EasyChair Preprint

№ 1112

Representative Bureaucracy in India: An empirical study of gender representation by K-12 teachers

Anita Dhillon and Kenneth J. Meier

June 8, 2019

Representative Bureaucracy in India?
An empirical study of gender representation by K-12 teachers[1]

Anita Dhillon
American University
ad3679a@student.american.edu

Kenneth J. Meier
American University
kmeier@american.edu

Abstract

A bureaucracy representative of historically oppressed groups in society has been linked to better outcomes for those groups, especially in key policy areas such as education, child welfare, and law enforcement. Most of the empirical work in education has used the case of the United States, finding that representation leads to academic improvements for the represented group. Would the same effects be seen in national contexts beyond the Western World? This paper examines how and when a female bureaucrat in the Indian education system leads to enhanced outcomes for female students, using data gathered by the Government of India on nearly 1.5 million schools, from 2014-15 to 2017-18. We find a modest representative effect of female teachers on the academic achievement of female students which is enhanced with more class days, longer teacher hours, and a smaller student teacher ratio. Additionally, in situations where active representation would be extremely difficult due to a severe lack of resources and opportunities (schools located in rural areas and with poor infrastructure) we also see a positive effect of female teachers on female student academic achievement, which possibly indicates the presence of symbolic representation via the role-model effect, in these environmental conditions.

[1] Draft Paper: to be presented at the 2019 Public Management Research Conference at the University of North Carolina, Chapel Hill

Introduction

Bureaucracies that are representative of minority groups in a society may positively affect policy outcomes for those groups, especially in areas such as education, child protection, and law enforcement (Favero and Molina 2018; J. A. Grissom, Kern, and Rodriguez 2015). While initial representative bureaucracy studies focused on race and ethnicity as the salient identity of representation, more recent work has found similar positive effects of a gender-representative bureaucracy in certain circumstances (Keiser et al. 2002). Parallel research also identifies conditions when a sharing of identities between bureaucrat and client leads to more positive outcomes for the client, including the salience of that shared identity and the level of discretion possessed by the bureaucrat, among others (Keiser et al. 2002; Meier 1975; Meier and Nigro 1976; Riccucci, Van Ryzin, and Jackson 2018; Theobald and Haider-Markel 2009; Wilkins and Keiser 2006).

Numerous empirical studies on representative bureaucracy exist, covering a variety of salient demographic characteristics, different levels of bureaucracy, and different types of governmental agencies (see Kennedy 2014). Most of the empirical work, however, studies public organizations and bureaucracies in the US and Western Europe. This severely limits how generalizable the representative bureaucracy theory is to national contexts beyond the 'western' world (but see Agyapong 2017; Song 2018; Zhang 2018). Additionally, this narrow contextual focus also means that the full range of organizational and environmental conditions that need to exist, so that representation in the bureaucracy leads to more responsive policy outcomes, has not been adequately explored and identified. Some of these theoretically important variables for bureaucratic action could include access to basic infrastructure, resources, organizational training, and time for the bureaucrats to adequately perform the critical aspects of their role as well as how embedded the cultural and societal norms are that exacerbate the oppression for the salient minority group in a particular society.

This paper adds to a nascent but growing body of research in representative bureaucracy that tests the theory in political and national contexts beyond that of the US or Western Europe. It does so by studying passive gender representation among street level bureaucrats and its effects on policy outcomes in India. Specifically, we seek to discern the basic internal and external environmental conditions needed (beyond values, attitudes or beliefs) for representing bureaucrats to have a positive effect on the needs of their represented group (either actively or symbolically). For ease of comparison to existing research, we study the Indian K-12 educational system, using multiple years of data gathered by the Government of India on nearly 1.5 million schools across the county, and analyze how and when could a female teacher in the Indian education system enhance the academic outcomes of female students.

The paper begins with a brief discussion of the theory of representative bureaucracy and its evolution, especially in the context of gender representation. This is followed by a summary of the Indian bureaucratic context, including the salience of gender as an identity in the country, its K-12 educational system, and the various internal and external environmental conditions that could possibly affect bureaucratic representation. We then introduce a theoretical framework that lays out our hypotheses on how these conditions would affect the outcomes of the represented group. After discuss the data and methodology of our analysis, we review the findings and identify various organizational factors that either enhance or dampen the policy effects of gender representation in a K-12 educational setting in India, ending with further research required to adequately test the conditions and contexts under which the representative bureaucracy theory will hold in practice.

Review of the representative bureaucracy literature

The term "Representative Bureaucracy" was introduced by Donald J. Kingsley in 1944, but Mosher's (1982) analysis informs its current interpretation: a bureaucrat's origins

and background shapes their values and actions, which in turn influences their bureaucratic decision-making. Mosher further added to the theory by differentiating between two types of representation. He defined passive representation to mean the degree to which the individuals in a bureaucracy mirror the entire population they serve and active representation to be when individual bureaucrats advocate for the interests of the population they represent, potentially leading to improved outcomes for this represented group (Mosher 1982, 14-15). Past research on representative bureaucracy traditionally falls under three broad thematic areas: identifying the extent (and determinants) of passive representation in different agencies, identifying the link between passive representation and the potential for active representation via values or attitudes, and studying the relationship between passive and active representation in terms of policy outcomes (Selden 1997). A more recent literature examines symbolic representation: the case where passive representation changes the attitudes or behaviors of the individuals who are represented (Riccucci, Van Ryzin and Levena 2014). There is a growing interest, however, in moving beyond merely identifying a relationship between passive and active representation to better understand why and under what conditions does passive representation lead to improved outcomes for the represented group (Meier 2019).

Multiple "conditions" have been posited for passive representation to turn into active representation. These include a shared value set between the bureaucrat and the group they represent (Meier 1975; Meier and Nigro 1976), the amount of discretion possessed by bureaucrats (Keiser et al. 2002), the salience of the policy area to the shared demographic trait (Wilkins and Keiser 2006), and the ability of bureaucrats to influence the actions of their colleagues/organization. Studies also identify several possible causal pathways between passive representation and improved outcomes, beyond active representation. One such way is through symbolic representation, where the mere presence of a representative bureaucrat is enough for clients in the represented group to change their behavior, thereby influencing outcomes without any action on the part of the bureaucrat (Riccucci, Van Ryzin, and Jackson

2018; Theobald and Haider-Markel 2009). Symbolic representation may occur when a bureaucrat serves as an aspirational role model for the target group that identifies with them, altering the behavior of target group members (Atkins and Wilkins 2013) or providing a signal that the target group's interests will be given consideration (Vinopal 2017).

Education policy and administration can particularly benefit from the empirical analysis of representative bureaucracy since the interactions, between teachers as street-level bureaucrats and students as the beneficiary group, are an apt setting for the aforementioned causal pathways and conditions to be tested. First, various identities could be considered salient in the context of education, including gender, race, ethnicity and religion, with systemic disparities in outcomes noted along these identity lines. Second, discretionary authority is a key aspect of teaching and added regulation or supervision can only dampen it to a certain extent. Third, teachers can have a powerful "role model" effect just based on the lengthy interaction time they have with students. Finally, in certain contexts teachers can also influence curriculum and pedagogy changes as well as the actions of other teachers within the same school via formal and informal interactions.

A majority of the early empirical analysis of representative bureaucracy focused on race and ethnicity as the salient shared identity, with research on gender salience gaining momentum only in the 2000s (Keiser et al. 2002; Riccucci and Meyers 2004). Riccucci and Meyers (2004) find that earlier studies on linking passive to active representation for women produced mixed results but theorized that this could be because they did not consider whether the policy area was gendered, and the institutional contexts surrounding the bureaucracy. Having considered these factors, their study (and others) find active representation for women in a variety of contexts from K-12 education (Keiser et al. 2002; Song 2018) to child welfare (Riccucci and Meyers 2004; Wilkins and Keiser 2006) to law enforcement (Andrews and Johnston Miller 2013; Riccucci, Van Ryzin, and Lavena 2014) and even job counseling (Guul 2018).

Most of the empirical work in education has used the case of the United States and focused on studying representation with regard to race and ethnicity (Grissom, Nicholson-Crotty, and Nicholson-Crotty 2009; Morton 2015; Nicholson-Crotty et al. 2017; Pitts 2007; Roch and Edwards 2015), with a more recent body of work examining gender as the salient identity (Dee 2005; Keiser et al. 2002; Stearns et al. 2016). A majority of these studies find that passive representation leads to improved policy outcomes with respect to test scores, better access to gifted programs by the target group, less biased teacher perceptions, and less biased use of disciplinary policies. The strength of these positive effects, however, vary across identities and geographical contexts.

