and non-Congress candidates ($CongressTotalClose_{dt}$) is controlled for in both stages. Thus, the model that is estimated is:

$$Y_{dt} = \alpha + \beta Congress_{dt} + \gamma CongressTotalClose_{dt} + \epsilon_{dt}$$

$$Congress_{dt} = \mu + \lambda CongressClose_{dt} + \kappa CongressTotalClose_{dt} + v_{dt}$$

where Y_{dt} is a binary variable which takes value 1 if any riots occurred in district d during election cycle t.[9] The margin of victory used to define close elections is 1% and the instrument is set to zero when it is not defined.[10]

The result of their main specification using the full sample is shown in column 1 of Panel A in Table 4. They find that an increase in Congress seat share in a district from 0 to 1 results in a 42.3 percentage point decrease in the probability of occurrence of any riot in that district. Their result is robust to the inclusion of district fixed effects which control for the permanent differences across districts in Congress party representation and the occurrence of riots (column 2 of Panel A). Though the magnitude of the effect drops but it is still statistically significant. Their estimate is also robust to the use of 2% victory margin as the definition of close elections (column 3 of Panel A).

[Insert Table 4 here]

In panel B of Table 4, we present the results from our estimations of the same specifications as in panel A, but with the sample now restricted to only those district-election cycles where there was at least one close election. We find that the effect is still significant in the baseline specification, but it fails to hold when we include district fixed effects or when we use 2% victory margin to define close elections.[11]

---

[9] In addition to the binary indicator for the occurrence of riots, they use three other dependent variables: logged number of riots, logged number of riot casualties and logged number of days of rioting. Our conclusions remain similar when we use any of these other measures.

[10] Their dataset and code are available at: http://dx.doi.org/10.1561/100.00015051

[11] Nellis, Weaver and Rosenzweig (2016) also implement a test in which they first restrict the sample to district-election cycles where there was exactly one close election between a Congress and a non-Congress candidate and then use a t-test for difference in means to estimate the effect of Congress candidates winning as compared to them losing those elections. They find a significant effect of victory of Congress candidates on the probability of occurrence of

To summarize, when we use the full sample for the main analysis (Panel A) and use the restricted sample only as a robustness exercise (column 1 of Panel B), we might conclude that Congress party legislators cause a decrease in the probability of Hindu-Muslim riots. However, when we use the sample where the instrument is valid for our main analysis (column 1 of Panel B) and perform further robustness tests on this sample (columns 2 and 3 of Panel B), then we fail to find support for a robust causal relationship between Congress party legislators and Hindu-Muslim riots.

Therefore, we believe that for the purpose of robust causal inference it is imperative that the main results and any robustness tests on those results should only be based on the sample where the instrument exogeneity holds.

V. Conclusion

In this article, we highlight the problems in the implementation of an instrumental variables methodology that is being used to estimate the causal impact of politician identity on development outcomes. We show that the instrument, which is constructed by aggregating the outcomes of close elections in a district, is defined only when there exist at least one constituency in the district that had a close election. Given that the absence of close elections in a district cannot be considered random, we challenge the current practice of setting the instrument to zero when it is undefined as it makes the instrument endogenous and therefore inappropriate for the purpose of establishing causality.

Using an example of the application of this empirical strategy by Bhalotra and Clots-Figueras (2014) to estimate the effect of gender identity of state legislators on health outcomes,

---

riots (column 1 of Panel C in Table 4). However, as before, our estimations show that the estimate is not significant when we add district fixed effects or use 2% victory margin to define close elections (see columns 2 and 3 of Panel C in Table 4).

we show that such a non-random assignment of the instrument to zero is not without consequences. We find that their result suggesting a causal effect of female legislators on neonatal mortality does not hold once we take into account that the instrument is valid only for a small subset of the sample.

Finally, we evaluate the results of some of the other works that have used a similar instrumentation strategy and find that the estimates do not always continue to remain significant when the sample is restricted to only those observations where the instrument is valid. Further, employing the dataset used by Nellis, Weaver and Rozensweig (2016) to study the effect of Congress party incumbency on Hindu-Muslim riots in India, we show that using the sub-sample where the instrument is valid merely as a robustness exercise while using the full sample with an invalid instrument for the main estimation might lead to incorrect causal inferences. While we cannot comment on whether the results of the other works that have used such an instrumentation strategy would be any different once the estimation sample is changed, but what we can argue is that it is premature to infer causality from instrumental variable estimates which, for a substantial part of the sample, rely on a non-randomly determined instrument.

