The Determinants of Household Level Fertility in India

Raghbendra Jha
Australian National University

ABSTRACT

Using NSS data for 1993-94 and 2004-05 this paper highlights the impact of growing incomes, social and household decisions of households, and regional and ethnic factors on patterns of household level fertility in India. These have helped determine the composition of India's young (aged 9 to 34) today. Demographic transition is well underway in India with rising incomes associated with fewer children and smaller family size. The number of women in the child-bearing age group significantly affects the number of children. Households with more women in the age group 26-35 have more children, are more likely to have children than not having them as well as having larger family size, *ceteris paribus*. Average education of females lowers household size whereas (instrumented) shares of expenditure on education and health have varying effects. The impact of a household being SC or ST varies by year and by the regression model chosen. Over both time periods Muslim households have more children and are more likely than the general population to have larger family sizes. Households in BIMARU states have more children and have larger family sizes as do urban households. Thus demographic transition has occurred unevenly across various groups in India.

Keywords: Gender Bias, Census, National Sample Survey, Demographic Transition, India
JEL Classification Code: J16, N35, O15, O53

Address all correspondence to:
Prof. Raghbendra Jha, ASARC,
Arndt–Corden Dept of Economics,
H.C. Coombs Building (09)
Australian National University,
Canberra, ACT 0200, Australia
Phone: + 61 2 6125 2683, Fax: + 61 2 6125 0443,
Email: r.jha@anu.edu.au

I. Introduction

When the results of the 2011 Census of India were announced two factors were most noticed: (i) a reduction in the total fertility rate from 2.9 in 2001 to 2.62 in 2011, and (ii) deterioration in the gender balance, i.e., the number of girls per 1000 boys between the censuses of 2001 and 2011. For 0-4, 5-9, and 0-6 year olds this fell from 939 to 891, 920 to 889, and 927 to 914 respectively.

While the first effect is usually taken as an indicator of the demographic transition associated with rising per capita incomes, the second is often cited as evidence of widespread gender bias in the Indian population. Considerable evidence (Jha et.al.. 2011) exists of sex selection tests and follow-up abortions if the fetus is found to be female. However, since the Census does not include household level characteristics, identifying such characteristics that increase the chance of feticide is difficult with this database. Chaudhri and Jha (2013) used household level data for the National Sample Survey (NSS) rounds of 1993-94 and 2004-05[1] to identify characteristics of households determining the number of girl children relative to boys and found, ironically enough, that higher education of mothers and higher prosperity are each associated with increasing gender bias. Only when the product of the two reaches a relatively high level does gender bias start coming down.

In this paper I ask a question complementary to that in Chaudhri and Jha (2013), i.e., what are the determinants of the number of children in the household, i.e., the household's fertility? In particular I investigate, for the same data set as in Chaudhri and Jha (2013), the determinants of the number of children (aged 0-14) at the household level.

An additional contribution of this paper is that it provides a profile of India's young from those aged 9 (born in 2004-05) to those who are 34 (those who were 14 in 1993-94). The plan of this paper is as follows. Section II discusses the data and empirical approach. Section III elaborates on the results and section IV concludes from a broad policy perspective.

II. Data and Methodology

The National Sample Survey is one of the largest and most comprehensive household/enterprise surveys conducted anywhere in the world. The 1st to the 10th rounds of the NSS were the formative years when many concepts of data collection and sampling were tested. The 28th Round of the survey conducted between October 1973 and June 1974 under the auspices of the newly formed National Sample Survey Organization (NSSO) marked a watershed in the development of the NSS. Variations in data collection methods have occurred almost routinely but a substantial degree of homogeneity has been retained for the Consumer survey as well as the Employment and Unemployment survey.

The basic design of sampling flowed is a stratified two-stage (for small villages and urban blocks). For larger villages and blocks and sub-block formation a three stage sampling procedure is used. Regardless of this difference in stratification the ultimate stage unit is the household or enterprise.

Consistent with Chaudhri and Jha (2013) we use NSS data for the 50th Round (1993-94) and 61st Round (2004-05) and model the determinants of the number of children aged 0-

[1] The intervening sample survey of 1999-00 could not be used because of well known problems of the consumption data for this round of the NSS with other rounds.

14 at the household level. Three approaches are taken. In the first an OLS regression with robust standard errors is run to explain the number of children. The second approach involves using a probit model to explain whether a household has at least one child. Finally, in a multinomial logit analysis I model the determinants of one, two, three, and four or more children for the household in comparison to the state of having no children.