Evidence supporting the representative bureaucracy theory in the US context, and the related policy implications, make a strong case for the study of representation in education in other national contexts. Would the same effects be seen in societies with varying degrees of identity-based relations, bureaucratic influence, administrative structures, and socio-economic development? An emerging body of work is attempting to answer these questions by studying the effects of passive representation by bureaucrats in the educational context of non-western countries, specifically Ghana, Korea and China (Agyapong 2017; Song 2018; Zhang 2018). All these studies focus on gender as the salient identity and find that passive female representation improves the test scores of female students, with the role model effect (Agyapong 2017) and the amount of teacher discretion (Song 2018) influencing the strength of this improvement. The contexts of the three relevant countries, however, only cover some variations in bureaucratic influence, organizational structure, and political institution setup, namely the more centralized Asian state with a Confucian model of bureaucracy and a relatively prosperous and stable West African state. Further studies on representation in other national contexts are, therefore, required to test the generalizability of the theory.

Gender and Education in India

India is an important national context for the study of representative bureaucracy given the myriad salient identity cleavages that exist in the country (the caste system, gender, religion, and linguistic groups to name a few), its unique political and governance system (a functioning democracy with lawmaking powers shared between the central government and the 29 states), and its tenuous socio-economic situation (vast social disparities across states, identity groups, and rural/urban areas). The issue of representing various identities occupies a prominent place in Indian politics, administration, and society. It has its roots in the British colonial rule, playing a part in the country's formation and subsequent partition based on religious lines, and leading to the constitutional provision of reservations for persons belonging to the historically disadvantaged Hindu castes and tribal groups (van Gool and de Zwart 2013). While not as politically visible, gender identity is also particularly salient in India, given the wide disparities between the social and economic outcomes of men and women, which has earned the country a ranking of 127 out of 160 countries on UNDP's Gender Inequality Index in 2018 (United Nations Development Programme 2018). Given the lack of a strong legislative/constitutional provision to address the gender disparities (something that exists for the other identity groups discussed above), it is therefore appropriate to begin researching gender representation among Indian bureaucrats and the effects it can have on policy outcomes.

Gender disparities start early in life for the female child in India; the sex ratio of 933 females per thousand males in India as per the 2011 National Census (Office of the Registrar General & Census Commissioner 2011) contrasts with a ratio of 1021 females per thousand males in the US. Given how large and diverse the country is, it is no surprise that these disparities are nuanced by caste, geographical location, income and urbanization, among other factors. On the economic front, Jayachandran's (2015) analysis on cross-national gender inequality finds that men are three times more likely than women to be part of the

labor force in India, one of the largest disparities noted across countries. Biased attitudes are also common with women enjoying little freedom of choice or control over their life and parents exhibiting a strong preference for male children (Jayachandran 2015). Gender based-violence is another entrenched action, with the National Family Health Survey finding that nearly 2 in 5 women report having experienced spousal violence (this number is higher for women who are employed) and more than 50% of women agreeing with one or more reasons that justify wife beating (NFHS Report, 2006).

The gender disparities extend to educational outcomes too with women's literacy levels 16 percentage points below men's, according to the 2011 Census. These disparities exist even at the school education level, with girl students between the ages 8-11 underperforming boy students on reading and math achievement tests (White et al. 2016). Also, while enrollment of girls in primary school has improved, they are less likely to continue education post primary school (5^{th} grade) than boys (Ministry of Human Resource Development 2018). The redistributive role education can play and the tremendous positive externalities of girl children being educated, especially in terms of promoting further gender equality (Malhotra, Pande, and Grown 2003), makes education in India a gendered policy area and an ideal policy area for this study.

The issue of gender inequality is a highly salient one for the Indian Government too, especially in education. At a broader level, the desire to reform the education system led the central government to pass the Right to Free and Compulsory Education (RTE) Act in 2009. The RTE Act is the legislative manifestation of an amendment to the Indian Constitution that grants every child the fundamental right to free, compulsory, and full-time education between the ages 6-14. Additionally, it also lays down norms for basic school infrastructure, student teacher ratios, teacher training requirements, and prohibitions of corporal punishment among other things (Ministry of Human Resource Development 2019). Beyond the RTE Act, and to specifically tackle the gender inequalities in education, a national campaign called "*Beti*

Bachao, Beti Padhao" (save the girl child, educate the girl child) was also launched in 2014. It aims to coordinate action between ministries to improve the sex ratio as well as ensure the survival and education of the girl child (Ministry of Women and Child Development 2018).

The Ministry of Human Resource Development reports regularly on gender differences in various aspects of education including student enrolment, dropouts, and teacher composition. The 2018 report showed around 50% of the teachers in the country were female and concluded that this number was too low in comparison with nations such as Russia (99% female teachers) and Brazil (89%). It suggested that an increase percentage of female teachers in the education workforce would, therefore, make the Indian education system more globally competitive and also lower drop-outs (Ministry of Human Resource Development 2018). This faulty causal leap notwithstanding,[2] female education outcomes and the gender compositions of the teacher workforce are clearly being prioritized by the Indian Government, making this study relevant not only in terms of strengthening theory but also in terms of enhancing policy actions in the country.

Despite the recent policy stressing gender equality in education, India remains a hard case for gender representation in schools for two reasons. First, India is a highly heterogeneous country with major demographic cleavages based on caste, religion, income, language and region that are highly salient. These divisions are the bases for political action in India and they are likely to take precedence over gender issues. Second, the educational environment which is resource poor and subject to massive variation across the country provides little leverage for those teachers seeking to improve the education of girls, and what benefits can be achieved in schools face the harsh realities of gender discrimination in the workplace and society. In combination, India might be considered a highly unfavorable

[2] The report presents no evidence that Brazil and Russia attain better educational results. There is some evidence in existing studies that women teachers are more effective than male teachers in elementary and secondary education (see Keiser et al. 2002), it is unclear if these results hold across countries.

context for gender representation in the schools, suggesting any impact would be small to nonexistent.

Theoretical Framework for the paper

The Indian gender context illustrates the challenges faced in equalizing educational opportunities for young women, but the structure of education policy in the country creates situations that can provide theoretical leverage on the study of representative bureaucracy. Although substantial research has linked female teachers to better outcomes for female students, these studies have not been able to distinguish whether this results from active representation (that is, the teachers act to benefit the disadvantaged student) or from symbolic representation (students see a teacher who looks like them and increase effort to please that teacher, a role model effect). The extreme variation among Indian schools in the demands of the teaching job and the resources available can provide some insight on these different theoretical processes. This leverage can be exploited by illustrating how the impact of passive representation on educational outcomes varies in different contexts.

Active representation requires a teacher to make additional efforts beyond the regular job to encourage or work with students; existing theory suggests that as task difficulty is reduced or as the basic job burdens decline, the bureaucrat has more freedom to act as a representative (Meier 2019). Unlike the situations in the US where task difficulty varies but within modest limits, the task difficulty of Indian teachers can be extreme. The school year ranges (two standard deviations either way from the mean) from 139 days a year to 297 days and teacher working hours range from 4 to 9 hours per day. Similarly, while the mean student teacher ratio is 28 (compared to approximately 12-15 in the US), it rises above 100 in many schools (see Table 1 for details). Logically a teacher has more opportunities for active representation in a school with a student teacher ratio of 29 or less and when a school has more instructional days/working hours, than when student teacher ratios are massive, and

schools meet infrequently. This suggests that active representation is more likely to be possible, and thus the relationship between female teachers and female student performance will be enhanced, when task difficulty is low. In short, passive representation should interact with task difficulty and resources and show a larger impact on outcomes when task difficulty is low. The operational hypothesis is that passive representation will have a stronger association with student outcomes when student to teacher ratios are lower, when teacher working hours are higher, and when schools meet for more class days.

