DISCUSSION PAPER SERIES

IZA DP No. 5687

Reconsidering Gender Bias in Intra-Household Allocation in India

Laura Zimmermann

April 2011

Reconsidering Gender Bias in Intra-Household Allocation in India

Laura Zimmermann
University of Michigan
and IZA

Discussion Paper No. 5687
April 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 5687
April 2011

ABSTRACT

Reconsidering Gender Bias in Intra-Household Allocation in India*

Detecting gender discrimination among children in the intra-household allocation of goods from household surveys has often proven to be difficult. This paper uses some of the commonly used techniques in this field to analyze education expenditures in India. Contrary to most previous research, I find evidence of discrimination against girls. Results at the all-India level are robust to the statistical method and the education expenditure measure, while they are more sensitive to changes in the analysis at the state level. In general, girls experience gender discrimination especially from age 10 onwards, with almost universal disadvantage in the amount of education expenditures in the group of 15-19 year olds.

JEL Classification: I24, I25, J16, O15

Keywords: gender discrimination, India, intra-household allocation, education expenditures

Corresponding author:

Laura Zimmermann
Department of Economics
University of Michigan
611 Tappan Street
Ann Arbor, MI 48109-1220
USA
E-mail: lvzimmer@umich.edu

* I thank Ashwini Deshpande, Suresh Tendulkar and participants at the Midwest Economics Association Annual Conference in Chicago for valuable comments and feedback.

1. Introduction

In a number of developing countries, girls have significantly worse education, health and nutrition outcomes than boys. One possible explanation of this discrepancy in child outcomes is gender discrimination in the intra-household allocation of resources: Parents may spend more on boys than on girls, both in monetary terms and in the time allocated to each child. Empirical as well as anecdotal evidence suggests that gender bias is considerable in some developing countries: Parents in India may wait longer before going to the hospital with their baby girl than they do with their baby boy, for example, leading to differences in health outcomes[1]. Rose (2000) finds that women work less after the birth of a boy and thus spend more time with their children in households with at least one boy than in households with only girls. In consequence, studying the existence and extent of gender bias in intra-household allocation has potentially important policy implications in order to improve outcomes for girls.

Unfortunately, most studies of gender discrimination in inputs such as consumption, health or education expenditures are hampered by the general unavailability of data. While there is enough data to study outcomes such as school enrollment or mortality, expenditure data is usually only available at the household level, rather than for each individual child. This makes it harder to analyze gender bias within households. Because of these data limitations, researchers have often used the Engel curve approach, which allows the use of household level data to try to get at gender bias within households. In the Engel curve approach, the household budget share of the good in question is regressed on log per-capita total expenditure, log household size, the shares of various age-sex groups and other relevant household characteristics. If there is gender bias in intra-household allocation, then the coefficient on the share of male members of a certain age group should be significantly different from the coefficient on the share of female members of the same age group.

By now, the Engel curve has been used to analyze gender bias in the allocation of a number of goods and in a variety of different countries, but strong results have been rare even in places where anecdotal evidence and measures of outcome variables suggest widespread discrimination against girls. This has lead to the questioning of the Engel curve approach both in terms of the statistical method used as well as the substantive question of whether discrimination can be picked up from household expenditure data or is only detectable in

[1] For this and other evidence on gender bias in developing countries see for example Asfaw, Klasen and Lamanna (2007), Klasen (1996), Messer (1997), Miller (1997), Sauerborn, Berman and Nougtara (1996)

other outcomes such as mortality. In general, however, the gender bias puzzle remains as most suggested alternatives do not lead to very different results.

Contrary to the mainstream literature, my analysis of gender discrimination in this paper leads to strong results, suggesting severe and widespread gender discrimination in intra-household allocation at least for the older age groups. I look at education expenditures in India, using nationally representative survey data from 2005. Gender bias is very robust to different measures of education expenditures and the statistical method used at the all-India level where discrimination works through both the decision to spend money on a boy but not on a girl (the extensive margin) and the decision to spend less money on the girl than on the boy, conditional on spending positive amounts on both children (the intensive margin). At the state level, there is some heterogeneity in the importance of the two channels across states and across age groups, but most discrimination occurs at the intensive margin. Results are sensitive to the education expenditure measure used in a minority of states, and their number declines as the age of the children increases. Pro-male bias also increases with age and becomes almost universal at the intensive margin for 15-19 year olds. In general, girls experience gender discrimination in the allocation of education expenditures especially from age 10 onwards.

These results suggest that the Engel curve approach is not generally failing to deliver results but may be sensitive to the level of analysis (country versus state level) and the specific measure of expenditure on a good when the analysis proceeds at the sub-national level.

The analysis in the rest of the paper proceeds as follows: Section 2 provides some background about the related literature. Section 3 explains the empirical strategy and the used dataset. Section 4 presents the results. Section 5 concludes.

2. Background Discussion

Ideally, researchers want to measure gender discrimination at the individual level by comparing household expenditures on each child. Given that expenditure information is often only available as an aggregate figure at the household level, however, the Engel curve approach has been widely used. Extending the traditional Engel curve to include household characteristics and composition variables yields a variant of the Working equation

$$b_i = \alpha + \beta \ln(x_i / n_i) + \gamma \ln(n_i) + \sum_{k=1}^{K-1} \delta_k (n_{ki} / n_i) + \eta z_i + \varepsilon_i \qquad (1)$$

where b is the budget share of the analyzed good, x is total expenditure, n is the household size, n_k is the number of people in the household that belong to age-sex class k, z consists of other socio-economic variables of interest, and ε is the error term (Deaton 1997).

To get more information about the sharing rule that households use to allocate resources between its members without imposing structural assumptions on household behavior, the literature relies on the use of exclusive goods, for instance goods that are only consumed by adults in the household, as dependent variables. For these goods, we know that only certain household members consume them. In my analysis, I will make use of children goods by considering educational expenditures. In this case, household consumption is children's consumption. If we further assume that additional household members act as negative income effects for these goods, but that there are no substitution effects, an additional person in the household has the same effect on the consumption of these goods as a reduction in income. The age-sex classes in the equation above then tell us the effect an additional member of a certain sex and age has on the consumption of the exclusive good.

Theoretically, the use of adult and children goods should lead to equivalent conclusions about intra-household bias against specific household members. Empirically, however, children goods have the advantage that they are measuring expenditures on children more directly than adult goods like tobacco or alcohol. When using adult goods, the analysis picks up whether parents cut back more strongly on the consumption of these goods when they have a boy of a certain age group rather than a girl, rather than measuring directly whether household expenditures on goods only consumed by children go up more for the boy than for the girl. The assumption that additional household members act just as income shocks and do not cause substitution effects may therefore be more likely to hold for children goods than for the indirect measurement through adult goods, which could change for reasons completely unrelated to additional household expenditures on children.

Equation (1) can be estimated empirically by using ordinary least squares on household level data. If there is gender bias in intra-household allocation, then the coefficient on the share of female household members should be significantly different from the coefficient on the share of male household members of the same age group.

By now, the Engel curve approach has been tested on a number of countries like Bangladesh, Burkina Faso, China, Cote d'Ivoire, India, Pakistan, Papua New Guinea, Sri Lanka or Thailand, and a variety of datasets and dependent variables, including adult goods like tobacco or alcohol, or other items like food or education expenditures. Unfortunately, strong results have been rare. The typical analysis is not able to detect strong gender bias even in countries where other indicators show severe discrimination against girls[2]. Individual level gender bias seems to somehow disappear at the household level. The case of education expenditures is not very different from the general picture: Subramanian and Deaton (1990) only find weak evidence of gender discrimination in rural Maharashtra for 10-14 year olds, but not for other age groups, while Lancaster, Maitra and Ray (2003) detect significant gender bias in their sample of rural Bihar and rural Maharashtra only for the age group of 10-16 year olds. Kingdon (2005) uses Indian survey data on 16 states but finds that the Engel curve approach does often not pick up existing individual level bias.

This widespread empirical failure of the Engel curve approach has led to a questioning of the reliability of the method by Case and Deaton (2003), but also to attempts to save the approach by trying to find the underlying reasons of this pattern. Various possibilities have been advanced: One potential cause is the functional form imposed by the OLS regression, so a number of papers have sought to relax this restriction. Bhalotra and Attfield (1998), for instance, estimate semi-parametric Engel curves and test whether potential effects are washed out by too large age categories, but still find hardly any difference between the treatment of children. Gong et al. (2005) also use semi-parametric analysis for different goods and conclude that gender bias seems to occur more through higher mortality rates and lower school enrollment for girls than boys, rather than through the consumption of common goods.

Another potential reason is the unreliability of household survey data. Gibson and Rozelle (2004) argue that the strong bias they find using the Engel curve approach for adult goods in Papua New Guinea might be driven by the large budget shares of these goods compared to other datasets, and the fact that their dataset was specifically designed for this analysis.

A third suggested reason in the literature is that aggregation at the household level leads gender bias to disappear. Kingdon (2005) and Aslam and Kingdon (2008) focus on education expenditures and are able to compare individual level and household level data to test whether individual level bias washes out at the household level. They indeed find stronger evidence of gender discrimination at the individual level rather than at the household level. Both papers also argue in favor of using a hurdle model that takes account of the censoring of the data, which is especially important in the education expenditure case where a significant number of households do not spend any money on education.

[2] See e.g. Ahmad and Murdoch 1993, Deaton 1989, Fuwa et al. 2006, Gibson 1997, Gibson and Rozelle 2004, Haddad and Reardon 1993, Himaz 2008, Lee 2008, Liu and Hsu 2004, Murdoch and Stern 1997, Subramaniam 1996

My approach in this paper is similar to Kingdon (2005) and Aslam and Kingdon (2008) insofar as I am also making use of education expenditure data and a hurdle model to take account of the censoring of the data. In contrast to these papers, however, I am able to look at the robustness of my results to changes in the measure of education expenditures by constructing two different variables, which show that especially in some states results are sensitive to the definition of education expenditures.

3. Data and Empirical Strategy

In this paper I use the India Human Development Survey (IHDS) from 2005, which is a nationally representative dataset, including 41,554 households from 1,503 villages and 971 urban neighborhood areas (Desai et al 2005). It includes both individual and household level responses on various topics such as education, employment, health, fertility, and gender relations.