The Probit model The dependent variable, $Y_{i,}$ in our case is a dichotomous variable taking the value 1 if household *i* has at least one child, 0 otherwise. In the probit model used in this paper it is postulated that the realization Y_i depends on a vector **X** of household and other characteristics. Any cumulative probability model meets the requirement $Y_i = P(\eta_i) = P(\alpha + \beta X_i)$ (1)

where *P* is the Cumulative Distribution Function pre-selected in advance, and α and β are parameters to be estimated. Since *P* is strictly increasing it is possible to invert the function in (1) to write: $P^{-1}(Y_i) = \eta_i = \alpha + \beta X_i$ (2)

where P^{-1} is the inverse of the CDF, i.e., the quantile function. Thus, we have a linear mode for a transformation of Y_i or, equivalently, a non-linear model for Y_i. The transformation P(.) is often chosen as the CDF of the unit normal distribution

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-(0.5)z^2} dz \qquad (3)$$

where π=3.141 and e=2.718 are the usual constants. When

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-(0.5)z^2} dz \qquad = \Phi(\alpha + \beta X_i) \quad (4)$$

we get the linear probit model. It is well-known (Greens, 2003) that marginal effects (defined below in the context of the multinomial logit model) are more meaningful than the coefficients themselves.

The Multinomial Logit Model

The multinomial unordered logit model for household type *j* is

$$\Gamma[Y_i = j] = \frac{e^{\beta'_j x_i}}{\sum_{k=1}^{4} e^{\beta'_k x_i}}, j = 0,1,2,3,4 \text{...}(5)$$

where $j = 0,1,2,3,4$ refers to type of household (Greene, 2003) since we believe the dependent variable is unordered. The estimated equations yield a set of probabilities for $j+1$ choices for a decision maker with characteristics x_i. Out of five choices, only four e parameter vectors are needed to determine all the four probabilities. The probabilities are given by

$$\Gamma[Y_i = j/x_i] = \frac{e^{\beta'_j x_i}}{1 + \sum_{k=1}^{J} e^{\beta'_k x_i}}, for \quad j = 0,1,.....J, \beta_0 = 0 \text{.....................................}(6)$$

In our case Type 0 = no children is the omitted or reference group. Other groups are one child, two children, three children and four or more children in the household. Further, β coefficients in this model are difficult to interpret, therefore, we compute marginal effects as

$$\delta_j = \frac{\partial \Gamma[Y_i = j]}{\partial x_i} = \Gamma[Y_i = j][\beta_j - \bar{\beta}]; j = 1,2,3,4 \text{...}(7)$$

Thus every sub-vector of β enters every marginal effect, both through the probabilities and through the weighted average that appears in δ_j. Standard errors are computed using the delta method.

The variables used in this paper are as follows. For the OLS model the dependent variable is the number of children aged 0-14 in the household, for the probit model the dependent variable is whether the household has at least one child aged 0-14, whereas for the multinomial logit model the dependent variables are whether (compared to having no children aged 0-14) the household has one, two, three or four or more children aged 0-14. The independent variables for all the models are number of females in child-bearing age group, i.e., aged 15-25, 26-35 and 36-49, instrumented values of number of girls aged 5-14 in the household who are not going to school,[2] average education (number of years of schooling) of females in the household, log of real monthly per capita expenditure of the household (2004-05 values deflated to 1993-94 values), instrumented share of education in total expenditure[3] of the household, instrumented share of health in total expenditure of the household, a dummy for a Scheduled Caste (SC) household, a dummy for a Scheduled Tribe (ST) household, a dummy for a Muslim household, a dummy for the BIMARU states of Bihar, Madhya Pradesh, Rajasthan and Uttar Pradesh and a dummy for rural households.

[2] In the probit model this variable is not significant and is dropped.

[3] It might be argued that the number of girls not in school, health and education expenditures could be endogenous to the number of children in the household. Young girls, it may be argued, may be needed to look after their younger siblings in the households, whereas health and, particularly, education expenditures could depend on the number of children in the household. Hence, I used instrumented values of these variables in the regressions. Results of the first stage regression are not included here to conserve space. The results were not very different from those when these variables are included directly in the regressions. In particular, the signs of the coefficients and their significance were unaltered. This indicates that the results are robust to alternative specifications.