The Indian context might also provide situations where active representation is difficult or impossible and thus the only possible influence of passive representation will be symbolic, through creating role models. Given the extreme range in incomes and wealth in India and the corresponding range in resources allocated to education, many schools serve a population that is impoverished and isolated from potential opportunities to change. Similarly, access to resources exacerbates the problem of wealth inequalities. Indian schools vary dramatically in resources particularly when infrastructure is considered. Some schools in India lack basics such as a library, electricity, or even a school building. In such situations where there are few prospects for students and even fewer prospects for female students, likely the only influence of the teacher who faces a challenging job and has few resources is as a role model. Those circumstances should be manifest in rural schools and in schools that have inadequate infrastructure. Representation, to the extent it shows up in such circumstances, is likely to reflect symbolic responses since active representation is likely to be limited. The operational hypothesis linked to symbolic representation suggests that passive representation will have a stronger association with educational outcomes where there are few possible role models other than teachers, that is, when schools lack infrastructure, and when the schools are in rural areas.

Data and Key Variables

The dataset used in this study comes from the Unified District Information System for Education (U-DISE) which is run by the National University of Educational Planning and Administration under the guidance of the Ministry of Human Resource and Development of India. U-DISE has been collecting data from all recognized schools in the country, covering basic infrastructure facilities, location, funding, management, student/teacher counts, demographic make-up, dropout rates, and end of year examination performance. It is the most comprehensive and disaggregated database available on primary and secondary educational institutions in India and currently includes information on 1.5 million schools from 2005 to 2016. The data is submitted annually by school principals to specialized staff at the district level, and two levels of consistency checks are conducted before it is uploaded to the information system (National Institute of Educational Planning and Administration n.d.). Due to restrictions in the availability of data for the dependent variable in this study, we utilize data from only four academic years: 2014-15 to 2017-18. Next, we describe the key variables used in the analysis, the summary statistics for which can be found in the appendix in Table 1.

Independent Variable

To measure passive gender representation amongst the teaching staff in schools, we generated a variable from the dataset that computed the percentage of female teachers. The unit of analysis is an individual school; hence, the independent variable measures percentage of female teachers in the school. Post 4th grade, students get to interact with most teachers in their school through subject teaching, extra-curricular and co-curricular activities. The variable will therefore measure the overall effect female teachers can have on students across the school (see Favero and Molina 2018 for organizational effects of representation) and is complemented by the dependent variables that measure overall academic performance and are detailed next. Despite the gender-skewed national labor pool in the country, on average

41% of teachers are female, which points to teaching being a gendered occupation as well in the country.

To measure the potential for active representation, we use three variables as proxies for task difficulty: the total number of class days for students in 8th grade in the year, the number of teachers working hours in a day, and the ratio of students to teachers in the school. To assess the conditions that might limit representation to symbolic representation, we use two environmental factors as variables. Firstly, we include a school-level infrastructure index which is a factor score (Eigenvalue 1.6) generated to measure the basic level of facilities available to students and teachers, including computer labs, library, playgrounds, and electricity. Second, we include a binary variable that denotes whether the school is located in a rural area of the country or not.

Dependent Variables

To measure student academic achievement, we used the percentage of girls scoring more than 60% in the 8th grade end of year examination. In the Indian education system, the 8th grade is considered a gateway grade after which students enter secondary school. While the U-DISE dataset has information on both the number of students who pass and those that score 60%+, disaggregated by gender, we focus on the latter measure of academic success. We do so because the 8th grade examinations are not standardized, and grading is at the discretion of the school. This may lead teachers to pass most students to the next grade unless there is a grave reason not to. The dataset confirms this, finding with a mean pass percentage of ~90% for both the examinations. Scoring 60% or above, however, is a way for the teacher to communicate academic progress and future academic success in the crucial secondary grades. There is more variation seen in this measure of academic success and hence it is used to generate the key dependent variable, which is the percentage of girls scoring 60%+ in the 8th grade examination, of the total number of girls that appeared for it. As can be seen from

Table 1, on average 62% girl students score 60%+ in the 8th grade examination, with considerable variation noticed across schools and years.

Control Variables

Controls that could affect the academic performance of students or any of the key independent variables were included so that we could test whether the gender representation relationship was spurious. These controls can be grouped into 3 sets: student, teacher, and school management/infrastructure characteristics.

First, a control for the caste-related diversity of the student population in each school is included by creating a Herfindahl index using the percentage of student population across Scheduled Caste (SC), Scheduled Tribe (ST), Other Backward Class (OBC) and non-disadvantaged group categories. Research has found that diversity in the student population affects academic achievement outcomes, although studies have not yet covered the Indian context (Bankston and Caldas 1996). Moreover, the three groups included are historically disadvantaged due to the caste system prevalent in India with disparities noticed between them and the rest of the population in education, income and location (Borooah 2005; Dunn 1993). Second, the educational qualifications of teachers are controlled for in the model. Teachers who are college graduates and who have professional teaching qualifications would possibly be better at teaching and, therefore, influence the academic performance of students.

Finally, four school characteristics are controlled for, including whether the school is run by the government or by a private body, whether the medium of instruction in the school is English, a constructed school quality index, and the sex ratio in the school, calculated as the ratio of female students to male students. Students in private schools are found to perform better academically than those in government schools in India (White et al. 2016). Additionally, government schools do not charge students any fees for attendance, which gives us some indication of the socio-economic status of their parents. English medium

schools are popular amongst certain parents because the reputed best higher education systems in India (post 12th grade) provide instruction in English. Therefore, students in English medium schools would possibly have higher chances of continuing education post 12th grade but could also struggle more academically if their mastery of the English language is poor. The school quality index, calculated to provide for the existence of a school management committee, school development plan, textbook provision by the school, and the existence of special student trainings is included to capture the overall quality of the school management/ governance, that could affect the academic scores of girl students. The sex ratio in the school is included to provide for any gender disparities in the student population that could also have representation effects on the students.

Methodology

Since the purpose of the study is to capture the variation of the representation effect between schools, the hypotheses were tested using a pooled Ordinary Least Square (OLS) regression model with time fixed effects. Additionally, the dependent variable measures 8th grade performance, therefore only those schools that had 8th grade classes in the 2014-18 time period were included as observations, leading to an N of ~1.25 million schools. The nature of the data collection process meant that we encountered some unrealistic outliers when new variables were generated (e.g. percentage values that were higher than a 100). The dataset was therefore cleaned to only include percentage values between 1-100 and variable values that fell within 4 standard deviations of the mean. For any given variable, this did not lead to a drop of more than 3% of the observations. Lastly, to avoid heteroscedasticity, standard errors were clustered by the school codes.

Findings and Discussion

Table 2 in the appendix presents the linear and non-linear model of the association between female teachers and female students' performance on the eighth-grade end of school

exam, without any interaction terms. A one percentage point increase in female teachers is associated with an increase of .08 percentage points in female students scoring 60 percent or better on the exam. Although the marginal effect is relatively small, the variance in percentage of female teachers is quite large (34.8%), and so the cumulative difference adds up, as can be seen in Figure 1. The impact of female teachers, however, is not the only or even the strongest determinant of female student test scores which also reflect the quantity of instruction, school infrastructure, the school type/characteristics and economic conditions. Since the modeling effort focused on trying to account for as many factors as possible to ensure the gender relationship was not spurious, we will not spend time in interpreting the coefficients for the control variables but rather focus on the representation relationships.

As noted above, the theoretical advantage of the Indian case for studying representative bureaucracy is that there is extensive variation in the context which allows greater precision in specifying contexts that can facilitate or hinder various forms of representation. Our first set of hypotheses concerns the task demands on the teacher and the situations where task demands, or resources are such that teachers can perceive that active representation will potentially be effective – three such moderating variables are considered here: number of instructional days, teacher working hours and the student teacher ratios. Our hypothesis is that passive representation is more likely to translate into active representation and influence female test scores when the teacher has more instructional days and working hours to spend with the student, and when the student teacher ratio is lower.

The interaction of percent female teachers with number of instructional days, teacher working hours, and the student-teacher ratio is presented in Table 3. All three interactions generate strong relationships in the predicted direction, that is, female teachers are associated with higher girls' test scores when there are more instructional days, when they have longer working hours in a day and when the student-teacher ratio is smaller. Because interactions

are at times difficult to interpret, we present a series of graphs that illustrate the marginal impact of female teachers at various levels of the two interactive relationships.

Figure 3 illustrates that the marginal effect of female teachers on female test scores more than doubles, from approximately .04 when students have 100 days of instruction in a year to approximately .09 when they have 250 days of instruction. The representative effect of female teachers, therefore, increases as the working conditions get more favorable to active representation. Similarly, in figure 4 we see a doubling of the marginal effect from when teachers work 4 hours (0.05) to when they work 8 hours (0.1). Lastly, figure 5 shows the marginal effect of female teachers at various ratios of students to teachers. In line with the moderating effects of instructional days and teacher working hours (albeit at a more modest level), as the student teacher ratio drops from very high (120+) to levels more akin to those found in the US (~15), the marginal effect increases from 0.06 to 0.08.