I limit the dataset to households with children aged 0 to 19, which leaves me with a final sample size of 32,263 households. In order to warrant the analysis of gender bias in education expenditures, which can be seen as inputs to achieve the output of school attainment, it makes sense to first ask whether we do see different levels of school attainment in the dataset. If there is no significant difference between the school performance of girls and boys, different expenditures by sex might actually be justifiable in order to achieve this result that we ultimately care about. If we find that girls have lower school attainment than boys, on the other hand, it then makes sense to analyze education expenditures, which might be the cause for the difference in output.

In the dataset, all children in the household between 8 and 11 were asked to do a test which checked children’s performance in the areas of reading, writing and mathematics. For these variables, we can analyze whether there are significant differences between girls and boys. The results are reported in table 1. As we can see, there are strong gender differences in the test results in all three categories, with girls doing significantly worse than boys[3].

This suggests that there seems to be some evidence for anti-girl discrimination in school attainment which warrants moving on to the question whether girls are also discriminated against in the intra-household allocation of education expenditures.

The IHDS collects expenditure data on education in a disaggregated way and for both the household as a whole as well as for each individual child. This allows me to create two

[3] These results are consistent with evidence on gender differences in the Indian education system as summarized by Kingdon (2007).

measures of household expenditures that are similar, even if not completely equivalent to each other. We get the first measure, called education expenditure measure 1 from now on, by summing up all the reported household expenditures on school and private tuition fees, school books and other educational articles (for example newspaper, library charges, stationery, and internet charges) in the past year. We get the second measure, called education expenditure measure 2 from now on, by summing up all the reported expenditures on school fees, private tuition, school books, uniforms, transportation and other school materials for each child in the household in the past year. This gives us an alternative measure of household educational expenditures that has a broader definition than the first one but is applicable for the same households. Consistent with the broader definition of measure 2, school expenditures with the second measure have a slightly higher budget share on average than expenditures with the first measure as reported in table 2. The budget shares of education expenditures both are relatively small with less than 5%.

Table 2 also presents summary statistics for other important household characteristics. As we can see, a bit over 70% of households with children aged 0-19 incur positive education expenditures, which means that censoring is important for this sample. The average household size is 5.8, and the sample has the expected composition by religion and caste. Over a fifth of the sample are households classified as living below the official Indian Planning Commission's poverty line. About two thirds of households in the sample live in rural areas.

A natural starting point for the empirical analysis is using the Engel curve approach to detect gender bias in intra-household allocation of education expenditures as in the previous literature, especially because more flexible functional forms in general do not lead to very different results. I therefore estimate (1) from above, using the budget share of education expenditures as dependent variable. The age categories included in the regression are 0-4, 5-9, 10-14, 15-19, 20-55 and over 55. Females over 55 are taken as the reference group. As education expenditures are made for school going children, we expect the coefficients for both males and females of the age groups 5-9, 10-14 and 15-19 to be positive, which means that holding household size constant, increasing the share of household members from one of these three age categories will increase the budget share of education expenditures. These are our main estimates of interest, as we want to test whether household education expenditures increase significantly less for a girl than for a boy of the same age.

We should expect to find a negative coefficient for 0-4 year olds who are not yet attending school: Holding household size constant, an increase in the share of very young children should decrease household expenditure on education. The coefficients on the shares of the adults in the regression equation will depend on whether an additional adult will spend more on education than a female over 55 years old. Household variables included in the regression are religion and caste dummies, state fixed effects, an indicator variable for rural households and the highest education level for a male and a female in the household. We also estimate equation (1) for 16 big Indian states separately.

As the above regression is estimated by using OLS, results are potentially biased if the regressors are not exogenous. This has sometimes been raised in the literature as a potential reason for the usual lack of empirical results with the Engel curve approach. Jensen (2002) for example argues that fertility behavior in response to son preference will be an important determinant of family size: If son preference is high, then parents may get more children than they otherwise would until they reach the desired number of sons. This will mean that on average girls will live in larger families than boys, which puts more strain on household resources. This may well mean that girls will tend to have worse outcomes than boys merely as a result of larger family size, rather than as a consequence of intra-household discrimination. A related cause of potential endogeneity of the household size would occur if parents choose the number of children based on their preferences for school expenditures, for example by having fewer children with higher education expenditures per child[4].

As we are mainly interested in testing the equality of coefficients on the share of male and female household members of a certain age group, endogeneity will be particularly worrying if it affects the age-sex groups. Suppose for example that a preference for sons is positively correlated with education spending, which may for example be true for members of the usually relatively conservative middle classes. Then the error term, which includes son preference, will be positively correlated with the share of boys at any age, but also positively correlated with education expenditures, thereby leading to an upward bias of the coefficients for boys. Similarly, son preference will be negatively correlated with the share of daughters at any age, leading to a downward bias of the corresponding coefficients. As the test for gender discrimination involves testing the equality of the estimated coefficients for the same age group for boys and girls, the omitted variable bias will tend to overstate pro-male bias.

If we believe Jensen (2002), however, it is not clear that this is necessarily the correct story to tell even in the presence of son preference. If parents continue to have children until they get the desired number of sons, then son preference may actually be positively correlated

[4] See for example Rosenzweig and Wolpin (2000) for a more detailed discussion of the quantity-quality tradeoff and its empirical importance

with the share of girls, especially at older ages. As long as son preference is positively correlated with education expenditures, this would tend to bias the estimates on the share of girls of a particular age group upwards rather than downwards. Depending on the size of the bias relative to the upward bias of the male coefficients, this may either reduce the overstating of pro-male bias or may even lead to an understating of pro-male bias if the upward bias of the girl coefficients is large enough. This ambiguity of the bias of girl coefficients in the presence of son preference means that it is not clear whether our estimates of gender discrimination will tend to over- or understate true differentials in the treatment of girls and boys, even assuming that all households prefer sons over daughters. Furthermore, some households may not have a preference for the gender of their children or may even prefer girls.

Systematic misreporting of education expenditures is another important concern. As long as errors in education expenditures are uncorrelated with the gender composition of the children in the household, misreporting will not bias the estimated coefficients. Suppose, however, that households with girls will tend to overreport education expenditures in order to conceal that they are spending little on education for girls. In this case, the estimated coefficients will tend to understate true gender discrimination within the household.

One of the other big problems with using OLS to estimate the Engel curve is the fact that a significant number of households incur zero expenditure on education. Censoring is therefore an important concern, which will bias the OLS results. To account for the censoring of the data I will therefore follow Kingdon (2005) and use a hurdle model to analyse gender bias at the extensive and the intensive margin. The advantage of using a hurdle model is that, in contrast to the Tobit model, it does not assume that the same mechanism underlies the choice of whether to incur positive education expenditures or not, and the choice of how much to spend on education conditional on incurring positive spending. The hurdle model therefore proceeds in two steps (Wooldridge 2002). First, the decision of whether to spend money on education or not is modelled by estimating a probit model, with an indicator variable for a positive budget share of education expenditures as the dependent variable. The second step is to model the decision of how much to spend on education, conditional on having decided to spend any money at all. I do so by estimating a conditional OLS regression using the observations with positive budget shares, assuming that the budget share of education follows a log-normal distribution for positive budget shares. As graphs 1 and 2 show in the appendix, this is not a terrible assumption for either measure of education expenditures.

Given this setup, calculating the marginal effects of both the probit and the conditional OLS part is straightforward. The marginal effect of interest in the conditional OLS part is the change in expected budget share of education expenditures for the positive observations of this budget share for a change in variable x, which is given by the following equation:

$$\frac{\partial E(b \mid x, b > 0)}{\partial x} = \beta \exp(x\beta + \sigma^2 / 2) \qquad (2)$$

A consistent estimator for σ is the standard error of the conditional OLS regression.

4. Results

Table 3 presents the marginal effects for the OLS and hurdle model for the whole sample, using both education expenditure measures. As we can see, the signs of the coefficients for all child age groups go in the expected direction: Marginal effects for 5-19 year old children are positive, whereas they are negative for 0-4 year olds. This means that holding household size constant, an additional school-aged child increases the budget share of education expenditures, whereas a young child will reduce education expenditures. Given that the average budget share of education expenditures for a household is a bit over 0.04, the magnitudes of the effects are quite large, especially for older children. With the exception of 20-55 year old women, all coefficients are statistically significantly different from zero. The negative signs of the coefficients for adult men suggest that men spend less on education than women.

Table 3 also reports the test statistics for tests of the equality of the coefficients for males and females of the same age group, using F tests for the OLS and the probit model, and using a two mean sample comparison t-test for the conditional OLS model. The results suggest that there is severe discrimination against girls from age 10 onwards in all parts of the analysis. Gender bias for all age groups is particularly strong at the intensive margin, however. It is not quite clear how to interpret the result of a significant difference of boys and girls in the 0-4 year age group: Holding household size constant, household education expenditures decline less for a young girl than for a young boy[5].

[5] We might think that this gives us a pro-girl bias, as household school fee expenditures are affected less for a baby girl than for a baby boy, which could mean that more likely money is spent on educational expenditures for

Suggestive evidence of discrimination against girls is quite robust to changes in the measure of education expenditures. Coefficients for measure 2, which is the broader measure of school expenditures, tend to be higher than those for measure 1, which is intuitive.

Tables 4, 5 and 6 present the corresponding results for 16 big Indian states in a condensed form by age. Table 4 for example reports the difference of the coefficients for the share of 5-9 year old boys and girls, and the significance level of the F or t test for the equality of the two coefficients. As the difference was calculated by subtracting the female coefficient from the estimated male coefficient, positive numbers represent higher expenditure on boys, whereas negative numbers represent higher expenditure on girls. Tables 5 and 6 are constructed analogously. The actual estimated coefficients as well as the test statistics are reported in the appendix.

As for the whole sample, there is little evidence for pro-male bias for 5-9 year olds in the OLS estimation and on the extensive margin as reported in table 4. Andhra Pradesh is the only state that displays robust anti-girl discrimination under changes of the dependent variable in the OLS regression. Karnataka and Rajasthan have suggested discrimination at the extensive margin in one of the two specifications, but not in the other, whereas Punjab and Tamil Nadu have non-robust pro-girl bias.

