Does Higher Economic Development Reduce Household Size? Evidence from India

Tripathi, Sabyasachi

Adamas University

6 January 2018

Does Higher Economic Development Reduce Household Size? Evidence from India

Dr. Sabyasachi Tripathi
Assistant Professor of Economics
Adamas University
Kolkata - 700126, West Bengal, India
Email: sabya.tripathi@gmail.com
Phone No: +91-9748076984

Abstract:
The paper investigates the impact of higher economic development on average household size in India. The analysis finds that increasing income measured by net per capita state domestic product and per capita consumption expenditure has a negative effect on average household size. Variables such as, higher level education, health outcomes, extent of inequality and urbanization has negative effect on the average household size. Lower level of poverty is associated with lower level family size in long run, whereas, infrastructure has a mix effect. Results show that different religious and social groups have an effect on family size in India. Smaller family size faces several problems such as child rearing, depression, separation, anxiety and land distribution dispute. The paper argues that the issue needs to bring into current development policies for changing social structure, land distribution process, and helping for maintaining appropriate balance between work and family which are missing currently.

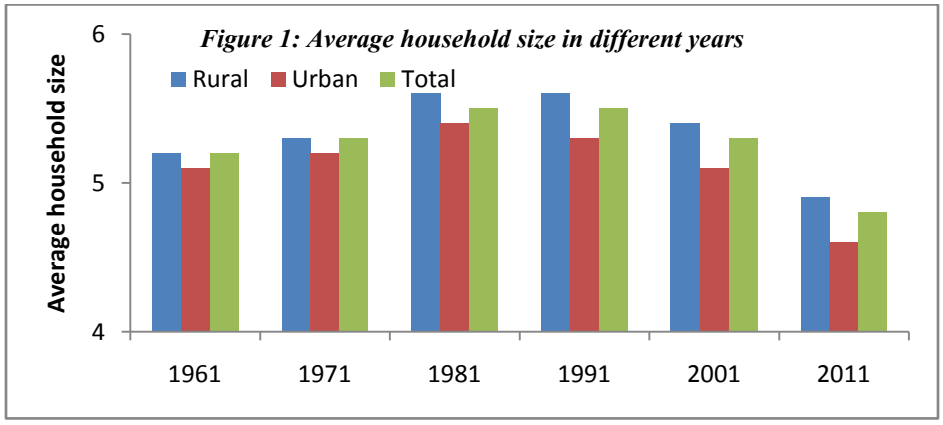
Key wards: Household size, Economic development, India

JEL Classification: D10, R20, O10

1. Introduction

The world is experiencing a reduction in average family size with higher economic development. For instance, the average household size in Australia (or France) has fallen from 4.5 (or 3.1) in 1911 (or 1968) to 2.53 (or 2.3) in 2006 (or 2011). The literature (e. g., Adams, 2010; Bongaarts, 2001) argued that the declining in family size is a consequence of increasing urbanization, industrialization, and educational expansion which are all part of higher development. While Indian economy witnessed an average growth rate of 7.4% per year between 2001 and 2011 through a transformation from agricultural based economy to industrialized/serviced based urban economy (Tripathi, 2013, Tripathi and Mahey, 2017), the average family size is declining from nuclear to single. While the total number of Indian households increased by 29 percent, the single person households have increased by about 35 percent between 2001 and 2011. Though, the number of nuclear families has increased from 135 million in 2001 to 172 million in 2011 but the percentage has remained almost same at 70% which is quite higher. This could be the reason of inadequacy of housing availability in urban areas or working couple choosing to live with the other family members for a support. Census data also shows that 56% of households in urban India had four or less members in 2011. This indicates that family size in India is declining.