References

Baskaran, Thushyanthan, Sonia Bhalotra, Brian Min, and Yogesh Uppal. 2018. "Women Legislators and Economic Performance". WIDER Working Paper 2018/47.

Beg, Sabrin. Forthcoming. "Favoritism and Flooding: Political Alignment and Allocation of Irrigation Water". *World Development*.

Besley, Timothy, and Stephen Coate. 1997. "An Economic Model of Representative Democracy". *Quarterly Journal of Economics* 112 (1): 85–114.

Bhalotra, Sonia, and Irma Clots-Figueras. 2014. "Health and the political agency of women". *American Economic Journal: Economic Policy* 6(2): 164-197.

Bhalotra, Sonia, Irma Clots-Figueras, Guilhem Cassan, and Lakshmi Iyer. 2014. "Religion, politician identity and development outcomes: evidence from India". *Journal of Economic Behavior & Organization* 104: 4-17.

Bhalotra, Sonia, Irma Clots-Figueras, and Lakshmi Iyer. 2018. "Religion and Abortion: The Role of Politician Identity". IZA Discussion Paper No. 11292.

Bhavnani, Rikhil R., and Francesca Jensenius. 2016. "Voting for development? ruling coalitions and literacy in India". https://faculty.polisci.wisc.edu/bhavnani/wp-content/uploads/2016/02/Bhavnani_Jensenius_2016.pdf.

Caughey, Devin, and Jasjeet S. Sekhon. 2011. "Elections and the Regression Discontinuity Design: Lessons from Close US House Races, 1942–2008". *Political Analysis* 19(4): 385-408.

Clots-Figueras, Irma. 2011. "Women in politics: Evidence from the Indian States". *Journal of Public Economics* 95(7-8): 664-690.

Clots-Figueras, Irma. 2012. "Are Female Leaders Good for Education? Evidence from India". *American Economic Journal: Applied Economics* 4(1): 212-244.

Downs, Anthony. 1957. "An Economic Theory of Democracy". New York: Harper and Row.

Eggers, Andrew C., Anthony Fowler, Jens Hainmueller, Andrew B. Hall, and James M. Snyder Jr. 2015. "On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races". *American Journal of Political Science* 59(1): 259-274.

Ferreira, Fernando, and Joseph Gyourko. 2009. "Do political parties matter? Evidence from US cities". *Quarterly Journal of Economics* 124(1): 399-422.

Grimmer, Justin, Eitan Hirsh, Brian Feinstein, and Daniel Carpenter. 2011. "Are Close Elections Random?". https://web.stanford.edu/~jgrimmer/CEF.pdf.

Lahoti, Rahul, and Soham Sahoo. 2017. "Are Educated Leaders Good for Education? Evidence from India". https://ssrn.com/abstract=2851895.

Lee, David S. 2008. "Randomized experiments from non-random selection in US House elections". *Journal of Econometrics* 142(2): 675-697.

Lee, YuJung Julia. 2018. "Gender, Electoral Competition, and Sanitation in India". *Comparative Politics* 50(4): 587-605.

Nellis, Gareth, Michael Weaver, and Steven Rosenzweig. 2016. "Do Parties Matter for Ethnic Violence? Evidence from India". *Quarterly Journal of Political Science* 11(3): 249-277.

Nellis, Gareth, and Niloufer Siddiqui. 2018. "Secular party rule and religious violence in Pakistan". *American Political Science Review* 112(1): 49-67.

Pettersson-Lidbom, Per. 2008. "Do parties matter for economic outcomes? A regression-discontinuity approach". *Journal of the European Economic Association* 6(5): 1037-1056.

Prakash, Nishith, Marc Rockmore, and Yogesh Uppal. 2015. "Do criminally accused politicians affect economic outcomes? Evidence from India". *Households in Conflict Network (HiCN), The Institute of Development Studies, University of Sussex.*

Rehavi, Marit. 2007. "Sex and Politics: Do Female Legislators Affect State Spending?". Unpublished.

Snyder, Jason. 2005. "Detecting Manipulation in U.S. House Elections". Unpublished.

Tandon, Sharad. 2015. "Taxation and Political Mobilization: Evidence from Andhra Pradesh". *Economic Development and Cultural Change* 63(3): 515-549.

Tandon, Sharad. 2018. "Election Outcomes and the Poor: Evidence from the Consumption of Scheduled Castes and Tribes in India". *Economic Development and Cultural Change* 66(2): 331-382.