Similar to the all-India level, a lot of discrimination seems to occur at the internal margin. With the exception of Maharashtra and Assam, the hypothesis of equal coefficients is rejected for all states at the 1% significance level and is robust to the two education expenditure measures used. In contrast to the aggregate level where the effect is anti-girl bias under both specifications, effects at the state level are more affected by the education expenditure measure. 10 out of 16 states display the same sign of the difference in both estimations. Andhra Pradesh, Haryana, Madhya Pradesh, Orissa, Rajasthan and Uttar Pradesh show evidence of anti-girl bias, whereas in Assam, Gujarat, Himachal Pradesh and Maharashtra there seems to be some evidence of favoring girls in education expenditures. For the remaining states, the estimated difference in the budget share of an increase in the share of boys and girls in this age group is positive with one measure of education expenditures, and negative with the other. This could be either due to bias in the estimators, especially where the difference between coefficients is small, or reflect different expenditure patterns for children in different states. Transportation, for example, is only included in measure 2, but not in

---

the small girl than for the small boy. Alternatively, this situation could indicate a pro-male bias: An additional 0-4 year old boy makes it less likely that there is positive expenditure on school fees, but this might be because the household now saves money for the later education for the baby boy and is thus more likely to stop school fee expenditures on the other (older and possibly female) children in the household.

measure 1. If school transportation is more costly for girls than for boys, then it is possible that we find significant pro-boy bias under measure 1 which does not include transportation, but do find suggested pro-girl bias with measure 2, i.e. the effect of transportation costs could dominate discrimination in other parts of education expenditure, for example in school fees. On the other hand, if parents spend more on parts included only in measure 2 just for boys, say by buying more expensive school uniforms, then the effect could go the other way, possibly turning a pro-girl bias into a pro-boy bias.

For 10-14 year olds, the results are reported in table 5. While the all-India OLS regression suggested pro-male bias in this age group, Punjab is the only state to display a highly statistically significant bias under both specifications. A similar situation arises for the probit models, where only Rajasthan seems to have robust gender bias. Again, most of the action seems to occur at the intensive margin. In contrast to the 5-9 year olds, there are only two states who switch the sign of the estimated bias, Himachal Pradesh and Tamil Nadu. 6 states have a consistently non-positive sign, 7 states a consistently non-negative sign, and Uttar Pradesh's difference in education expenditure effects in insignificant for both specifications.

For 15-19 year olds, the results are reported in table 6. The patterns are qualitatively similar to the situation for 10-14 year olds in that most results again occur at the intensive margin rather than at the extensive margin or in the OLS regression. However, more states now display statistically highly significant differences in the Engel curve approach, and all of them are anti-girl biases. At the intensive margin, most states now have consistently non-negative differences, the only exceptions being Haryana and West Bengal. Bihar is the only state to switch signs from one education expenditure measure to the other.

Overall, these results display interesting patterns at the state level. Gender bias in general seems to be much stronger at the internal margin than at the external margin, i.e. across most states parents do not choose to spend positive amounts on their sons but not on their daughters, but rather discriminate in the amount spent on their children. At every age group, the sign of the gender bias is robust to the use of different education expenditure measures for the majority of states, and the number of states that experience a change in the sign of the bias from one specification to the other declines with the age of the children. Together with increased action at the extensive margin and in the Engel curve approach for older ages, this seems to suggest that anti-girl discrimination is more pervasive the older the child. This is qualitatively the same conclusion that was reached at the aggregate all-India level.

5. Conclusion

The empirical analysis in my paper has shown evidence of gender discrimination in the intra-household allocation of education expenditures in India. The used nationally representative dataset allowed me to analyze gender bias both at the all-India level as well as at the state level. Similar to some of the previous literature on education expenditures, I used not only the Engel curve approach, but also a hurdle model to take account of the censoring of education expenditures. The richness of the dataset also allowed me to construct two different measures of household education expenditures which were used to test the robustness of the results.

In contrast to a number of previous papers, I find that gender bias in intra-household allocation is generally stronger and more widespread. Anti-girl discrimination seems to increase with age and becomes almost universal at the intensive margin for 15-19 year olds. This finding is quite robust to the level of analysis (all-India versus state level) and the measure of education expenditure used. While I find strong results with the Engel curve approach at the all-India level, most effects at the state level are only visible when using the hurdle model, as most discrimination at the state level seems to work through the intensive margin channel: Parents often spend less on education-related goods and services for their daughters than for their sons, rather than spending positive amounts on sons and nothing on their daughters. Results in this paper also suggest that even for the same sample of households, gender discrimination may be sensitive to the definition of the expenditure of the good in question, especially if analysis proceeds at a relatively disaggregated level like the state level and for age groups where discrimination is not almost universal.

One reason for these strong results in comparison to previous papers may lie in the quality of the data. The dataset used here collected detailed expenditure data on education-related goods and services both at the household level and at the individual level, which may have increased the reliability of the data in comparison to other datasets which just ask for one lump-sum number of education expenditure in the previous year. Asking about school materials and other costs such as transportation and uniforms may have reminded household respondents of their expenditures on these items which are most likely infrequent purchases and therefore more prone to recollection error.

References

Ahmad and Murdoch, Identifying Sex Bias in the Allocation of Household Resources: Evidence from Limited Household Surveys from Bangladesh, 1993

Asfaw, Abay, Klasen, Stephan and Lamanna, Francesca. Intra-Household Disparities in Children's Medical Care before Death in India. IZA DP No. 2586, 2007.

Aslam, Monazza and Kingdon, Geeta Gandhi. Gender and Household Education Expenditure in Pakistan. *Applied Economics*, 40(20):2573 - 2591, 2008.

Bhalotra, Sonia and Attfield, Cliff. Intrahousehold Resource Allocation in Rural Pakistan: A Semiparametric Analysis. *Journal of Applied Econometrics*, 13(5):463-480, 1998.

Case, Anne and Deaton, Angus. Consumption, Health, Gender, and Poverty. Working Paper no. 3020, World Bank Poverty Reduction and Economic Management Network, Washington, DC., 2003.

Deaton, Angus. Looking for Boy-Girl Discrimination in Household Expenditure Data. *The World Bank Economic Review*, 3(1):1-15, 1989.

Deaton, Angus. *The Analysis of Household Surveys*, Ch.4, 1997

Desai, Sonalde, Reeve Vanneman, and National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS), 2005 [Computer file]. ICPSR22626-v5. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-06-22. doi:10.3886/ICPSR22626

Fuwa, Nobuhiko, Ito, Seiro, Kubo, Kensuke, Kurosaki, Takashi, and Sawada, Yasuyuki. Gender Discrimination, Intrahousehold Resource Allocation, and Importance of Spouses' Fathers: Evidence on Household Expenditure from Rural India. *The Developing Economies*, XLIV(4):398-439, 2006.

Gibson, John. Testing for Boy-Girl Discrimination with Household Expenditure Data: Results for Papua New Guinea. *Applied Economics Letters*, 4(10):643-46, 1997.

Gibson, John and Rozelle, Scott. Is it Better to be a Boy? A Disaggregated Outlay Equivalent Analysis of Gender Bias in Papua New Guinea. *Journal of Development Studies*, 40(4):115-136, 2004.

Gong, Xiaodong, van Soest, Arthur and Zhang, Ping. Sexual Bias and Household Consumption: A Semiparametric Analysis of Engel Curves in Rural China. *Journal of Applied Econometrics*, 20: 509-527, 2005.

Haddad, Lawrence and Reardon, Thomas. Gender Bias in the Allocation of Resources within Households in Burkina Faso: a Disaggregated Outlay Equivalent Analysis. *Journal of Development Studies*, 29(2):260-276, 1993.

Himaz, Rozana. Intrahousehold Allocation of Education Expenditure and Returns to Education: The Case of Sri Lanka. University of Oxford Department of Economics Discussion Paper Series No.393, 2008.

Jensen, Robert. Equal Treatment, Unequal Outcomes? Generating Gender Inequality through Fertility Behavior. Unpublished manuscript, Department of Economics, Brown University, 2002.

Kingdon, Geeta Gandhi. Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure. *Economic Development and Cultural Change*, 53:409 - 451, 2005.

Kingdon, Geeta Gandhi. The Progress of School Education in India. Global Poverty Research Group WPS 071, 2007.

Klasen, Stephan. Nutrition, Health and Mortality in Sub-Saharan Africa: Is There a Gender Bias? *Journal of Development Studies,* 32(6):913-932, 1996.

Lancaster, Geoffrey, Maitra, Pushkar and Ray, Ranjan. Endogenous Power, Household Expenditure Patterns, and New Tests of Gender Bias: Evidence from India. Technical Report Discussion Paper no. 20/03, Department of Economics, Monash University, 2003.

Lee, Yiu-fai Daniel. Do Families Spend More on Boys than on Girls? Empirical Evidence from Rural China. *China Economic Review*, 19:80-100, 2008.

Liu, Jin-Long and Hsu, Ching-Chun. Economies of scale, gender discrimination, and cost of children. *Applied Economics Letters,* 11(6):377-382, 2004.

Messer, Ellen. Intra-Household Allocation of Food and Health Care: Current Findings and Understandings – Introduction. *Social Science and Medicine*, 44(11): 1675-1684, 1997.

Miller, Barbara D. Social Class, Gender and Intra-Household Food Allocations to Children in South Asia. *Social Science and Medicine*, 44(11): 1685-1695, 1997.

Murdoch, Jonathan J. and Stern, Hal S. Using mixture models to detect sex bias in health outcomes in Bangladesh. *Journal of Econometrics*, 77:259-276, 1997.

Rose, Elaina. Gender Bias, Credit Constraints and Time Allocation in Rural India. *The Economic Journal*, 110(465):738-758, 2000.

Rosenzweig, Mark R. and Wolpin, Kenneth I.. Natural "Natural Experiments" in Economics. *Journal of Economic Literature*, 38(4):827-874, 2000.

Sauerborn, R., Berman, P. and Nougtara, A.. Age bias, but no gender bias, in the intra-household resource allocation for health care in rural Burkina Faso. *Health Transition Review*, 6:131-145, 1996.

Subramaniam, Ramesh. Gender-Bias in India: The Importance of Household Fixed-Effects. *Oxford Economic Papers*, 48:280-299, 1996.

Subramanian, Shankar and Deaton, Angus. Gender Effects in Indian Consumption Patterns. Discussion Paper no.147, Research Program in Development Studies, Woodrow Wilson School, Princeton University, 1990.