Table 3 focused on situations where the task demands of teaching were not so severe that a teacher might be able to engage in active representation. In Table 4, we examine cases where active representation is difficult, but teachers might be one of the few positive role models available to female students – cases with little educational infrastructure, and schools located in rural area. Using figures to illustrate the results, figure 6 shows that female teachers in rural areas can increase the percentage of girls scoring 60%+ by nearly ten percentage points. Interestingly, in urban schools, an increase in female teacher percentage has a negative effect of the academic performance of the girl students in those schools. This finding will need to be explored further to unpack the reason behind this reverse effect in the 15% of urban area schools in the country. A similar relationship for the marginal influence of female teachers in schools with different infrastructure provisions is shown in figure 6. The marginal effect of female teachers more than doubles for schools with the lowest score on the infrastructure index (0.13), in comparison to schools with the highest score (0.06).

These findings show that if we consider a spectrum of environmental/organizational conditions that affect representation in education, on one end we have factors indicating quality, that may strengthen active representation efforts of teachers. On the other end we have factors that indicate the most basic resources required to facilitate learning for the salient identity group, the absence of which may strengthen the symbolic advantages of having a bureaucracy be representative of that population in need.

Conclusion

This study takes us a step closer towards understanding the generalizability of the representative bureaucracy theory, and the conditions under which passive representation leads to either active or symbolic representation. While we did find a representative effect of female teachers on the academic achievement of female students, the effect was modest, which points to the presence of other factors that influence the relationship. We analyzed some of these factors and found evidence of enhanced active representation when the teachers have more instructional days with the students and a smaller student teacher ratio. Additionally, we found that in situations where active representation would be difficult (schools located in rural areas, and with poor infrastructure), we still see a positive effect of female teachers on female student academic achievement. This supports our hypothesis that the role model effect of female teachers is likely to be higher in the absence of any other role models for the female students.

There are a few limitations of this study. Firstly, a preliminary analysis of the data at the state-level shows considerable variation in representation, which points to a need to unpack the dataset and explore the heterogenous effects in the different Indian states. Secondly, we cannot say with utmost certainty that active and symbolic representation plays out the way it is stated in the paper. Interviews and surveys of current teachers in the Indian education system will strengthen the theoretical propositions and empirical findings of this

paper. Thirdly, given the salience of the gender identity and its long-lasting effects on Indian women through their life, it would be interesting to explore the effect female teachers have on girl students beyond academic achievement. This will paint a richer picture of the span of control of K-12 teachers in India. Finally, it could be argued that understanding representative bureaucracy in the unique national context of India is not enough to generalize the theory to other developing country democracies, South Asia or nations that were former British colonies. Future research can tackle each of these areas to further enhance our understanding of representation and the conditions under which it occurs.

References

Agyapong, Elijah. 2017. "Representative Bureaucracy: Examining the Eeffects of Female Teachers on Girls' Education in Ghana." *International Journal of Public Administration* 41(16): 1–13. https://doi.org/10.1080/01900692.2017.1388255.

Andrews, Rhys, and Karen Johnston Miller. 2013. "Representative Bureaucracy, Gender, and Policing: The Case of Domestic Violence Arrests in England." *Public Administration* 91(4): 998–1014.

Atkins, Danielle N., and Vicky M. Wilkins. 2013. "Going beyond Reading, Writing, and Arithmetic: The Effects of Teacher Representation on Teen Pregnancy Rates." *Journal of Public Administration Research and Theory*.

Bankston, Carl, and Stephen J. Caldas. 1996. "Majority African American Schools and Social Injustice: The Influence of de Facto Segregation on Academic Achievement." *Social Forces*.

Borooah, Vani K. 2005. "Caste , Inequality , and Poverty in India." 9(3): 399–414.

Dee, Thomas S. 2005. "A Teacher like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review*.

Dunn, Dana. 1993. "Gender Inequality in Education and Employment in the Scheduled Castes and Tribes of India." *Population Research and Policy Review* 12(1): 53–70.

Favero, Nathan, and Angel Luis Molina. 2018. "Is Active Representation an Organizational-Level Process? The Indirect Effect of Bureaucrats on Clients They Don't Directly Serve." *American Review of Public Administration* 48(1): 3–17.

van Gool, Bas, and Frank de Zwart. 2013. "Politics of Representative Bureaucracy in India." In *Representative Bureaucracy in Action*, Edward Elgar Publishing, 188–203. http://www.elgaronline.com/view/9780857935984.00020.xml.

Grissom, J. A., E. C. Kern, and L. A. Rodriguez. 2015. "The 'Representative Bureaucracy' in Education: Educator Workforce Diversity, Policy Outputs, and Outcomes for Disadvantaged Students." *Educational Researcher* 44(3): 185–92. http://edr.sagepub.com/cgi/doi/10.3102/0013189X15580102.

Grissom, Jason A., Jill Nicholson-Crotty, and Sean Nicholson-Crotty. 2009. "Race, Region, and Representative Bureaucracy." *Public Administration Review*.

Guul, Thorbjørn Sejr. 2018. "The Individual-Level Effect of Gender Matching in Representative Bureaucracy." *Public Administration Review* 78(3): 398–408.

Jayachandran, Seema. 2015. "The Roots of Gender Inequality in Developing Countries." *Ssrn*.

Keiser, Lael R, Vicky M Wilkins, Kenneth J Meier, and Catherine A Holland. 2002. "Lipstick and Logarithms : Gender , Institutional Context , and Representative Bureaucracy Published by : American Political Science Association Lipstick and Logarithms : Gender , Institutiona Representative Bureaucracy." *The American Political Science Review* 96(3): 553–64. https://www.jstor.org/stable/pdf/3117929.pdf?refreqid=excelsior%3Ac7cbf8df1ceaa923dc3ea6b9cfeba4cb.

Kennedy, Brandy. 2014. "Unraveling Representative Bureaucracy: A Systematic Analysis of the Literature." *Administration and Society* 46(4): 395–421.

Malhotra, Anju, Rohini Pande, and Caren Grown. 2003. "Impact of Investments in Female Education on Gender Equality." *International Center for Research on Women*. http://siteresources.worldbank.org/INTGENDER/Resources/ImpactInvestmentsFemaleEdu.pdf.

Meier, Kenneth J. 2018. "Theoretical Frontiers in Representative Bureaucracy: New Directions for Research." *Perspectives on Public Management and Governance* (August): 1–18. https://academic.oup.com/ppmg/advance-article/doi/10.1093/ppmgov/gvy004/5037752.

Meier, Kenneth J, Sharon H Mastracci, and Kristin Wilson. 2006. "Gender and Emotional Labor in Public Organizations : An Empirical Examination of the Link to Performance Author (s): Kenneth J . Meier , Sharon H . Mastracci and Kristin Wilson Published by : Wiley on Behalf of the American Society for Public Administra." 66(6): 899–909.

Meier, Kenneth John. 1975. "Representative Bureaucracy: An Empirical Analysis." *American Political Science Review*.

Meier, Kenneth John, and Lloyd G. Nigro. 1976. "Representative Bureaucracy and Policy Preferences: A Study in the Attitudes of Federal Executives." *Public Administration Review* 36(4): 458. http://www.jstor.org/stable/974854?origin=crossref.

Ministry of Human Resource Development. 2018. *Educational Statistics 2018*. http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s7567/list.html.

Ministry of Human Resource Development, Government of India. 2019. "Right to Education." https://mhrd.gov.in/rte (March 23, 2019).

Ministry of Women and Child Development. 2018. "Beti Bachao Beti Padhao." http://www.wcd.nic.in/bbbp-schemes (December 11, 2018).

Morton, Tabitha S.M. 2015. "Improving Educational Outcomes for Latinos: A Study of the Interactive Policy Effects of Representative Bureaucracy and Personnel Stability." *Journal of Latinos and Education* 14(4): 301–19. http://dx.doi.org/10.1080/15348431.2015.1091316.

Mosher, Frederick C. 1982. *Democracy and the Public Service*. 2nd ed. New York: Oxford University Press.