III. Results

In Table 1 I present results on OLS estimation with robust standard errors of the number of children in the household.

Table 1 about here.

Three sets of results are presented – one for 1993-94, the second for 2004-05 whereas the final one presents differences in the coefficients for the respective variables across the two years and the significance of these differences. The key results are as follows. In both years the number of children is significantly higher the greater the number of females in the childbearing age (15-49) in the household. Indeed this effect peaks with the age group 26-35 of potential mothers and is lower for younger women (15-25) and older women (36-49). The (instrumented) number of girls (5-14) not in school is also a strongly significant determinant of the number of children. Average education of adult females in the household, computed as the total number of years of schooling of these women divided by the number of such women, lowers the number of children in the household. Similarly, the higher the log of Monthly Per Capita Income (MPCE) (figures for 2004-05 are deflated to make them comparable with 1993-94 figures) the lower is the number of children. (Instrumented) shares of health and education in the household budget have a positive impact. Scheduled Caste households have fewer children in both years whereas Scheduled Tribe households had fewer children in 1993-94 but more children in 2004-05. Muslim households had more children in both years with the coefficients for this dummy weighty and strongly significant for both years. Households in the BIMARU states had more children than those in non-BIMARU states whereas rural households had fewer children than urban households.

It is interesting also to note the differences in coefficients across the two time periods. Compared to 1993-94 women in the child-bearing ages had fewer children in 2004-05 with the drop being highest for women in the age group 26-35. The impact of (instrumented) number of girls not in school went up over time as did that of the (instrumented) shares of education and health expenditures. Higher MPCE led to a further drop in the number of children. SC and ST households had significantly higher number of children as did households in the BIMARU states. The change in the number of children in Muslim households was insignificant. The impact of average education of females went up over time.

Similar results are found for the probit model modeling whether the household in question has at least one child aged 0 to 14. The numbers of women in child bearing age group 15-25, 26-35, and 36-49 are each strongly significant. The second age group has the largest such marginal effect. Higher education of females in the household leads to a drop in the probability of having a child. Log(MPCE) has negative and significant marginal effects in both years. (Instrumented) shares of education and health have positive and significant marginal effects in both years. Compared to the general population marginal effects for SC and ST households were negative and significant. The marginal effects for Muslim households were positive and significant for both years and the dummy for rural households was negative and significant for both years. Differences between the marginal effects for both years are negative and significant for number of females aged 15-25, 26-35 and 36-49, log (MPCE) and rural household dummy. This difference is positive and significant for education share of total expenditure, health share of total expenditure, SC, ST households, Muslim households, BIMARU dummy and rural households. The results are shown in Table 2.

Table 2 here.

Further, insight is provided by the results on the estimation of the multinomial logit model. In both years the higher number of females in the 15-25 age group in the household the higher the likelihood of having one child and the lower the likelihood of having two or three children; however, the probability of having four or more children is higher in both years. The higher number of females aged 26-35 the lower the probabilities of having one child in both years, of having two children in 1993-94 whereas the probabilities of having two children in 2004-05, three children in both years, and four more or more children in both years are all higher. The higher the number females aged 36-49 the lower the probabilities of having one or two children and the higher the probabilities of having three or four or more children in both years. The probability of having one child was positively associated with the number of girls (5-14) not in school in 1993-94 but positively associated in 2004-05. The probabilities of having two, three, four or more children are positively affected by the (instrumented) number of girls between 5 and 14 not in school in both years. The higher the education of females in the household the higher were the chances of having one or two children in both years. Households with less educated women ended up having three or four or more children in both 1993-94 and 2004-05. The higher the (instrumented) share of education in a household's expenditure the lower are the chances of having one child in both years and of having two children in 1993-94. In 2004-05 this effect was positive. In both years the higher the share of education in total expenditures the greater are the chances of having three or four or more children. A similar pattern holds for (instrumented) health expenditure. The fact that a household had SC background increased its chance of having one child in 2004-05, reduced its chances of having two children in both years, and its chances of having three or four or more children in 1993-94. The fact that a household had ST background lowered the probability of it having one child in 2004-05, of having two children in 2004-05, of having three children in 1993-94 and of having four or more children in 1993-94. ST households had greater probability of having two children in 1993-94, of having three children in 2004-05 and of having four or more children in 2004-05. In both years Muslim households had lower probabilities of having one or two children and higher probabilities of having three or four or more children. In both years households in BIMARU states had lower probabilities of having one or two children and higher probabilities of having three or four or more children. A rural household had higher probabilities of having one child in both years and four or more children in 2004-05. In all other cases the impact of this variable was negative.