Wooldridge, Jeffrey M. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press, 2002.

Table 1: Gender bias in school performance and enrolment for children between 8-11

	Girls	Boys	p-value of difference (two sided test)
Reading	2.57 (.02) N = 5832	2.67 (.02) N = 6598	0.0
Writing	.68 (.01) N = 5800	.71 (.01) N = 6523	0.0
Mathematics	1.51 (.01) N = 5812	1.64 (.01) N = 6569	0.0

Notes:
* significant at 10%, ** significant at 5%, *** significant at 1%
Reported are the means, with standard deviation in parenthesis; N is the number of observations
All school performance tests were available in English, Hindi, or the official state languages.
The reading level was scored as follows: 0 (cannot read), 1 (letter), 2 (word), 3 (paragraph), 4 (story). The writing level was scored as follows: 0 (cannot write), 1 (writes with two or fewer mistakes). The mathematics level was scored as follows: 0 (cannot), 1 (number), 2 (subtraction), 3 (division)

Table 2: Summary statistics

budget share education expenditure measure 1	.042 (.063)
budget share education expenditure measure 2	.049 (.069)
positive expenditure under measure 1	71.18
positive expenditure under measure 2	72.0
Muslim	12.27
Brahmin	5.49
Other Backward Castes (OBC)	39.40
Scheduled Castes (SC)	20.55
Scheduled Tribes (ST)	8.25
other castes	26.3
rural	65.52
below the poverty line	22.23
household size	5.8 (2.39)
household income	55157 (84401)
N	32263

Note: Standard deviations in parentheses; for dummy variables, percentages are reported
The used poverty line here is the official poverty line by the Indian Planning Commission for 2005, which is based on 1970s calculations of income needed to achieve minimal calorie consumption and has been adjusted by price indices. It varies by state and urban/rural areas.

Table 3 All-India results for variables of interest

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.039	-.299	-.092	-.038	-.326	-.099
	(.006)***	(.036)***	(.000)***	(.006)***	(.035)***	(.000)***
Share male 5-9	.058	1.016	.017	.074	1.021	.033
	(.006)***	(.036)***	(.000)***	(.006)	(.035)***	(.000)***
Share male 10-14	.092	1.054	.082	.118	1.077	.112
	(.005)***	(.036)***	(.000)***	(.006)***	(.035)***	(.001)***
Share male 15-19	.103	.467	.134	.125	.399	.164
	(.006)***	(.035)***	(.001)***	(.006)	(.034)***	(.001)***
Share male 20-55	-.027	-.233	-.019	-.037	-.275	-.032
	(.006)***	(.035)***	(.000)***	(.006)***	(.034)***	(.000)***
Share male over 55	-.031	-.220	-.020	-.037	-.268	-.019
	(.007)***	(.046)***	(.000)***	(.008)***	(.044)***	(.000)***
Share female 0-4	-.036	-.283	-.079	-0.35	-.281	-.091
	(.006)***	(.036)***	(.000)***	(.006)	(.035)***	(.000)***
Share female 5-9	.049	1.001	.007	.064	1.003	.022
	(.006)***	(.036)***	(.000)***	(.006)***	(.035)***	(.000)***
Share female 10-14	.077	.940	.072	.104	.986	.102
	(.005)***	(.036)***	(.000)***	(.006)***	(.036)***	(.000)***
Share female 15-19	.075	.300	.100	.098	.286	.128
	(.005)***	(.034)***	(.000)***	(.006)***	(.033)***	(.000)***
Share female 20-55	.003	.033	-.017	.004	-.004	-.015
	(.005)	(.032)	(.000)***	(.006)	(.031)	(.000)***
test statistic 0-4	0.41	0.33	-22.85***	0.25	2.82*	-12.82***
test statistic 5-9	3.81*	0.26	110.91***	4.97**	0.35	56.92***
test statistic 10-14	13.15***	13.70***	18.80***	10.54***	8.45***	13.68***
test statistic 15-19	45.38***	42.11***	40.71***	38.10***	20.74***	36.56***
test statistic 20-55	28.39***	63.28***	-18.01***	46.61***	71.44***	-7.39***
R squared/pseudo Rsquared	0.24	0.33	0.31	0.26	0.35	0.30
Log likelihood		-13017.5			-12360.8	
N	32263	32263	22957	32263	32263	23221

Note: test statistic gives F statistic for OLS and probit and t statistic for conditional OLS; significance levels: * 10%, ** 5%, *** 1%

Table 4: Difference of coefficient estimates for 5-9 year old boys and girls and significance level

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Andhra Pradesh	.032**	-.030	.032***	.052***	.012	.084***
Assam	-.008	.109	-.003*	-.002	.084	-.020***
Bihar	.003	.079	-.010***	-.015	.018	.009***
Gujarat	-.025	-.082	-.067***	-.014	-.117	-.065***
Haryana	.011	-.053	.032***	.018	-.078	.053***
Himachal Pradesh	.005	.007	-.013***	.017	-.203	-.003***
Karnataka	.023	.201**	.016***	.008	.053	-.021***
Kerala	-.012	.145	.001***	-.012	.194	-.004***
Maharashtra	-.006	-.122	-.001	-.014	-.119	-.033***
Madhya Pradesh	.024*	.091	.027***	.007	.091	.028***
Orissa	.015	-.074	.008***	.022	-.050	.023***
Punjab	.010	-.096	-.017***	.016	-.390*	.010***
Rajasthan	.013	.107	.034***	.024	.235**	.044***
Tamil Nadu	-.012	-.300**	-.007***	.012	-.178	.035***
Uttar Pradesh	.006	-.011	.010***	.013	-.005	.028***
West Bengal	.002	.072	-.012***	.007	.049	.001***
India	.008*	.015	.010***	.010**	.018	.011***

Note: significance levels: * 10%, ** 5%, *** 1%

Table 5: Difference of coefficient estimates for 10-14 year old boys and girls and significance level

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Andhra Pradesh	.017	.123	-.005***	.008	-.116	-.003
Assam	-.005	-.211	-.004	.049***	-.173	.034***
Bihar	.022	.222	.057***	.019	.285	.051***
Gujarat	.011	.254*	-.001	.001	.024	-.026***
Haryana	-.048**	.093	-.044***	-.042*	-.087	-.033***
Himachal Pradesh	-.021	-.426**	-.008**	.012***	-.146	.0143***
Karnataka	-.013	-.097	-.024***	.003	.112	-.009***
Kerala	.019	-.090	.012***	.057*	.338	.0671***
Maharashtra	.016	.116	-.004***	.007	.150	-.007***
Madhya Pradesh	.034***	.221*	.049***	.022*	.249**	.049***
Orissa	.010	.044	.011***	.020	-.082	.034***
Punjab	.068***	-.054	.102***	.097***	.047	.112***
Rajasthan	.048***	.413***	.050***	.019	.371***	.020***
Tamil Nadu	.002	.002	.014***	-.002	.098	-.024***
Uttar Pradesh	-.004	.091	.061	-.006	.071	.000
West Bengal	.022	.129	-.016***	.023	.105	-.023***
India	.015***	.115***	.010***	.014***	.091***	.010***

Note: significance levels: * 10%, ** 5%, *** 1%

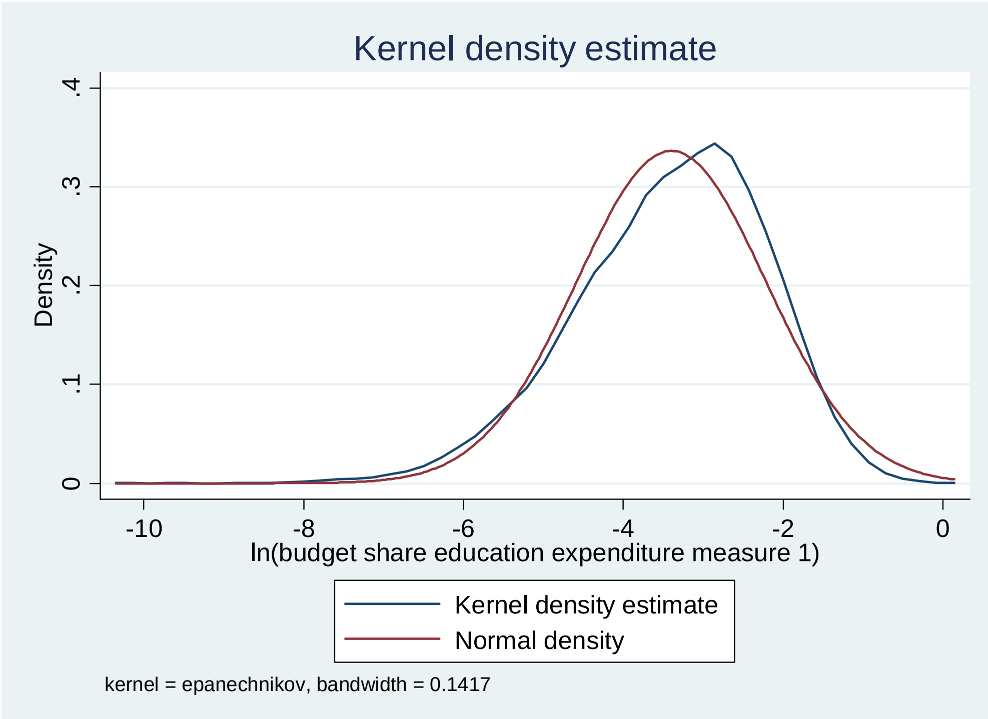
Table 6: Difference of coefficient estimates for 15-19 year old boys and girls and significance level

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Andhra Pradesh	.054***	.223**	.063***	.061***	.134	.098***
Assam	-.004	.085	.014***	-.020	-.131	.008***
Bihar	-.015	.455*	.007***	-.012	.580***	-.006**
Gujarat	.033**	.289**	.083***	.037*	.235**	.075***
Haryana	-.024	.057	-.011***	-.014	.037	-.0126***
Himachal Pradesh	.081	.100	.084***	.084***	-.017	.101***
Karnataka	.033**	.265***	.020***	.045***	.113	.059***
Kerala	.007	-.045	.017***	-.0201	.002	.001
Maharashtra	.035	.154*	.057***	.0476***	.035	.047***
Madhya Pradesh	.039***	.146	.054***	.035***	.156	.074***
Orissa	.031**	.146	.059***	-.008	-.016	.039***
Punjab	.028	.029	.020***	.046*	.013	.027***
Rajasthan	.019	.322***	.058***	.015	.388***	.054***
Tamil Nadu	.032	-.087	.041***	.031	.098	.031***
Uttar Pradesh	.028**	.167*	.038***	.044***	.132	.054***
West Bengal	.026	.1823*	-.023***	.022	.166*	-.016***
India	.028***	.167***	.036***	.028***	.113***	.036***