National Institute of Educational Planning and Administration. "Evolving DISE: Strengthening of EMIS in India." http://schoolreportcards.in/SRC-New/AboutDISE/AboutDISE.aspx (March 18, 2019).

Nicholson-Crotty, Sean et al. 2017. "Exploring the Causal Story: Gender, Active Representation, and Bureaucratic Priorities." *Journal of Public Administration Research and Theory* 77(4): 911–19. https://academic.oup.com/jpart/article/28/4/506/5041982.

Office of the Registrar General & Census Commissioner. 2011. "Sex Ratio." *2011 Census Data*. http://censusindia.gov.in/Census_Data_2001/India_at_glance/fsex.aspx (February 6, 2019).

Oxford Poverty and Human Development Initiative. 2018. "Global Multidimensional Poverty Index." https://ophi.org.uk/multidimensional-poverty-index/ (March 18, 2019).

Pitts, David W. 2005. "Diversity, Representation, and Performance: Evidence about Race and Ethnicity in Public Organizations." *Journal of Public Administration Research and Theory* 15(4): 615–31.

Riccucci, Norma M., and Marcia K. Meyers. 2004. "Linking Passive and Active Representation: The

Case of Frontline Workers in Welfare Agencies." *Journal of Public Administration Research and Theory* 14(4): 585–97.

Riccucci, Norma M., Gregg G. Van Ryzin, and Cecilia F. Lavena. 2014. "Representative Bureaucracy in Policing: Does It Increase Perceived Legitimacy?" *Journal of Public Administration Research and Theory* 24(3): 537–51.

Riccucci, Norma M, Gregg G Van Ryzin, and Karima Jackson. 2018. "Representative Bureaucracy, Race, and Policing: A Survey Experiment." *Journal of Public Administration Research and Theory* 28(4): 506–18. https://academic.oup.com/jpart/article/28/4/506/5041982.

Roch, Christine H., and Jason Edwards. 2015. "Representative Bureaucracy and School Discipline: The Influence of Schools' Racial Contexts." *American Review of Public Administration* 47(1): 58–78.

Selden, Sally Coleman. 1997. *The Promise of Representative Bureaucracy: Diversity and Responsiveness in a Government Agency*. M.E. Sharpe.

Song, Miyeon. 2018. "Gender Representation and Student Performance: Representative Bureaucracy Goes to Korea." *American Review of Public Administration*.

Stearns, Elizabeth et al. 2016. "Demographic Characteristics of High School Math and Science Teachers and Girls' Success in STEM." *Social Problems*.

Theobald, Nick A., and Donald P. Haider-Markel. 2009. "Race, Bureaucracy, and Symbolic Representation: Interactions between Citizens and Police." *Journal of Public Administration Research and Theory*.

United Nations Development Programme. 2018. "Human Development Reports." http://hdr.undp.org/en/countries/profiles/IND (December 11, 2018).

Vinopal, Katie. 2017. "Understanding Individual and Organizational Level Representation: The Case of Parental Involvement in Schools." *Journal of Public Administration Research and Theory* 28 (1): 1-15.

White, G., M. Ruther, J.R. Kahn, and D. Dong. 2016. "Educational Inequality in India: An Analysis of Gender Differences in Reading and Mathematics." *Journal of Research in Gender Studies* 6(2): 153–82.

Wilkins, Vicky M., and Lael R. Keiser. 2006. "Linking Passive and Active Representation by Gender: The Case of Child Support Agencies." *Journal of Public Administration Research and Theory* 16(1): 87–102.

Zhang, Youlang. 2018. "Representative Bureaucracy, Gender Congruence, and Student Performance in China." *International Public Management Journal* 7494: 1–22.

Appendix

Table 1. Descriptive Statistics

Variable	Description		Mean.	Std. Dev.	Min	Max	Observations
Key Independent Variable							
Female Teachers (%)	Percent of total teachers in the school that are female	overall	41.454	34.828	0	100	N = 6060797
		between		33.749	0	100	n = 1703674
		within		9.757	-33.546	116.454	T-bar = 3.55749
Dependent Variable							
Girls Scoring 60%+ in grade 8 (%)	Percent of girls that scored more than 60% in 8th grade exam	overall	61.612	35.196	0	100	N = 1510986
		between		30.247	0	100	n = 569538
		within		20.467	-13.389	136.611	T-bar = 2.653
Active Representation Variables							
Instructional Days	Number of instructional days for students in the year	overall	89.046	109.624	0	250	N = 6049093
		between		104.883	0	250	n = 1715537
		within		34.336	-98.454	276.547	T-bar = 3.52606
Teacher Working Hours	No. of hours the teachers work in a day	overall	2.637	3.236	0	12	N = 6133640
		between		3.091	0	10	n = 1720765
		within		1.028	-4.863	10.637	T-bar = 3.56448
Student Teacher Ratio	Ratio of students to teachers in school	overall	27.157	21.139	0	150	N = 5985367
		between		20.454	0	150	n = 1692853
		within		8.749	-82.843	136.989	T-bar = 3.53567
Symbolic Representation Variables							
Rural School	Dummy: 1 means the school is located in a rural area	overall	.846	.362	0	1	N = 6133791
		between		.371	0	1	n = 1720771
		within		.043	.096	1.596	T-bar = 3.56456
Infrastructure Index	Factor: Measures availability of computers, library, playground and electricity	overall	.181	.956	-1.679	2.031	N = 6133791
		between		.937	-1.679	2.031	n = 1720771
		within		.240	-2.601	2.963	T-bar = 3.56456
Additional controls							
Teachers with Graduate Degrees	Percent of teachers in school that have a graduate degree	overall	66.361	37.254	0	100	N = 6060797
		between		34.903	0	100	n = 1703674
		within		14.124	-8.639	141.361	T-bar = 3.55749
Teachers with Professional Qualification	Percent of total teachers in school that have a professional teacher qualification	overall	80.982	33.266	0	100	N = 6060797
		between		32.571	0	100	n = 1703674
		within		10.096	5.982	155.982	T-bar = 3.55749
Government School	Dummy: 1 means the school is managed by a govt body (0 is a private body)	overall	0.717	.451	0	1	N = 6133791
		between		.463	0	1	n = 1720771
		within		.033	-.034	1.467	T-bar = 3.56456
English medium school	Dummy: 1 means the medium of instruction in the school is English	overall	.102	.303	0	1	N = 6133791
		between		.315	0	1	n = 1720771
		within		.053	-.648	.852	T-bar = 3.56456
School sex ratio	Ratio of the total girls in the school to total boys	overall	1.183	8.474	0	2649	N = 5932074
		between		14.565	0	2649	n = 1671486
		within		1.933	-785.149	1181.85	T-bar = 3.54898
Student to classroom ratio	Ratio of total students to total classrooms in the school	overall	29.498	31.318	0	300	N = 5608988
		between		31.121	0	300	n = 1628249
		within		11.237	-157.502	249.248	T-bar = 3.4448
School Quality Index	Factor: measures presence of a management committee, development plan, student special training & textbooks	overall	.002	.999	-1.924	.807	N = 6133791
		between		1.027	-1.924	.807	n = 1720771
		within		.149	-2.047	2.049	T-bar = 3.56456
Caste Herfindahl Index	Index: Measures the size of each caste group in the school as an indicator of diversity	overall	.652	.232	.25	1	N = 6076329
		between		.222	.25	1	n = 1708232
		within		.073	.089	1.213	T-bar = 3.55709