Across the two years differences between the responses are positive and significant for number of females aged 15-25 in the case of one child, two children, and three children and negative and significant for four or more children. For women aged 26-35 the difference between the two years is positive and significant for one child and three children and negative and significant for four or more children. Differences for females aged 36-49 were positive and significant for two children and negative and significant for four or more children. Differences for number of girls aged 5-14 and not in school were positive and significant for two children and for three children. Differences for average education level of females were positive and significant for one child and three children and negative and significant for two children. Differences for log(MPCE) were positive and significant for one child and negative and significant for two children and three children. Differences for (instrumented) education share of total expenditure are positive and significant for two children, three children and four or more children and negative and significant for one child. Differences for (instrumented) health share of total expenditure are positive and significant for one child, two children, three children and four or more children. Differences for SC households are positive and significant for three children and four or more children. Differences for ST households are positive and significant for three children and four or more children and negative and significant for one child and two children. Differences for Muslim households were negative and significant for one child and four or more children, and positive and significant for two children, and three children. Differences for BIMARU states were positive and significant for three children and four or more children and negative and significant for one child and two children. Differences for rural households were positive and significant for two children and negative and significant for one child and positive and significant for two children.

Results for the multinomial model are shown in Table 3.

Table 3 about here.

IV. Conclusions

This paper has highlighted the impact of growing incomes, social and household decision factors of households and regional factors in determining the patterns of household level fertility in 1993-

94 and 2004-05 and, hence, in determining the composition of India's youth (aged 9 to 34) today. Demographic transition is well underway in India with rising incomes associated with fewer children and smaller family size. As expected the number of women in the child-bearing age group has a significant effect on the number of children. This effect is particularly strong for women aged 26-35. Households with more women in this age group have more children, are more likely to have children than not having them as well as having larger family size, *ceteris paribus*. Average education of females lowers household size whereas (instrumented) shares of expenditure on education and health have varying effects. The impact of a household being SC or ST varies by year and by the regression model chosen. Over both time periods Muslim households have more children and are more likely than the general population to have larger family sizes. Households in BIMARU states have more children and have larger family sizes as do urban households. The full demographic transition is, therefore, yet to set in over Muslim households and those in BIMARU states.

References

Chaudhri, D.P. and R. Jha (2013) "India's Gender Bias in Child Population, female education and growing prosperity" *International Review of Applied Economics*, vol. 27, no.1, pp. 23-43.

Greene, W. H (2003), *Econometric Analysis*. 5th Edition, Delhi, Pearson Education, India.

Jha, P., Kesler, M., Kumar, R., Ram, F., Ram, U., Aleksandrowicz, L., Bassani, D., Chandra, S. and J. Banthia (2011) "Trends in selective abortions of girls in India: analysis of nationally representative birth histories from 1990 to 2005 and census data from 1991 to 2011" *Lancet,* vol. 377, Issue 9781,

Table 1 - OLS model, Robust SE
(Dependent variable is number of children aged 0 to 14)

	1993-94	2004-05	Coef. difference between 2005-1993
No. of females aged 15-25	0.245***	0.150***	-0.095***
	(0.00812)	(0.00701)	
No. of females aged 26-35	1.329***	1.244***	-0.085***
	(0.0133)	(0.0114)	
No. of females aged 36-49	0.401***	0.344***	-0.057***
	(0.0126)	(0.00967)	
INumber of girls (5-14) not in school	0.953***	0.990***	0.037***
	(0.00885)	(0.0121)	
Average education level of females	-0.0548***	-0.0362***	0.02***
	(0.00152)	(0.00131)	
Log(MPCE)	-0.490***	-0.718***	-0.22***
	(0.00848)	(0.00896)	
IEducation share of total expenditure	0.0306***	0.0420***	0.012***
	(0.000796)	(0.000823)	
IHealth share of total expenditure	0.00297***	0.00989***	0.007***
	(0.000278)	(0.000422)	
Scheduled caste	-0.106***	-0.0180*	0.09***
	(0.0110)	(0.00989)	
Scheduled tribe	-0.0579***	0.0917***	0.141***
	(0.0134)	(0.0118)	
Muslim	0.357***	0.368***	0.011
	(0.0147)	(0.0127)	
Bihar, MP, UP or Rajasthan	0.224***	0.387***	0.163***
	(0.00889)	(0.00875)	
Dummy = 1 if rural	-0.0321***	-0.0514***	-0.02*
	(0.00888)	(0.00789)	
Constant	3.576***	5.293***	
	(0.0533)	(0.0607)	
Observations	108429	118548	
F_statistics	4103.1	4401.7	
p_value	0.000	0.000	
Adjusted R2	0.384	0.393	