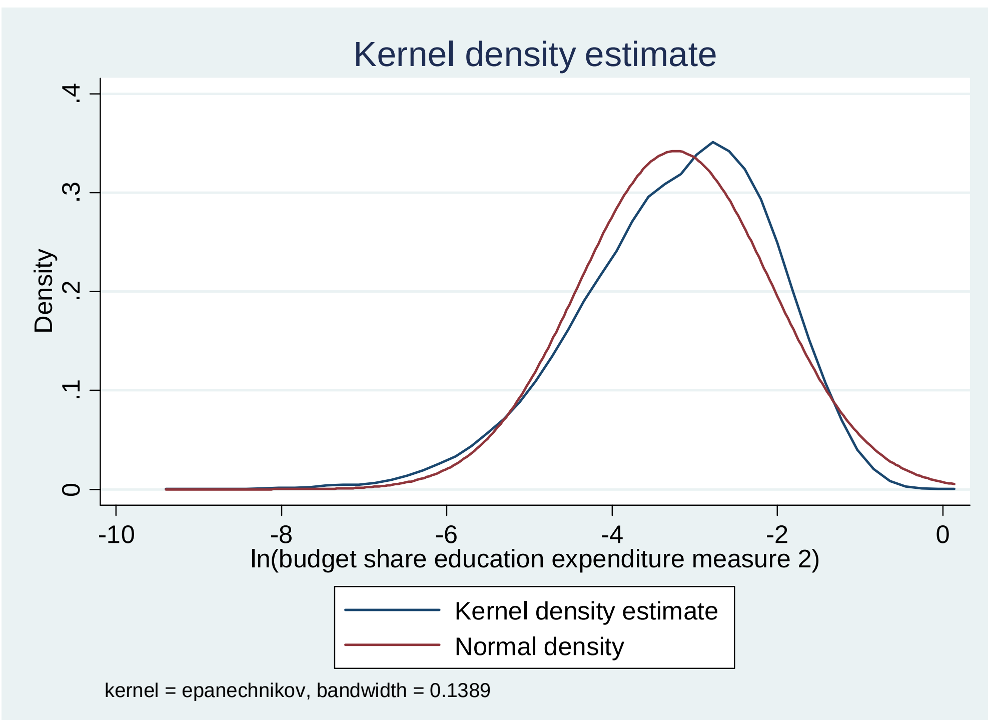
Note: significance levels: * 10%, ** 5%, *** 1%

Appendix

Graph 1: Kernel density estimate for education expenditure measure 1



Graph 2: Kernel density estimate for education expenditure measure 2



State-specific results for the parameters of interest (same comments apply as for all-India results)
Andhra Pradesh

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.058	-.394	-.102	-.049	-.588	-.092
	(.021)***	(.146)***	(.002)***	(.026)*	(.150)***	(.002)***
Share male 5-9	.042	1.069	.030	.080	.926	.086
	(.021)**	(.156)***	(.001)***	(.025)***	(.160)***	(.002)***
Share male 10-14	.043	.929	.030	.082	.631	.096
	(.021)**	(.153)***	(.001)***	(.025)***	(.154)***	(.002)***
Share male 15-19	.097	.253	.149	.134	.029	.260
	(.021)***	(.146)*	(.003)***	(.026)***	(.150)	(.005)***
Share male 20-55	-.042	-.290	-.020	-.051	-.427	-.036
	(.022)*	(.152)*	(.000)***	(.027)*	(.156)***	(.001)***
Share male over 55	-.058	-.327	-.072	-.063	-.583	-.026
	(.029)**	(.196)*	(.000)***	(.035)*	(.201)***	(.001)***
Share female 0-4	-.058	-.345	-.102	-.056	-.545	-.100
	(.021)***	(.147)**	(.002)***	(.026)**	(.151)***	(.002)***
Share female 5-9	.010	1.099	-.001	.028	.907	.002
	(.021)	(.158)***	(.000)***	(.025)	(.160)***	(.000)***
Share female 10-14	.026	.805	.035	.075	.747	.100
	(.020)	(.146)***	(.001)***	(.024)***	(.153)***	(.002)***
Share female 15-19	.043	.030	.087	.073	-.105	.162
	(.020)**	(.141)	(.002)***	(.025)***	(.145)	(.003)***
Share female 20-55	-.011	-.021	-.024	-.002	-.090	-.022
	(.019)	(.131)	(.001)***	(.023)	(.135)	(.000)***
test statistic 0-4	.00	.21	.00	0.15	0.15	2.84***
test statistic 5-9	4.62**	0.05	49.20***	8.23***	0.02	47.68***
test statistic 10-14	1.37	0.99	-4.90***	0.20	0.88	-1.39
test statistic 15-19	12.12***	4.86**	16.82***	10.43***	1.71	15.67***
test statistic 20-55	2.09	3.84**	5.40***	3.54*	5.73**	-16.17***
R squared/pseudo Rsquared	0.23	0.34	0.30	0.22	0.32	0.25
Log likelihood		-714.6			-740.9	
N	1817	1805	1306	1817	1817	1246

Assam

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	.004	-.008	-.007	.001	.052	-.159
	(.034)	(.388)	(.000)***	(.042)	(.378)	(.005)***
Share male 5-9	.056	1.483	.034	.054	1.499	-.100
	(.029)*	(.321)***	(.001)***	(.036)	(.312)***	(.003)***
Share male 10-14	.086	1.109	.080	.104	1.180	-.006
	(.029)***	(.328)***	(.003)***	(.036)***	(.320)***	(.000)***
Share male 15-19	.070	.760	.095	.097	.727	.048
	(.029)**	(.327)**	(.003)***	(.036)***	(.319)***	(.002)***
Share male 20-55	.035	.119	.048	.024	.056	-.050
	(.031)	(.347)	(.002)***	(.038)	(.338)	(.002)***
Share male over 55	.073	.511	.114	.054	.129	-.015
	(.038)*	(.417)	(.004)***	(.046)	(.405)	(.000)***
Share female 0-4	.030	.378	.037	.006	.367	-.142
	(.032)	(.354)	(.001)***	(.039)	(.346)	(.005)***
Share female 5-9	.064	1.374	.038	.056	1.415	-.080
	(.030)**	(.327)***	(.001)***	(.036)	(.320)***	(.003)***
Share female 10-14	.091	1.320	.084	.055	1.354	-.041
	(.029)***	(.329)***	(.003)***	(.036)	(.320)***	(.001)***
Share female 15-19	.074	.675	.081	.117	.858	.040
	(.029)**	(.323)**	(.003)***	(.036)***	(.314)***	(.001)***
Share female 20-55	.016	.357	.004	.024	.239	-.049
	(.027)	(.305)	(.000)***	(.036)	(.297)	(.002)***
test statistic 0-4	1.17	2.05	-36.96***	0.02	1.39	-2.53**
test statistic 5-9	0.27	0.33	-1.94*	0.01	0.20	-4.92***
test statistic 10-14	0.10	1.26	-0.95	6.76***	0.88	25.81***
test statistic 15-19	0.07	0.21	3.50***	0.95	0.50	4.11***
test statistic 20-55	0.76	0.94	28.50***	0.00	0.58	-0.43
R squared/pseudo Rsquared	0.25	0.23	0.31	0.24	0.25	0.38
Log likelihood		-359.7			-352.9	
N	743	743	500	743	743	467

Bihar

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.044	-.403	-.127	-.025	-.426	-.082
	(.029)	(.230)*	(.003)***	(.030)	(.219)*	(.002)***
Share male 5-9	.035	.512	.005	.038	.484	.028
	(.028)	(.225)***	(.000)***	(.029)	(.213)**	(.001)***
Share male 10-14	.055	.625	.069	.101	.822	.125
	(.028)**	(.222)***	(.002)***	(.029)***	(.211)***	(.003)***
Share male 15-19	.049	.376	.050	.067	.450	.068
	(.029)*	(.236)	(.001)***	(.030)**	(.225)**	(.002)***
Share male 20-55	-.009	-.472	-.021	-.047	-.519	-.058
	(.029)	(.229)**	(.001)***	(.030)	(.217)**	(.001)***
Share male over 55	.008	-.383	.020	-.007	-.482	-.005
	(.038)	(.309)	(.000)***	(.039)	(.292)*	(.000)***
Share female 0-4	-.029	-.508	-.022	-.023	-.493	-.023
	(.029)	(.230)**	(.001)***	(.030)	(.219)**	(.001)***
Share female 5-9	.030	.432	.015	.053	.466	.049
	(.028)	(.226)*	(.000)***	(.029)*	(.215)**	(.001)***
Share female 10-14	.032	.403	.012	.082	.537	.074
	(.028)	(.227)*	(.000)***	(.029)***	(.217)**	(.002)***
Share female 15-19	.064	-.079	.043	.080	-.130	.074
	(.028)**	(.230)	(.001)***	(.029)***	(.220)	(.002)***
Share female 20-55	-.005	-.186	-.026	.007	-.198	.004
	(.026)	(.206)	(.001)***	(.027)	(.195)	(.000)***
test statistic 0-4	0,56	0.49	-30.23***	0.02	0.21	-27.88***
test statistic 5-9	0,06	0.27	-23.49***	0.55	0.02	-14.82***
test statistic 10-14	1,5	2.13	30.17***	1.10	3.70*	14.07***
test statistic 15-19	0,5	6.40**	4.04***	0.31	11.21***	-2.50**
test statistic 20-55	0,02	1.95	5.88***	3.88**	2.62	-42.72***
R squared/pseudo Rsquared	0,28	0.27	0.30	0.31	0.32	0.26
Log likelihood		-511.2			-483.2	
N		1169	831	1169	1169	792