Uppal, Yogesh. 2009. "The disadvantaged incumbents: estimating incumbency effects in Indian state legislatures". *Public Choice* 138(1/2): 9-27.

Table 1: Electoral Dataset of Bhalotra and Clots-Figueras (2014)

Panel A : Descriptive Statistics		
Unit of observation: District in an election year, elections between 1967-1998 in 246 districts	Mean	SD
Proportion of seats won by females ($Female$)	0.0364	0.0740
Proportion of seats that had close election (3.5% margin) between females and males ($TotalClose$)	0.0084	0.0333
Proportion of seats won by females in close elections (3.5% margin) against males ($FemaleClose$)	0.0041	0.0227
Proportion of districts that had at least one close election (3.5% margin) between a female and a male ($TotalClose \neq 0$)	0.0745	0.2626
Observations	1,826	

Panel B : Breakup of the district-election year cells

	No close election ($TotalClose = 0$)	At least one close election ($TotalClose \neq 0$)	Total
No female legislator ($Female = 0$)	1,300	44	1,344
At least one female legislator ($Female \neq 0$)	390	92	482
Total	1,690	136	1,826

Notes: Panel A of the table presents the descriptive statistics of the electoral dataset of BC. These are the same as reported in Table 1B of their paper. Here, we present only the key variables that are of interest. In Panel B we show the breakup of the district-year observations by presence and absence of close elections and by presence and absence of female legislators.

Table 2: Female legislators and neonatal mortality – IV estimates

Dependent Variable: neonatal mortality	2SLS (1)	2SLS (2)	2SLS (3)
Fraction of seats in the district won by females ($Female$)	- 0.2061*** (0.078)	- 0.2289** (0.091)	+ 0.310 (0.22)
Observations	71,498	29,979	5,181
Number of mothers	18,754	9,989	2,105
Estimation Sample	Full	At least one female-male election	At least one *close* female-male election
First-stage regressions. Dependent variable: Fraction of seats in the district won by females ($Female$)			
Fraction of seats in the district won by females in close elections against males ($FemaleClose$)	0.8964*** (0.0789)	0.8900*** (0.0966)	0.8702*** (0.0429)
Fraction of seats in the district contested by females in close elections against males ($TotalClose$)	- 0.4495*** (0.0559)	- 0.4698*** (0.0692)	- 0.4192*** (0.0639)
First-stage F-statistic	129.1	84.9	411.3

Notes: The table presents the IV estimates for the effect of female political representation on neonatal mortality in India using BC's dataset. Robust standard errors clustered at the district level are reported in parentheses. Close elections are defined as those in which winner won against the runner-up by less than 3.5 percent of votes. This presents results from BC's preferred specification which includes mother and cohort fixed effects, state specific linear trends, district and individual level controls, and a control for the fraction of constituencies in the district that had close elections between male and female candidates ($TotalClose$). Column 1 estimates the model using the full estimation sample with the instrument set to zero where it is not defined (this is the same result as presented in BC's Table 2, column 6 of Panel A). Column 2 estimates the model on a sub-sample of those district-election years where there was at least one female-male election (this is the same result as presented in BC's Table 3, column 8 of Panel A). Column 3 estimates the model on the sub-sample of those district-election years where the instrument is defined, i.e. where there was at least one *close* female-male election.
** Significance at 5 percent level. *** Significance at 1 percent level.

Table 3: Female political representation and neonatal mortality for district-years with and without close elections

	At least one close election ($TotalClose \neq 0$)		No close election ($TotalClose = 0$)	
	Mean	SD	Mean	SD
Neonatal Mortality	0.0608	0.2390	0.0642	0.2452
Proportion of seats won by females ($Female$)	0.0964	0.0830	0.0333	0.0673
Observations	7,121		64,377	

Notes: This table presents the means of the dependent and the main independent variable of BC's model for district-election years with and without close elections.