Table 2. Pooled OLS regression analysis with time fixed effects

Dependent Variable: % girls scoring 60%+ in 8th grade exam

VARIABLES	(1) Linear	(2) Lagged DV	(3) Quadratic
Female teachers (%)	0.0821***	0.0522***	0.224***
	(0.00137)	(0.00133)	(0.00400)
Lagged girls scoring 60%+ (%)		0.406***	
		(0.00134)	
Student teacher ratio	-0.0551***	-0.0487***	-0.0460***
	(0.00196)	(0.00197)	(0.00197)
Class to student ratio	-0.0767***	-0.0415***	-0.0830***
	(0.00110)	(0.00105)	(0.00111)
Teacher work hours (per day)	3.051***	2.166***	3.048***
	(0.0341)	(0.0369)	(0.0340)
Instructional days (per year)	-0.0748***	-0.0363***	-0.0745***
	(0.00110)	(0.00125)	(0.00110)
English medium school	-1.639***	-0.414***	-1.407***
	(0.123)	(0.125)	(0.123)
Caste Herfindahl Index	4.989***	3.071***	5.207***
	(0.182)	(0.174)	(0.182)
Infrastructure index	4.859***	2.848***	4.722***
	(0.0429)	(0.0412)	(0.0430)
School quality index	-0.764***	0.160***	-0.700***
	(0.0581)	(0.0571)	(0.0581)
Rural school	0.623***	0.669***	0.331***
	(0.104)	(0.102)	(0.104)
Government school	-15.20***	-9.424***	-15.52***
	(0.125)	(0.122)	(0.125)
Teachers with graduate degrees (%)	-0.123***	-0.0815***	-0.121***
	(0.00127)	(0.00126)	(0.00127)
Teachers with teaching qualification (%)	0.161***	0.0981***	0.166***
	(0.00146)	(0.00153)	(0.00147)
School sex ratio	-0.0263***	-0.0175***	-0.0221***
	(0.00246)	(0.00196)	(0.00236)
Academic year 2015-16	-2.806***	-7.305***	-2.828***
	(0.0671)	(0.100)	(0.0670)
Academic year 2016-17	-1.340***	-5.173***	-1.364***
	(0.0695)	(0.101)	(0.0695)
Academic year 2017-18	5.051***		4.963***
	(0.0898)		(0.0898)
Squared % female teacher			-0.00157***
			(4.06e-05)
Constant	61.03***	37.70***	59.11***
	(0.302)	(0.331)	(0.307)
Observations	1,242,766	689,591	1,242,766
R-squared	0.132	0.282	0.134