Gujarat

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.032	-.285	-.062	-.025	-.330	-.077
	(.024)	(.188)	(.001)***	(.027)	(.165)**	(.002)***
Share male 5-9	.016	.845	-.026	.062	.966	.035
	(.023)	(.180)***	(.001)***	(.026)**	(.162)***	(.001)***
Share male 10-14	.083	1.194	.083	.109	1.100	.136
	(.022)***	(.176)***	(.002)***	(.025)***	(.162)***	(.003)***
Share male 15-19	.058	.446	.120	.108	.286	.197
	(.023)**	(.178)**	(.003)***	(.026)***	(.158)*	(.004)***
Share male 20-55	.000	-.307	.050	-.032	-.418	.064
	(.024)	(.189)	(.001)***	(.028)	(.165)**	(.001)***
Share male over 55	-.009	-.131	.046	.003	-.220	.119
	(.033)	(.255)	(.000)***	(.037)	(.225)	(.002)***
Share female 0-4	-.033	-.370	-.032	-.025	-.403	.011
	(.023)	(.185)**	(.001)***	(.026)	(.162)**	(.000)***
Share female 5-9	.041	.927	.041	.076	1.083	.010
	(.024)*	(.182)***	(.001)***	(.027)***	(.168)***	(.002)***
Share female 10-14	.072	.940	.084	.109	1.077	.162
	(.023)***	(.182)***	(.002)***	(.026)***	(.172)***	(.003)***
Share female 15-19	.025	.157	.036	.071	.051	.122
	(.023)	(.178)	(.001)***	(.026)***	(.158)	(.003)***
Share female 20-55	-.020	-.264	-.008	-.012	-.250	.005
	(.021)	(.166)	(.000)***	(.024)	(.145)*	(.000)***
test statistic 0-4	0.00	0.34	-18.54***	0.00	0.34	-54.36***
test statistic 5-9	2.00	0.35	-59.84***	0.54	0.73	-29.31***
test statistic 10-14	0.40	3.18*	-0.50	0.00	0.03	-5.98***
test statistic 15-19	3.96**	5.19**	28.97***	3.82*	4.65**	15.49***
test statistic 20-55	0.72	0.06	50.06***	0.59	1.20	44.39***
R squared/pseudo Rsquared	0.22	0.31	0.31	0.27	0.41	0.30
Log likelihood		-748.5			-612.3	
N	1602	1602	925	1602	1602	885

Haryana

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.045	-.420	-.086	-.038	-.371	-.072
	(.032)	(.202)**	(.001)***	(.033)	(.177)**	(.001)***
Share male 5-9	.076	.774	.041	.092	.902	.070
	(.029)***	(.191)***	(.001)***	(.030)***	(.174)***	(.001)***
Share male 10-14	.099	.907	.073	.126	1.055	.106
	(.028)***	(.188)***	(.001)***	(.030)***	(.173)***	(.002)***
Share male 15-19	.088	.351	.108	.124	.408	.142
	(.029)***	(.186)*	(.002)***	(.030)***	(.163)**	(.003)***
Share male 20-55	-.012	-.274	-.016	-.018	-.166	.013
	(.030)	(.194)	(.000)***	(.032)	(.171)	(.000)***
Share male over 55	.044	-.168	.003	.016	.019	.005
	(.040)	(.263)	(.000)***	(.042)	(.236)	(.000)***
Share female 0-4	-.050	-.405	-.134	-.026	-.245	-.062
	(.031)	(.199)**	(.002)***	(.033)	(.174)	(.001)***
Share female 5-9	.065	.827	.008	.074	.980	.017
	(.030)**	(.199)***	(.000)***	(.032)**	(.179)***	(.000)***
Share female 10-14	.146	.814	.116	.167	1.142	.140
	(.029)***	(.195)***	(.002)***	(.030)***	(.182)***	(.002)***
Share female 15-19	.112	.294	.119	.138	.370	.155
	(.029)***	(.188)	(.002)***	(.030)***	(.166)**	(.003)***
Share female 20-55	.019	.083	-.037	.016	.197	-.028
	(.027)	(.171)	(.001)***	(.029)	(.151)	(.001)***
test statistic 0-4	0.03	0.01	17.74***	0.17	0.79	-5.68***
test statistic 5-9	0.24	0.11	45.46***	0.55	0.29	41.55***
test statistic 10-14	4.96**	0.36	-18.51***	3.43*	0.31	-10.66***
test statistic 15-19	1.21	0.17	-4.06***	0.38	0.09	-3.38***
test statistic 20-55	0.99	3.54*	29.54***	1.11	5.06**	74.36***
R squared/pseudo Rsquared	0.22	0.26	0.27	0.26	0.37	0.25
Log likelihood		-618.4			-502.6	
N	1393	1393	987	1393	1393	963

Himachal Pradesh

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.090	-.124	-.167	-.095	-.165	-.164
	(.040)**	(.154)	(.003)***	(.043)**	(.145)	(.003)***
Share male 5-9	.085	1.262	.027	.095	1.070	.018
	(.038)**	(.175)***	(.000)***	(.041)**	(.154)***	(.000)***
Share male 10-14	.119	.792	.127	.172	1.018	.132
	(.038)***	(.164)***	(.002)***	(.040)***	(.179)***	(.002)***
Share male 15-19	.231	.739	.264	.249	.741	.259
	(.039)***	(.153)***	(.005)***	(.041)***	(.145)***	(.005)***
Share male 20-55	-.029	-.012	-.006	-.020	-.046	-.026
	(.039)	(.148)	(.000)***	(.041)	(.138)	(.000)***
Share male over 55	-.018	.088	.030	-.020	.046	.004
	.(047)	(.187)	(.001)***	(.051)	(.173)	(.000)***
Share female 0-4	-.075	-.158	-.092	-.067	-.075	-.086
	(.040)*	(.153)	(.002)***	(.043)	(.143)	(.001)***
Share female 5-9	.080	1.256	.039	.078	1.273	.021
	(.040)**	(.186)***	(.001)***	(.043)*	(.177)***	(.000)***
Share female 10-14	.140	1.219	.135	.160	1.164	.118
	(.037)***	(.203)***	(.003)***	(.040)***	(.196)***	(.002)***
Share female 15-19	.150	.638	.180	.164	.758	.158
	(.037)***	(.151)***	(.003)***	(.040)***	(.148)***	(.003)***
Share female 20-55	.017	.104	-.024	.014	.065	-.053
	(.035)	(.133)	(.000)***	(.038)	(.124)	(.001)***
test statistic 0-4	0.19	0.08	-21.37***	0.60	0.71	-24.13***
test statistic 5-9	0.02	0.00	-14.41***	0.21	1.33	-5.20***
test statistic 10-14	0.47	4.04**	-2.19**	0.13	0.43	4.65***
test statistic 15-19	7.42***	0.70	14.09***	6.92***	0.02	19.07***
test statistic 20-55	1.25	0.61	38.30***	0.62	0.66	26.30***
R squared/pseudo Rsquared	0.27	0.49	0.30	0.28	0.54	0.29
Log likelihood		-275.6			-244.8	
N	1040	1040	817	1040	1040	806

Karnataka

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.063	-.463	-.083	-.063	-.526	-.117
	(.019)***	(.110)***	(.001)***	(.021)***	(.117)***	(.002)***
Share male 5-9	.054	1.144	-.022	.068	1.059	-.004
	(.018)***	(.115)***	(.000)***	(.020)***	(.117)***	(.000)***
Share male 10-14	.049	1.185	.021	.093	1.237	.057
	(.018)***	(.112)***	(.000)***	(.019)***	(.117)***	(.001)***
Share male 15-19	.110	.570	.129	.138	.419	.198
	(.018)***	(.103)***	(.002)***	(.019)***	(.108)***	(.003)***
Share male 20-55	-.043	-.226	-.030	-.048	-.317	-.052
	(.018)**	(.102)**	(.000)***	(.019)**	(.107)***	(.001)***
Share male over 55	-.037	-.145	-.059	-.039	-.209	-.054
	(.022)	(.129)	(.001)***	(.024)	(.135)	(.001)***
Share female 0-4	-.063	-.515	-.085	-.056	-.479	-.101
	(.019)***	(.111)***	(.001)***	(.020)***	(.116)***	(.001)***
Share female 5-9	.031	.943	-.038	.061	1.006	.017
	(.018)*	(.109)***	(.001)***	(.020)***	(.114)***	(.000)***
Share female 10-14	.062	1.281	.045	.090	1.125	.066
	(.018)***	(.118)***	(.001)***	(.019)***	(.117)***	(.001)***
Share female 15-19	.077	.305	.110	.093	.306	.139
	(.017)***	(.103)***	(.002)***	(.019)***	(.108)***	(.002)***
Share female 20-55	.027	-.024	.007	.001	-.043	-.030
	(.017)	(.096)	(.000)***	(.018)	(.100)	(.000)***
test statistic 0-4	0.00	0.30	12.8	0.16	0.23	-7.19***
test statistic 5-9	1.29	4.14**	24.67***	0.24	0.28	-88.25***
test statistic 10-14	0.83	0.76	-32.56***	0.03	1.08	-7.63***
test statistic 15-19	6.05**	12.13***	7.93***	9.69***	2.06	17.69***
test statistic 20-55	16.14***	4.46**	-83.07***	6.80***	7.41***	-26.77***
R squared/pseudo Rsquared	0.20	0.3857	0.25	0.20	0.36	0.22
Log likelihood		-1139.8			-1221.6	
N	3008	3008	2084	3008	3008	1923