Standard errors in parentheses
* $p <= 0.10$, ** $p <= 0.05$, *** $p <= 0.01$

Table 2 - Probit model, Robust SE
(Dependent variable = 1 if HH has at least one child aged 0 to 14)

	1993-94	2004-05	Coef. difference between 2005-1993
No. of females aged 15-25	0.228***	0.149***	-0.079***
	(0.00808)	(0.00742)	
No. of females aged 26-35	1.192***	1.164***	-0.028*
	(0.0154)	(0.0155)	
No. of females aged 36-49	0.195***	0.152***	-0.043***
	(0.0111)	(0.0105)	
Average education level of females	-0.0276***	-0.00717***	0.02***
	(0.00191)	(0.00177)	
Log(MPCE)	-0.488***	-0.724***	-0.24***
	(0.0102)	(0.0126)	
IEducation share of total expenditure	0.0262***	0.0399***	0.013***
	(0.00105)	(0.00110)	
IHealth share of total expenditure	0.00203***	0.00755***	0.005***
	(0.000345)	(0.000542)	
Scheduled caste	-0.0724***	-0.0389***	0.04*
	(0.0137)	(0.0121)	
Scheduled tribe	-0.0333**	0.0507***	0.08***
	(0.0160)	(0.0140)	
Muslim	0.162***	0.219***	0.057***
	(0.0167)	(0.0146)	
Bihar, MP, UP or Rajasthan	0.0814***	0.173***	0.092***
	(0.0105)	(0.0102)	
Dummy = 1 if rural	-0.0538***	-0.0928***	-0.04***
	(0.0106)	(0.00994)	
Observations	93903	109818	203721
chi2	10798.2	13821.6	24445.1
p_value	0.000	0.000	0.000
Pseudo R2	0.159	0.191	0.175

Marginal effects; Standard errors in parentheses
* $p <= 0.10$, ** $p <= 0.05$, *** $p <= 0.01$

Table 3 - Multinomial logit model, Robust SE

(Dependent variable = 0 if HH has no child aged 0 to 14, = 1 if HH has 1 child aged 0 to 14, = 2 if HH has 2 children aged 0 to 14, = 3 if HH has 3 children aged 0 to 14, = 4 if HH has more than 3 children aged 0 to 14)

	1993-94	2004-05	Coef. difference between 2005-1993
Dependent variable = 1 versus 0			
No. of females aged 15-25	0.003***	0.027***	0.024***
	(0.0151)	(0.0135)	
No. of females aged 26-35	-0.255***	-0.209***	0.05**
	(0.0295)	(0.0277)	
No. of females aged 36-49	-0.047***	-0.028***	0.02
	(0.0215)	(0.0196)	
INumber of girls (5-14) not in school	-0.145	0.192	0.33
	(.)	(.)	
Average education level of females	0.019*	0.018***	0.001***
	(0.00369)	(0.00323)	
Log(MPCE)	0.133***	0.180***	0.05***
	(0.0195)	(0.0211)	
IEducation share of total expenditure	-0.009***	-0.012***	-0.01***
	(0.00225)	(0.00231)	
IHealth share of total expenditure	-0.001***	-0.002***	0.001***
	(0.000696)	(0.00108)	
Scheduled caste	0.035	0.011*	-0.02
	(0.0273)	(0.0245)	
Scheduled tribe	0.015	-0.016***	-0.03**
	(0.0316)	(0.0273)	
Muslim	-0.058***	-0.060***	-0.01*
	(0.0341)	(0.0298)	
Bihar, MP, UP or Rajasthan	-0.031*	-0.048***	-0.01***
	(0.0211)	(0.0211)	
Dummy = 1 if rural	0.016**	0.007***	-0.01***
	(0.0210)	(0.0193)	
Dependent variable = 2 versus 0			
No. of females aged 15-25	-0.012***	-0.016***	0.00***
	(0.0160)	(0.0149)	
No. of females aged 26-35	-0.014***	0.056***	0.07
	(0.0292)	(0.0284)	
No. of females aged 36-49	-0.051***	-0.057***	0.006***
	(0.0227)	(0.0213)	
INumber of girls (5-14) not in school	0.046***	0.399***	0.35***
	(0.0279)	(0.0372)	
Average education level of females	0.006***	0.001***	-0.01***
	(0.00393)	(0.00344)	
Log(MPCE)	0.026***	-0.018***	-0.04***
	(0.0215)	(0.0230)	
IEducation share of total expenditure	-0.000***	0.002***	0.002***
	(0.00231)	(0.00235)	
IHealth share of total expenditure	-0.000***	0.000***	0.000***
	(0.000698)	(0.00112)	