Table 4: Congress party representation and the occurrence of Hindu-Muslim riots

Dependent variable: riot (1 = riot occurred, 0 = otherwise)			
Panel A: Full Sample	2SLS Baseline (1)	2SLS District FE (2)	2SLS 2% margin (3)
Fraction of seats in the district won by Congress ($Congress$)	-0.423***	-0.273**	-0.207**
	(0.152)	(0.124)	(0.091)
Observations	2,871	2,871	2,871
First-stage F-statistic	58.8	48.4	157.9

Panel B: Restricted Sample (at least one close election)	2SLS Baseline (1)	2SLS District FE (2)	2SLS 2% margin (3)
Fraction of seats in the district won by Congress ($Congress$)	-0.355**	-0.230	-0.127
	(0.154)	(0.159)	(0.092)
Observations	838	838	1,382
First-stage F-statistic	58.9	35.1	166.8

Panel C: Exactly One Close Election	Baseline (1)	District FE (2)	2% margin (3)
Congress candidate won the election	-0.076**	-0.042	-0.012
	(0.032)	(0.048)	(0.027)
Observations	644	644	831

Notes: The estimation sample is the one used by Nellis, Weaver, and Rosenzweig (2016). The first two panels of this table present IV estimates for the effect of Congress party representation on a binary riot outcome. Robust standard errors clustered at the district level are reported in parentheses. All regressions in Panel A and B include a control for the fraction of constituencies in the district that had close elections between Congress and non-Congress candidates. Panel A presents IV estimates with the full estimation sample, i.e. with the instrument set to zero where it is not defined. Panel B presents IV estimates with a sub-sample of those district-election years where there was at least one close election between a Congress and a non-Congress candidate. In panel C the sample is first restricted to those district-election years where there was exactly one close election between a Congress and a non-Congress candidate. Then the binary riot outcome is regressed on a binary variable that is equal to 1 if the Congress candidate won that single close election. Column 1 shows the baseline estimates with the margin of victory used to define close elections as 1%. Columns 2 and 3 show robustness checks on the specification in column 1. Column 2 includes district fixed effects. Column 3 changes the margin of victory used to define close elections to 2%.

** Significance at 5 percent level. *** Significance at 1 percent level.

Appendix Table A1: List of works that have used the close elections IV strategy to examine the impact of politician identity on development outcomes

Author (year)	Type of work	Research Question	Main Estimation Sample (Full/Restricted)#	Remarks
Clots-Figueras (2011)	Journal of Public Economics	Effect of gender of elected legislators on public goods, policy and expenditure	Full	
Clots-Figueras (2012)	American Economic Journal: Applied Economics	Effect of gender of elected legislators on educational outcomes	Full	Robustness for restricted sample with at least one close election is reported in the paper. The estimates are not significant for this sub-sample.
Bhalotra and Clots-Figueras (2014)	American Economic Journal: Economic Policy	Effect of gender of elected legislators on health outcomes	Full	Our replication with restricted sample shows that the results do not hold.
Bhalotra, Clots-Figueras, Cassan, and Iyer (2014)	Journal of Economic Behavior & Organization	Effect of religious identity of elected legislators on health and educational outcomes	Full	
Tandon (2015)	Economic Development and Cultural Change	Effect of ruling party affiliation of elected legislators on collection of tax revenue in the state of Andhra Pradesh in India	Full	Robustness for restricted sample with at least one close election reported in the paper. The results hold.
Nellis, Weaver, and Rosenzweig (2016)	Quarterly Journal of Political Science	Effect of party affiliation of elected legislators on Hindu-Muslim riots	Full	Our replications with restricted sample show that the results are not robust to the inclusion of district fixed effects or to the change in margin used to define close elections.

Nellis and Siddiqui (2018)	American Political Science Review	Effect of secular-party incumbency on religious conflict in Pakistan	Full	
Tandon (2018)	Economic Development and Cultural Change	Effect of party affiliation of members of parliament on consumption of SCST households in India	Restricted	
Lee (2018)	Comparative Politics	Effect of gender of elected leaders on access to sanitation	Full	
Beg (forthcoming)	World Development	Effect of ruling party affiliation on allocation of river waters in Pakistan	Full	
Rehavi (2007)	Unpublished	Effect of gender of state legislators on the distribution of state spending in US	Full	
Bhavnani and Jensenius (2016)	Working Paper	Effect of ruling party affiliation of legislators on literacy	Full	
Lahoti and Sahoo (2017)	Working Paper	Effect of educated leaders on educational outcomes	Full	
Bhalotra, Clots-Figueras, and Iyer (2018)	IZA Working Paper	Effect of religious identity of elected legislators on sex-selective abortion	Full	Robustness for restricted sample with at least one close election is reported in the paper. The estimates are not significant for this sub-sample.

[#] Full - full sample with instrument set to zero where not defined; Restricted - sample restricted to those observations where the instrument is defined