Kerala

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.068	-.222	-.105	-.074	-.201	-.131
	(.043)	(.179)	(.002)***	(.040)*	(.159)	(.002)***
Share male 5-9	.0772	1.473	.016	.101	1.472	.028
	(.037)**	(.160)***	(.000)***	(.034)***	(.143)***	(.000)***
Share male 10-14	.134	1.479	.127	.211	2.235	.183
	(.037)***	(.168)***	(.002)***	(.034)***	(.262)***	(.003)***
Share male 15-19	.178	1.136	.170	.205	.948	.220
	(.036)***	(.155)***	(.003)***	(.033)***	(.138)***	(.003)***
Share male 20-55	.037	.073	.058	-.002	-.025	-.008
	(.035)	(.140)	(.001)***	(.032)	(.121)	(.000)***
Share male over 55	.024	.086	.022	-.010	-.096	-.018
	(.044)	(.182)	(.000)***	(.041)	(.160)	(.000)***
Share female 0-4	-.047	-.036	-.059	-.049	-.101	-.077
	(.044)	(.179)	(.001)***	(.040)	(.155)	(.001)***
Share female 5-9	.089	1.329	.016	.113	1.278	.033
	(.040)**	(.169)***	(.000)***	(.036)***	(.145)***	(.000)***
Share female 10-14	.115	1.568	.115	.154	1.897	.117
	(.035)***	(.174)***	(.002)***	(.033)***	(.245)***	(.002)***
Share female 15-19	.171	1.181	.153	.225	.946	.219
	(.038)***	(.161)***	(.002)***	(.035)***	(.142)***	(.003)***
Share female 20-55	.014	.187	.001	.009	.065	.005
	(.033)	(.134)	(.000)***	(.031)	(.114)	(.000)***
test statistic 0-4	0.27	1.27	-25.62***	0.46	0.54	-25.03***
test statistic 5-9	0.11	0.84	2.72***	0.14	1.72	-7.11***
test statistic 10-14	0.34	0.23	4.69***	3.64*	0.98	21.68***
test statistic 15-19	0.06	0.13	5.00***	0.56	0.00	0.13
test statistic 20-55	0.36	0.58	65.38***	0.09	0.53	-97.31***
R squared/pseudo Rsquared	0.17	0.39	0.16	0.23	0.52	0.22
Log likelihood		-421.9			-319.7	
N	1152	1152	823	1152	1152	800

Maharashtra

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.024	-.675	-.077	-.035	-.611	-.078
	(.014)*	(.120)***	(.001)***	(.015)**	(.114)***	(.001)***
Share male 5-9	.020	.910	-.025	.021	.873	-.020
	(.013)	(.117)***	(.000)***	(.015)	(.113)***	(.000)***
Share male 10-14	.059	1.113	.037	.075	1.322	.057
	(.013)***	(.119)***	(.001)***	(.014)***	(.122)***	(.001)***
Share male 15-19	.089	.509	.098	.104	.409	.117
	(.013)***	(.111)***	(.002)***	(.015)***	(.106)***	(.002)***
Share male 20-55	-.025	-.369	-.018	-.042	-.376	-.032
	(.013)*	(.111)***	(.000)***	(.015)***	(.105)***	(.000)***
Share male over 55	-.036	-.402	-.038	-.051	-.307	-.060
	(.018)**	(.147)***	(.001)***	(.019)***	(.140)**	(.001)***
Share female 0-4	-.015	-.399	-.052	-.023	-.348	-.068
	(.014)	(.117)***	(.001)***	(.015)	(.112)***	(.001)***
Share female 5-9	.026	1.031	-.025	.034	.992	.013
	(.013)**	(.115)***	(.000)***	(.015)**	(.109)***	(.000)***
Share female 10-14	.044	1.000	.041	.068	1.172	.063
	(.013)***	(.120)***	(.001)***	(.014)***	(.122)***	(.001)***
Share female 15-19	.054	.355	.070	.056	.374	.070
	(.013)***	(.113)***	(.001)***	(.014)***	(.109)***	(.001)***
Share female 20-55	-.008	-.080	-.016	-.015	-.057	-.025
	(.013)	(.104)	(.000)***	(.014)	(.099)	(.000)***
test statistic 0-4	0.71	7.48***	-15.33***	0.84	7.74***	-6.60***
test statistic 5-9	0.31	1.40	-1.21	1.20	1.46	-93.92***
test statistic 10-14	1.33	1.11	-3.74***	0.44	1.48	-5.23***
test statistic 15-19	11.64***	3.29*	12.95***	17.79***	0.19	23.26***
test statistic 20-55	1.48	6.42**	-6.38***	3.04*	8.93***	-12.46***
R squared/pseudo Rsquared	0.23	0.43	0.29	0.24	0.47	0.27
Log likelihood		-915.7			-842.5	
N	2536	2536	1693	2536	2536	1641

Madhya Pradesh

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.020	-.304	-.074	-.021	-.171	-.082
	(.018)	(.133)**	(.001)***	(.017)	(.131)	(.001)***
Share male 5-9	.051	1.120	.011	.061	1.235	.032
	(.018)***	(.130)***	(.000)***	(.016)***	(.129)***	(.001)***
Share male 10-14	.096	1.166	.070	.100	1.288	.079
	(.018)***	(.133)***	(.001)***	(.017)***	(.131)***	(.001)***
Share male 15-19	.091	.391	.119	.095	.477	.122
	(.018)***	(.131)***	(.002)***	(.017)***	(.129)***	(.002)***
Share male 20-55	-.026	-.133	-.051	-.022	-.015	-.064
	(.019)	(.137)	(.001)***	(.018)	(.135)	(.001)***
Share male over 55	-.035	-.037	-.056	-.017	.130	-.075
	(.025)	(.177)	(.001)***	(.023)	(.175)	(.001)***
Share female 0-4	-.025	-.392	-.074	.025	-.311	-.087
	(.018)	(.133)***	(.001)***	(.0174)	(.131)**	(.001)***
Share female 5-9	.027	1.030	-.016	.054	1.144	.004
	(.018)	(.132)***	(.000)***	(.017)***	(.130)***	(.000)***
Share female 10-14	.061	.946	.021	.078	1.039	.030
	(.018)***	(.134)***	(.000)***	(.017)***	(.133)***	(.001)***
Share female 15-19	.052	.244	.065	.060	.321	.048
	(.018)***	(.131)*	(.001)***	(.017)***	(.128)**	(.001)***
Share female 20-55	.019	.223	.005	.033	.240	.014
	(.017)	(.119)*	(.000)***	(.017)**	(.117)**	(.000)***
test statistic 0-4	0.15	0.78	0.10	0.07	2.02	2.53**
test statistic 5-9	3.33*	0.71	77.21***	0.32	0.72	51.06***
test statistic 10-14	6.74***	3.59*	37.29***	3.07*	4.53**	34.24***
test statistic 15-19	7.51***	1.93	22.21***	6.81***	2.28	33.32***
test statistic 20-55	5.15**	6.42**	-60.83***	8.69***	3.48*	-70.50***
R squared/pseudo Rsquared	0.21	0.36	0.30	0.22	0.37	0.28
Log likelihood		-875.8			-861.0	
N	2306	2306	1656	2306	2306	1606

Orissa

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.044	-.314	-.080	-.031	-.308	-.059
	(.018)**	(.138)**	(.002)***	(.020)	(.144)**	(.001)***
Share male 5-9	.056	1.226	.030	.078	1.249	.074
	(.017)***	(.136)***	(.001)***	(.020)***	(.139)***	(.001)***
Share male 10-14	.081	1.046	.087	.121	1.037	.148
	(.017)***	(.134)***	(.002)***	(.020)***	(.137)***	(.003)***
Share male 15-19	.068	.366	.114	.050	.222	.129
	(.018)***	(.136)***	(.002)***	(.021)**	(.141)	(.002)***
Share male 20-55	.007	-.257	.043	-.004	-.210	.038
	(.018)	(.135)*	(.001)***	(.021)	(.140)	(.001)***
Share male over 55	-.027	-.288	.021	-.015	-.292	.035
	(.022)	(.165)*	(.000)***	(.025)	(.175)*	(.001)***
Share female 0-4	-.031	-.206	-.078	-.013	-.188	-.013
	(.019)*	(.145)	(.002)***	(.022)	(.150)	(.000)***
Share female 5-9	.041	1.301	.022	.055	1.299	.051
	(.018)**	(.140)***	(.000)***	(.020)***	(.145)***	(.001)***
Share female 10-14	.071	1.002	.076	.101	1.119	.114
	(.018)***	(.145)***	(.001)***	(.020)***	(.150)***	(.002)***
Share female 15-19	.038	.220	.055	.058	.238	.089
	(.017)**	(.135)	(.001)***	(.020)***	(.140)*	(.002)***
Share female 20-55	-.003	.147	-.046	.023	.016	.011
	(.016)	(.121)	(.001)***	(.018)	(.125)	(.000)***
test statistic 0-4	0.75	0.93	-1.03	1.13	1.05	-41.66***
test statistic 5-9	1.07	0.34	11.01***	1.87	0.15	14.12***
test statistic 10-14	0.48	0.12	4.95***	1.53	0.40	9.94***
test statistic 15-19	4.74**	2.11	24.16***	0.22	0.02	13.82***
test statistic 20-55	0.32	10.12***	73.14***	1.75	2.99*	37.29***
R squared/pseudo Rsquared	0.25	0.47	0.30	0.24	0.44	0.28
Log likelihood		-531.3			-572.7	
N	1626	1626	1134	1626	1626	1078

Punjab

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.043	-.408	-.048	-.043	-.384	-.071
	(.034)	(.191)**	(.001)***	(.036)	(.156)**	(.001)***
Share male 5-9	.113	.781	.127	.146	.850	.113
	(.034)***	(.195)***	(.002)***	(.036)***	(.173)***	(.002)***
Share male 10-14	.154	.506	.216	.224	.777	.228
	(.033)***	(.189)***	(.003)***	(.035)***	(.172)***	(.004)***
Share male 15-19	.098	.118	.161	.156	.123	.170
	(.032)***	(.180)	(.002)***	(.034)***	(.149)	(.003)***
Share male 20-55	-.057	-.397	-.022	-.068	-.350	-.073
	(.034)*	(.185)**	(.000)***	(.036)*	(.153)**	(.001)***
Share male over 55	-.047	-.308	-.013	-.081	-.363	-.090
	(.043)	(.241)	(.000)	(.046)*	(.197)*	(.001)***
Share female 0-4	-.040	-.454	-.013	-.042	-.307	-.067
	(.037)	(.205)**	(.000)***	(.040)	(.165)*	(.001)***
Share female 5-9	.103	.876	.144	.130	1.240	.103
	(.035)***	(.205)***	(.002)***	(.038)***	(.197)***	(.002)***
Share female 10-14	.086	.560	.114	.126	.730	.115
	(.032)***	(.188)***	(.002)***	(.034)***	(.168)***	(.002)***
Share female 15-19	.070	.089	.142	.111	.110	.143
	(.034)**	(.189)	(.002)***	(.036)***	(.157)	(.002)***
Share female 20-55	-.039	-.265	-.005	-.057	-.340	-.092
	(.031)	(.171)	(.000)***	(.033)*	(.137)**	(.001)***
test statistic 0-4	0.01	0.08	-45.49***	0.00	0.35	-2.17**
test statistic 5-9	0.12	0.25	-5.83***	0.23	3.82*	4.08***
test statistic 10-14	7.20***	0.11	27.73***	13.01***	0.08	27.31***
test statistic 15-19	1.24	0.04	6.01***	2.83*	0.01	7.52***
test statistic 20-55	0.27	0.51	-50.64***	0.09	0.00	9.88***
R squared/pseudo Rsquared	0.23	0.24	0.24	0.33	0.39	0.31
Log likelihood		-587.7			-422.7	
N	1323	1319	961	1323	1319	946