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 1: Linear Regression Model

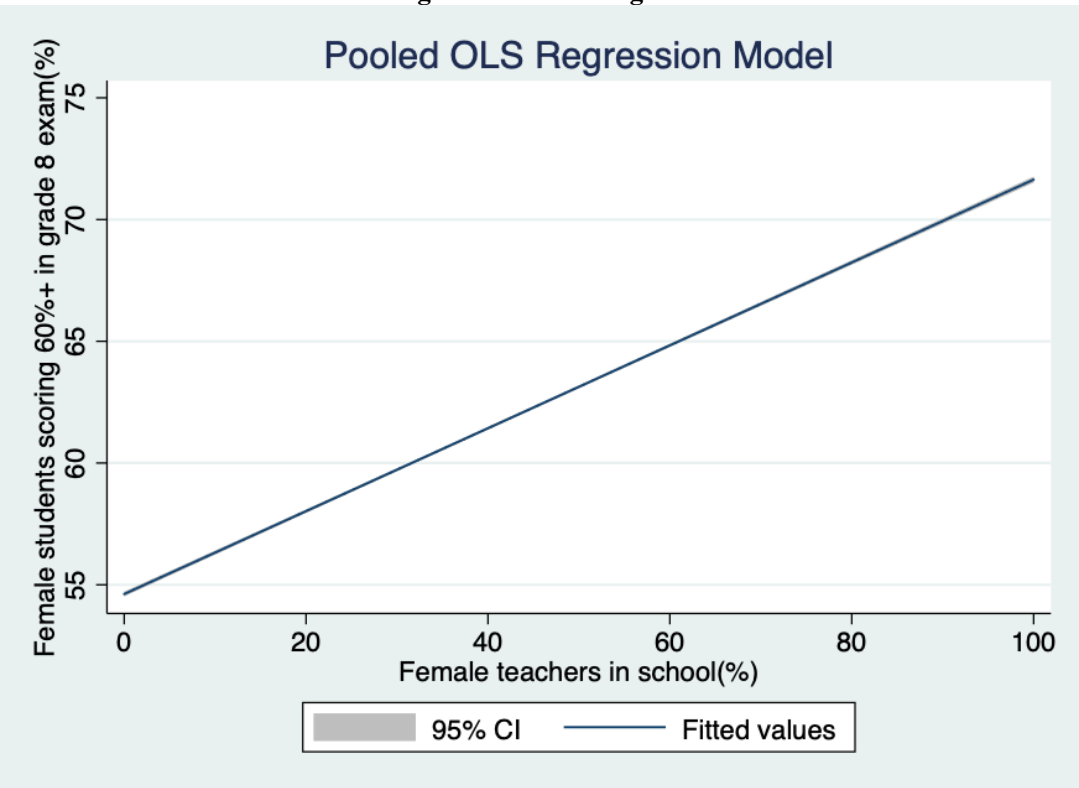

Figure 2: Quadratic Regression Model

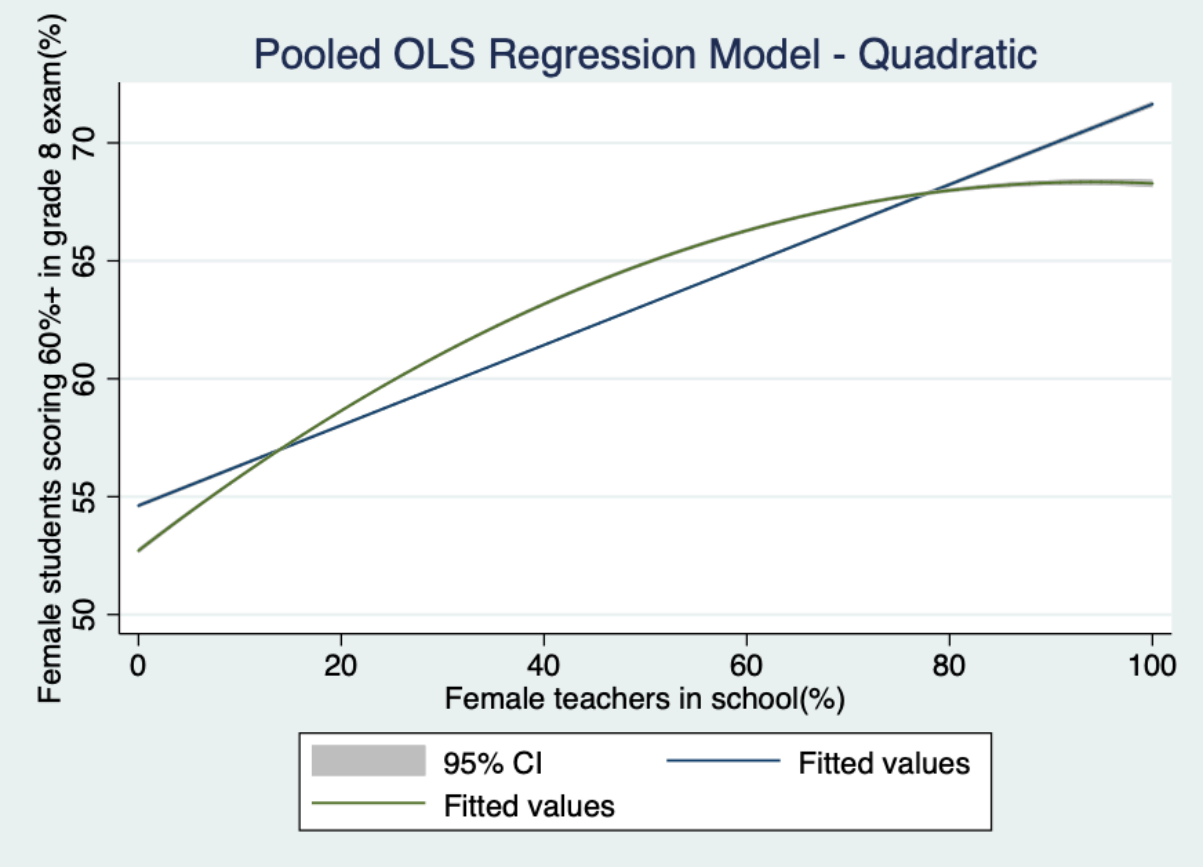

Table 3. Interactions for active representation

Dependent Variable: % girls scoring 60%+ in 8th grade exam

VARIABLES	(1) Instructional days	(2) Teacher work hours	(3) Student teacher ratio
Female teachers (%)	0.0102	-0.00911	0.0875***
	(0.00645)	(0.00559)	(0.00211)
Student teacher ratio	-0.0553***	-0.0547***	-0.0487***
	(0.00196)	(0.00196)	(0.00278)
Class to student ratio	-0.0766***	-0.0771***	-0.0765***
	(0.00110)	(0.00110)	(0.00110)
Teacher work hours (per day)	3.053***	2.437***	3.047***
	(0.0341)	(0.0505)	(0.0341)
Instructional days (per year)	-0.0873***	-0.0731***	-0.0747***
	(0.00156)	(0.00111)	(0.00110)
English medium school	-1.616***	-1.663***	-1.636***
	(0.123)	(0.123)	(0.123)
Caste Herfindahl Index	4.982***	4.930***	4.994***
	(0.182)	(0.182)	(0.182)
Infrastructure index	4.856***	4.850***	4.857***
	(0.0429)	(0.0429)	(0.0429)
School quality index	-0.746***	-0.777***	-0.759***
	(0.0581)	(0.0581)	(0.0581)
Rural school	0.620***	0.580***	0.616***
	(0.104)	(0.104)	(0.104)
Government school	-15.21***	-15.16***	-15.20***
	(0.125)	(0.125)	(0.125)
Teachers with graduate degrees (%)	-0.123***	-0.123***	-0.123***
	(0.00127)	(0.00127)	(0.00127)
Teachers with teaching qualification(%)	0.161***	0.161***	0.161***
	(0.00146)	(0.00146)	(0.00146)
School sex ratio	-0.0263***	-0.0261***	-0.0260***
	(0.00246)	(0.00244)	(0.00246)
Academic year 2015-16	-2.810***	-2.786***	-2.805***
	(0.0671)	(0.0671)	(0.0670)
Academic year 2016-17	-1.341***	-1.311***	-1.337***
	(0.0695)	(0.0695)	(0.0695)
Academic year 2017-18	5.058***	5.070***	5.050***
	(0.0898)	(0.0898)	(0.0898)
Female teachers X Instructional days	0.000330***		
	(2.87e-05)		
Female teachers X Teacher working hrs		0.0142***	
		(0.000836)	
Female teachers X Student teacher ratio			-0.000187***
			(5.43e-05)
Constant	63.78***	64.62***	60.84***
	(0.385)	(0.368)	(0.308)
Observations	1,242,766	1,242,766	1,242,766
R-squared	0.132	0.132	0.132

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Figure 3: Interaction with Instructional days