Scheduled caste	-0.011***	-0.027***	-0.01
	(0.0277)	(0.0249)	
Scheduled tribe	0.002*	-0.027**	-0.02**
	(0.0322)	(0.0287)	
Muslim	-0.046***	-0.035***	0.01**
	(0.0341)	(0.0300)	
Bihar, MP, UP or Rajasthan	-0.028***	-0.047***	-0.02**
	(0.0213)	(0.0212)	
Dummy = 1 if rural	-0.002***	-0.009***	0.007**
	(0.0215)	(0.0200)	
Dependent variable = 3 versus 0			
No. of females aged 15-25	-0.014***	-0.011***	0.003***
	(0.0189)	(0.0187)	
No. of females aged 26-35	0.126***	0.138***	0.01*
	(0.0323)	(0.0326)	
No. of females aged 36-49	0.024***	0.037***	0.01
	(0.0267)	(0.0265)	
INumber of girls (5-14) not in school	0.125***	0.272***	0.15**
	(0.0273)	(0.0362)	
Average education level of females	-0.013***	-0.011***	0.002***
	(0.00492)	(0.00449)	
Log(MPCE)	-0.087***	-0.129***	-0.04***
	(0.0260)	(0.0301)	
IEducation share of total expenditure	0.005***	0.008***	0.003***
	(0.00257)	(0.00270)	
IHealth share of total expenditure	0.000***	0.002***	0.002***
	(0.000797)	(0.00141)	
Scheduled caste	-0.014***	0.009	0.02***
	(0.0312)	(0.0292)	
Scheduled tribe	-0.012***	0.034***	0.04***
	(0.0372)	(0.0346)	
Muslim	0.019***	0.046***	0.03***
	(0.0374)	(0.0343)	
Bihar, MP, UP or Rajasthan	0.014***	0.045***	0.03***
	(0.0239)	(0.0240)	
Dummy = 1 if rural	-0.014***	-0.006***	0.01
	(0.0251)	(0.0246)	
Dependent variable $\geq$ 4 versus 0			
No. of females aged 15-25	0.026***	0.014***	-0.01***
	(0.0212)	(0.0213)	
No. of females aged 26-35	0.159***	0.114***	-0.05***
	(0.0367)	(0.0378)	
No. of females aged 36-49	0.077***	0.063***	-0.01***
	(0.0312)	(0.0318)	
INumber of girls (5-14) not in school	0.126***	0.139***	0.01
	(0.0269)	(0.0355)	
Average education level of females	-0.013***	-0.010***	0.003
	(0.00624)	(0.00599)	
Log(MPCE)	-0.078***	-0.095***	-0.02
	(0.0314)	(0.0392)	
IEducation share of total expenditure	0.005***	0.005***	0.000***

	(0.00281)	(0.00310)	
IHealth share of total expenditure	0.001***	0.001***	0.000***
	(0.000882)	(0.00162)	
Scheduled caste	-0.012***	0.004	0.01***
	(0.0358)	(0.0341)	
Scheduled tribe	-0.005**	0.015***	0.02***
	(0.0439)	(0.0424)	
Muslim	0.086***	0.065***	-0.02***
	(0.0395)	(0.0375)	
Bihar, MP, UP or Rajasthan	0.046***	0.062***	0.02***
	(0.0267)	(0.0270)	
Dummy = 1 if rural	-0.000***	0.000***	0.000
	(0.0297)	(0.0297)	
Observations	108429	118548	
Pseudo R2	0.155	0.169	

Standard errors in parentheses
* $p <= 0.10$, ** $p <= 0.05$, *** $p <= 0.01$