Rajasthan

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.053	-.239	-.121	-.066	-.222	-.119
	(.024)**	(.155)	(.002)***	(.026)**	(.149)	(.002)***
Share male 5-9	.0350	.892	.022	.053	.955	-.017
	(.022)	(.144)***	(.000)***	(.024)**	(.139)***	(.000)***
Share male 10-14	.077	1.057	.072	.085	1.100	.028
	(.022)***	(.148)***	(.001)***	(.024)***	(.143)***	(.000)***
Share male 15-19	.061	.424	.103	.062	.448	.056
	(.022)***	(.144)***	(.002)***	(.024)***	(.139)***	(.001)***
Share male 20-55	-.064	-.293	-.093	-.082	-.312	-.147
	(.024)***	(.155)*	(.002)***	(.026)***	(.149)**	(.002)***
Share male over 55	-.065	-.149	-.063	-.080	-.191	-.152
	(.033)**	(.212)	(.001)***	(.035)**	(.204)	(.002)***
Share female 0-4	-.052	-.156	-.109	-.054	-.101	-.137
	(.023)**	(.152)	(.002)***	(.025)**	(.145)	(.002)***
Share female 5-9	.022	.785	-.012	.029	.720	-.061
	(.022)	(.144)***	(.000)***	(.024)	(.139)***	(.001)***
Share female 10-14	.030	.644	.021	.066	.729	.009
	(.023)	(.152)***	(.000)***	(.025)***	(.148)***	(.000)***
Share female 15-19	.042	.102	.045	.047	.060	.002
	(.022)*	(.145)	(.001)***	(.024)*	(.140)	(.000)***
Share female 20-55	-.029	.142	-.050	-.028	.023	-.106
	(.021)	(.138)	(.001)***	(.023)	(.132)	(.002)***
test statistic 0-4	0.00	0.49	-3.98***	0.38	1.18	6.39***
test statistic 5-9	0.62	0.87	76.86***	1.75	4.48**	44.77***
test statistic 10-14	7.47***	10.26***	38.18***	1.00	8.55***	42.62***
test statistic 15-19	1.14	8.26***	29.32***	0.63	12.97***	62.51***
test statistic 20-55	2.04	7.77***	-23.13***	4.22**	5.11**	-14.35***
R squared/pseudo Rsquared	0.24	0.29	0.27	0.22	0.32	0.23
Log likelihood		-829.7			-780.2	
N	1990	1990	1448	1990	1990	1408

Tamil Nadu

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.038	.048	-.102	-.055	-.039	-.165
	(.034)	(.146)	(.002)***	(.033)	(.152)	(.003)***
Share male 5-9	.057	1.108	-.010	.076	1.096	.012
	(.031)*	(.146)***	(.000)***	(.031)**	(.151)***	(.000)***
Share male 10-14	.071	1.403	.049	.082	1.327	.034
	(.031)**	(.161)***	(.001)***	(.031)***	(.164)***	(.001)***
Share male 15-19	.135	.493	.150	.138	.475	.136
	(.033)***	(.143)***	(.003)***	(.032)***	(.148)***	(.002)***
Share male 20-55	.026	-.113	.062	.006	-.211	.049
	(.034)	(.144)	(.001)***	(.033)	(.149)	(.001)***
Share male over 55	.002	-.223	.007	-.007	-.300	.009
	(.043)	(.185)	(.000)***	(.042)	(.189)	(.000)***
Share female 0-4	-.034	.073	-.098	-.048	-.150	-.127
	(.034)	(.151)	(.002)***	(.034)	(.156)	(.002)***
Share female 5-9	.068	1.404	-.002	.065	1.274	-.022
	(.033)**	(.167)***	(.000)***	(.033)*	(.169)***	(.000)***
Share female 10-14	.060	1.401	.035	.084	1.229	.058
	(.031)*	(.157)***	(.001)***	(.031)**	(.155)***	(.001)***
Share female 15-19	.103	.581	.110	.108	.475	.105
	(.032)***	(.144)***	(.002)***	(.035)***	(.149)***	(.002)***
Share female 20-55	-.022	.037	-.032	-.025	.026	-.040
	(.032)	(.136)	(.001)***	(.032)	(.142)	(.001)***
test statistic 0-4	0.02	1.27	-1.57	0.05	0.96	-10.54***
test statistic 5-9	0.22	3.88**	-36.77***	0.22	1.41	79.35***
test statistic 10-14	0.22	0.00	11.92***	0.01	0.44	-20.75***
test statistic 15-19	1.91	0.73	10.96***	1.74	0.00	10.67***
test statistic 20-55	1.95	1.17	67.55***	0.83	1.69	82.28***
R squared/pseudo Rsquared	0.18	0.42	0.24	0.19	0.40	0.22
Log likelihood		-457.4			-474.7	
N	1367	1367	1005	1367	1367	971

Uttar Pradesh

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.023	-.093	-.041	-.000	-.136	-.022
	(.018)	(.120)	(.001)***	(.020)	(.117)	(.000)***
Share male 5-9	.054	.929	.049	.080	.907	.085
	(.017)***	(.116)***	(.001)***	(.019)***	(.114)***	(.001)***
Share male 10-14	.090	.912	.101	.106	.886	.135
	(.017)***	(.119)***	(.002)***	(.019)***	(.118)***	(.002)***
Share male 15-19	.086	.403	.130	.135	.370	.180
	(.018)***	(.121)***	(.002)***	(.020)***	(.119)***	(.003)***
Share male 20-55	-.042	-.212	-.026	-.041	-.264	-.021
	(.019)**	(.121)*	(.000)***	(.020)**	(.119)**	(.000)***
Share male over 55	-.039	-.098	.012	-.028	-.141	.031
	(.025)	(.166)	(.000)***	(.028)	(.163)	(.001)***
Share female 0-4	-.017	-.001	-.039	-.005	-.020	-.027
	(.018)	(.118)	(.001)***	(.020)	(.116)	(.000)***
Share female 5-9	.049	.940	.039	.067	.912	.057
	(.017)***	(.118)***	(.001)***	(.019)***	(.116)***	(.001)***
Share female 10-14	.094	.820	.098	.112	.815	.134
	(.017)***	(.121)***	(.002)***	(.019)***	(.119)***	(.002)***
Share female 15-19	.058	.236	.092	.091	.238	.126
	(.018)***	(.116)**	(.001)***	(.019)***	(.115)**	(.002)***
Share female 20-55	-.008	.090	-.003	.007	.076	.006
	(.016)	(.105)	(.000)***	(.018)	(.102)	(.000)***
test statistic 0-4	0.20	1.36	-2.23**	0.11	1.21	7.76***
test statistic 5-9	0.22	0.01	10.10***	0.89	0.00	16.46***
test statistic 10-14	0.09	0.87	0.97	0.22	0.53	0.06
test statistic 15-19	4.33**	3.58*	15.15***	9.08***	2.33	14.75***
test statistic 20-55	4.03**	7.16***	-57.03***	6.25**	9.51***	-74.61***
R squared/pseudo Rsquared	0.31	0.33	0.33	0.29	0.34	0.33
Log likelihood		-1065.0			-1036.9	
N	2898	2898	2209	2898	2898	2182

West Bengal

	Measure 1			Measure 2		
	OLS	Probit	Conditional OLS	OLS	Probit	Conditional OLS
Share male 0-4	-.048	-.191	-.241	-.025	-.109	-.223
	(.025)*	(.142)	(.004)***	(.024)	(.143)	(.004)***
Share male 5-9	.049	1.089	-.034	.080	1.171	.006
	(.024)**	(.147)***	(.001)***	(.023)***	(.147)***	(.000)***
Share male 10-14	.141	.907	.131	.170	.978	.175
	(.024)***	(.144)***	(.002)***	(.024)***	(.145)***	(.003)***
Share male 15-19	.126	.312	.138	.153	.395	.183
	(.024)***	(.139)**	(.003)***	(.024)***	(.139)***	(.003)***
Share male 20-55	-.020	-.288	-.012	.005	-.206	.020
	(.024)	(.139)**	(.000)***	(.024)	(.140)	(.000)***
Share male over 55	-.065	-.509	-.037	-.032	-.388	.000
	(.029)**	(.166)***	(.001)***	(.029)	(.167)**	(.015)
Share female 0-4	-.052	-.314	-.216	-.031	-.240	-.201
	(.025)**	(.141)**	(.004)***	(.025)	(.142)*	(.004)***
Share female 5-9	.047	1.017	-.021	.073	1.123	.005
	(.024)*	(.153)***	(.000)***	(.024)***	(.154)***	(.000)***
Share female 10-14	.120	.778	.148	.147	.874	.198
	(.024)***	(.144)***	(.003)***	(.024)***	(.145)***	(.004)***
Share female 15-19	.100	.130	.161	.131	.230	.198
	(.023)***	(.132)	(.003)***	(.023)***	(.132)*	(.004)***
Share female 20-55	.031	.054	-.031	.064	.161	.013
	(.023)	(.127)	(.001)***	(.023)***	(.128)	(.000)***
test statistic 0-4	0.04	1.30	-4.3***	0.09	1.44	-3.94***
test statistic 5-9	0.01	0.28	-16.66***	0.14	0.13	8.71***
test statistic 10-14	1.59	1.23	-4.51***	1.80	0.79	-4.59***
test statistic 15-19	2.33	3.50*	-5.90***	1.62	2.81*	-3.14***
test statistic 20-55	4.88**	6.90***	31.73***	6.49**	7.75***	16.79***
R squared/pseudo Rsquared	0.32	0.36	0.40	0.32	0.34	0.39
Log likelihood		-669.2			-683.3	
N	1804	1804	1327	1804	1804	1314