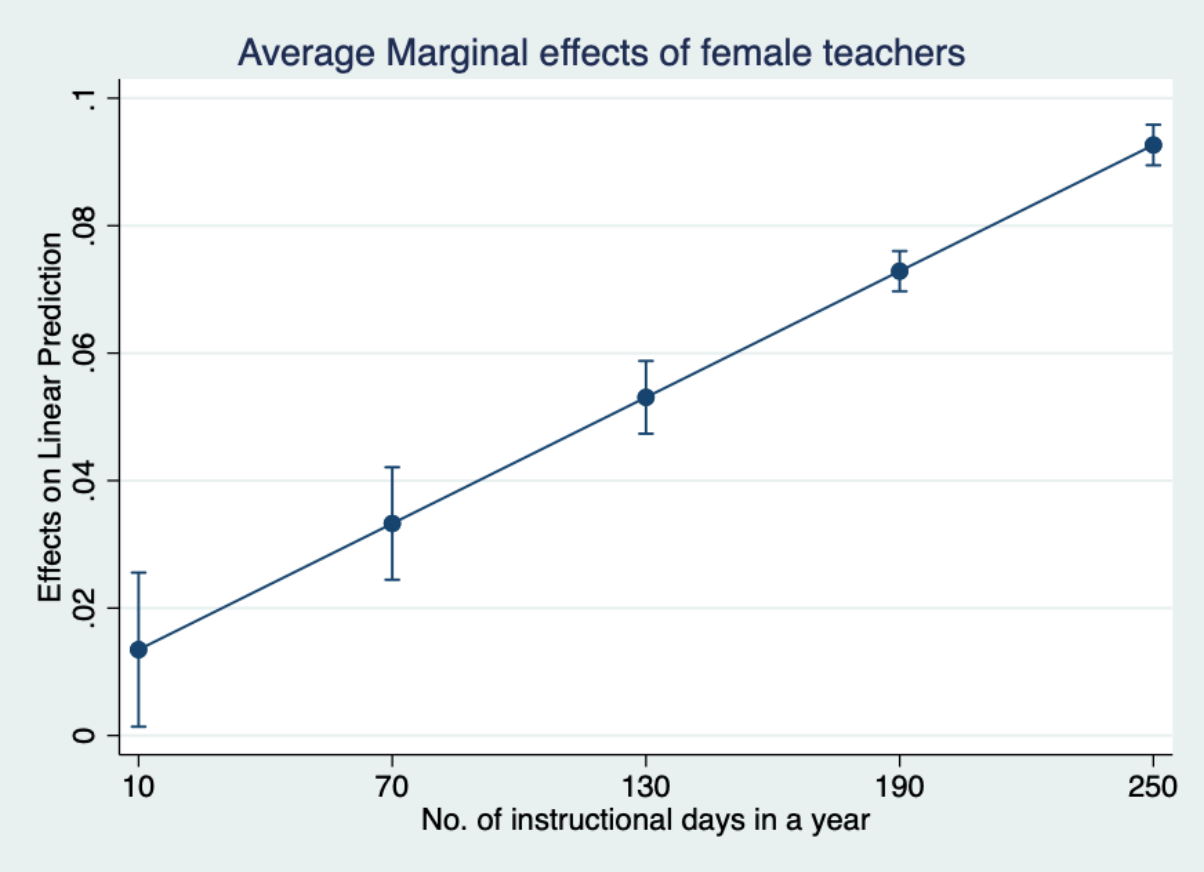

Figure 4: Interaction with teacher working hours

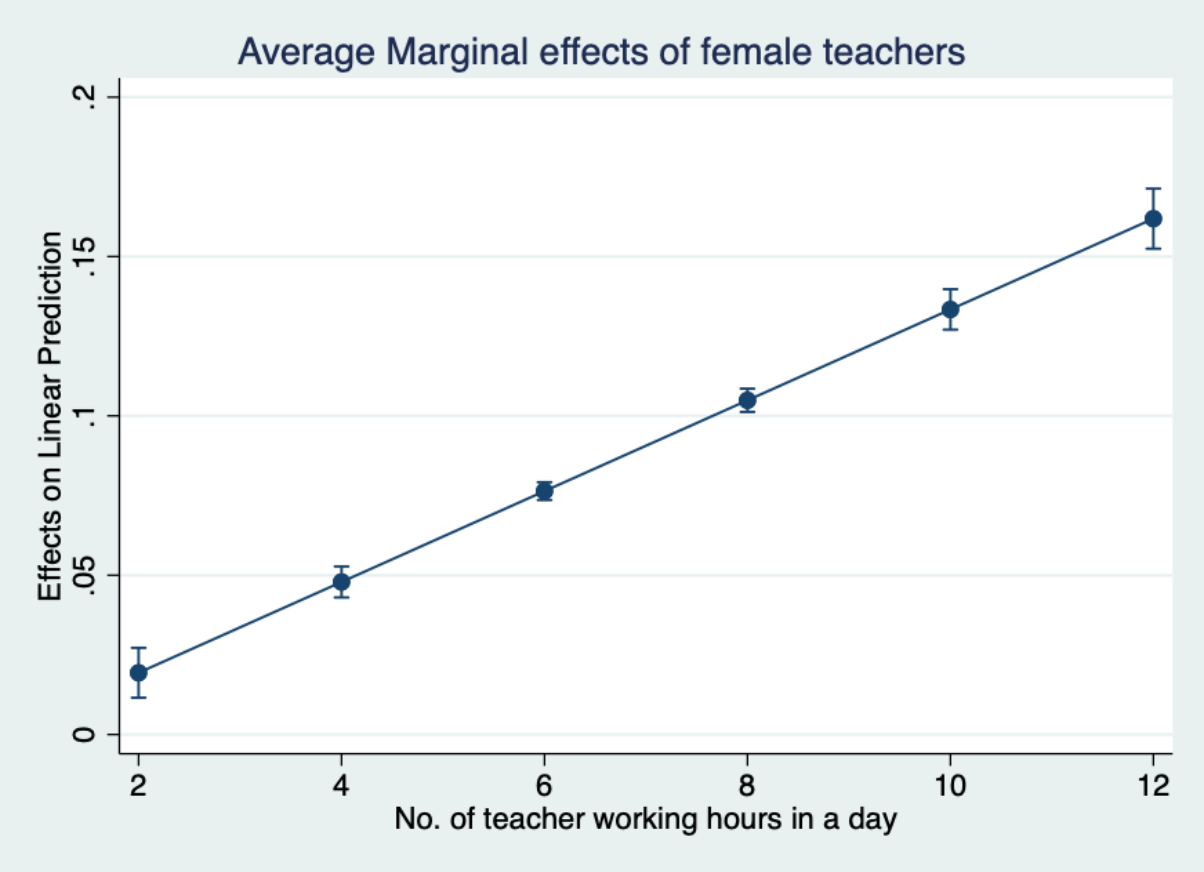

Figure 5: Interaction with student teacher ratio

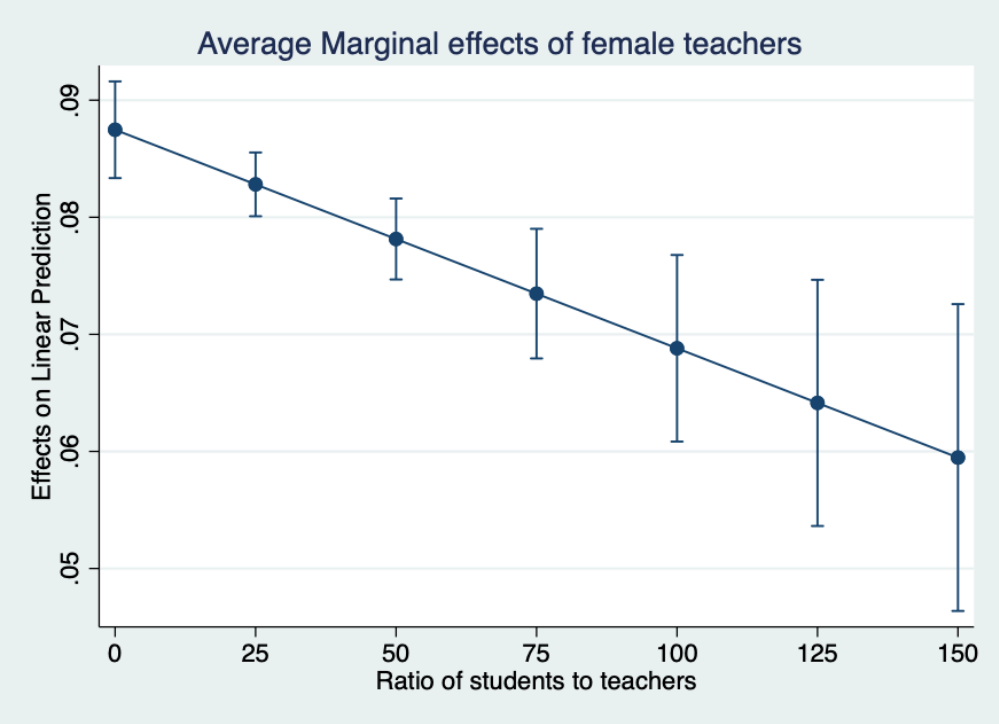

Table 4: Results of Interactions for Symbolic Representation

Dependent Variable: % girls scoring 60%+ in 8th grade exam

VARIABLES	(1) Rural schools	(2) Infrastructure Index
Female teachers (%)	-0.0147***	0.0921***
	(0.00302)	(0.00159)
Student teacher ratio	-0.0534***	-0.0541***
	(0.00196)	(0.00196)
Class to student ratio	-0.0779***	-0.0776***
	(0.00110)	(0.00110)
Teacher work hours (per day)	3.027***	3.064***
	(0.0340)	(0.0340)
Instructional days (per year)	-0.0741***	-0.0750***
	(0.00110)	(0.00110)
English medium school	-1.362***	-1.492***
	(0.123)	(0.123)
Caste Herfindahl Index	5.222***	5.110***
	(0.182)	(0.183)
Infrastructure index	4.810***	5.597***
	(0.0429)	(0.0703)
School quality index	-0.763***	-0.774***
	(0.0580)	(0.0581)
Rural school	-6.370***	0.577***
	(0.227)	(0.104)
Government school	-15.13***	-15.25***
	(0.125)	(0.125)
Teachers with graduate degrees (%)	-0.123***	-0.123***
	(0.00127)	(0.00127)
Teachers with teaching qualification (%)	0.161***	0.161***
	(0.00146)	(0.00146)
School sex ratio	-0.0247***	-0.0256***
	(0.00246)	(0.00244)
Academic year 2015-16	-2.795***	-2.813***
	(0.0670)	(0.0670)
Academic year 2016-17	-1.335***	-1.353***
	(0.0695)	(0.0695)
Academic year 2017-18	5.041***	5.030***
	(0.0897)	(0.0898)
Female teachers X Rural area school	0.116***	
	(0.00330)	
Female teachers X Infrastructure index		-0.0181***
		(0.00130)
Constant	67.03***	60.71***
	(0.346)	(0.303)
Observations	1,242,766	1,242,766
R-squared	0.133	0.132

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 6: Interaction with rural area schools

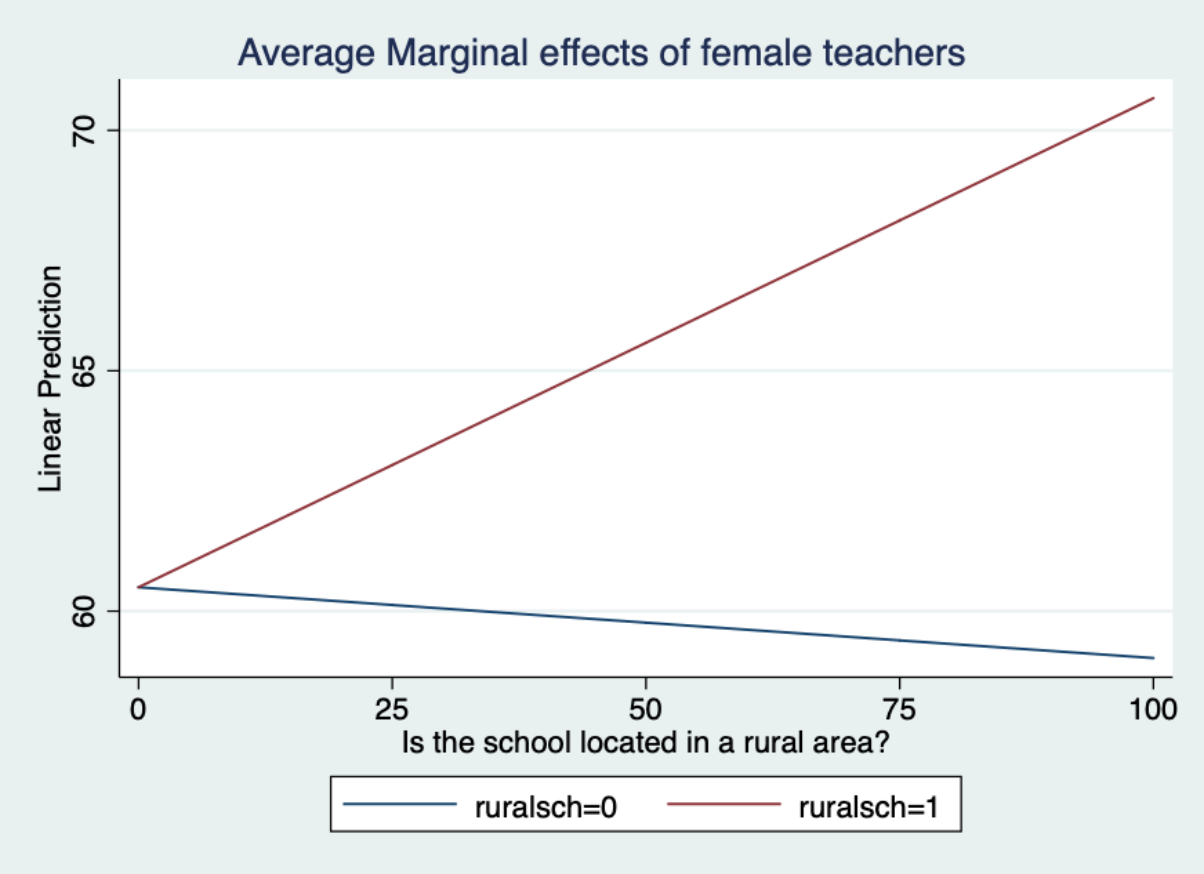

Figure 7: Interaction with infrastructure index