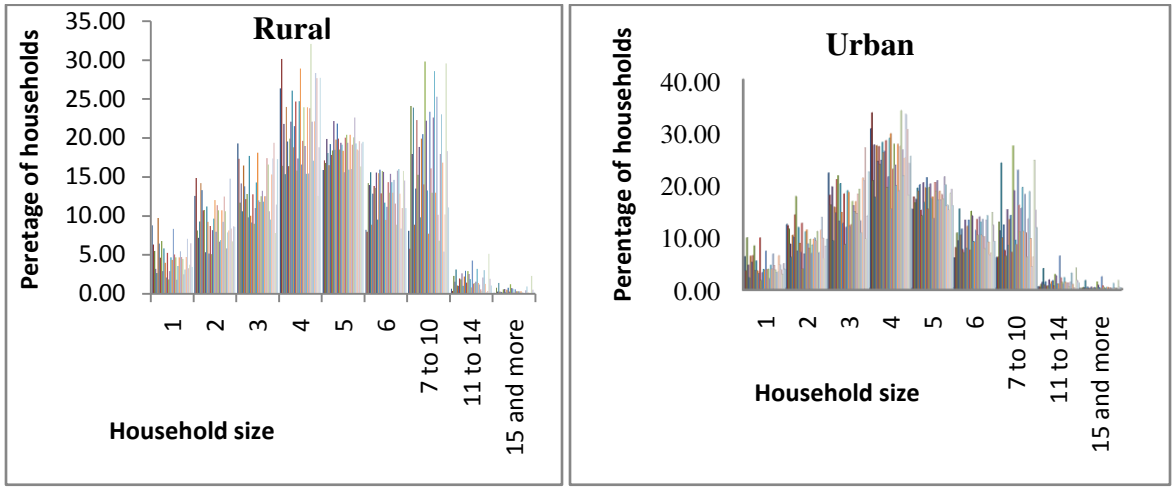
The figure 1 clearly shows that a decreasing trend of average household size in India is mostly from 1991. The average number of urban household size has declined more sharply than rural household size. This phenomenon could be attributed to increasing rate of higher education, different location of jobs from home and increasing urban economic pressure.



Source: Author's calculation using data from various year of Census

Below Figure 2 shows the percentage distribution of different members of households in 28 states and 7 Union Territories (UTs) in India in 2011. The figure shows that 21% of households in rural areas have 4 family members; followed by 18.8 percent households have 5 and 16.78 percent households have 7 to 10 family members. On the other hand, 25.38 % of urban households have 4 family members, followed by 17.87 % of households have 5 and 16.43% of households have 3 family members. This clearly indicates that urban households have lower number of members than rural households. The maximum number urban households consist of 3 to 5 family members whereas rural households have 4 to 7 family members.

Figure 2: Percentage distribution of different number of household members in different Indian States/UTs in 2011



Source: Author's using data from different Census period

Now the question arises, whether a nuclear family is better or a joint family. Obviously, there are some associated positive and negative effects. Nuclear families provide full freedom of decision making to a couple and encourage interdependence between them. In contrary, joint families can provide financial help, emotional help and a sense of community. Most importantly, the problems in child-rearing, marital conflicts and helplessness during illness can be avoided by living in a joint family. In fact, the increasing number of divorce, separation, depression and anxiety among Indians could be the reasons of lowering of household size in India.

In this paper, we try to understand the impact of higher level of economic development process on average family size in India. Higher level of development is measured by several ways; rate of urbanization, educational expansion, economic pressure, improvement in infrastructure, level of national income, poverty rate, extent of inequality and several other variables which depict the

development process. In addition to that, as different religious and social groups in India play an important role in various government decisions, we have also considered them in our study. We use both cross sectional and panel data models to estimate the impact of development on average household size in India from the period of 1991 to 2011. The study period from 1991 is chosen as India started its reform from 1991. The study has been conducted by considering aggregated (i.e., all India state level) and disaggregated level at household level separately. Data also have been used at aggregate and disaggregate level. Aggregate level data at all India/rural/urban is sourced mainly from Census of India, whereas household level data is taken from National Sample Survey data on consumption expenditure in 2011-12. The main contribution of this paper is the quantitative assessment of higher level of development on average household size. It is important to assess this relationship as both higher or lower household sizes have some benefits and some shortcomings. If higher economic development reduces family size than we need to suggest the policy makers to reduce the negative effects of this consequence, so that we can achieve the highest benefits of economic development. This stands as the main purpose of this paper.

The paper adopts the following structure. The next section reviews the related literature to find out the research gap. Section 3 introduces model, data as well the estimation strategy. Section 4 examines the higher economic development on average household size and discusses the results. Sections 5 and 6 present the major conclusions and policy suggestions, respectively.

2. Review of literature

Starting from Malthus (1798), economist tried to link with economic development on the population growth. According to third stage of demographic transition model, low birth-rate and low death rate are achieved when development process of a country shifts from agrarian to industrialized economy through fast urbanization process.

The linking between household size and economic development are not much highlighted. For instance, Barro and Becker (1989) modeled the choices of fertility and intergenerational transfers on economic growth and macroeconomic performance. This seminal contribution does not consider family size per se. Rios (1991) the demographic transition in Latin America has resulted in increased family size rather than the Western European model of reduced family size. Economic development in Latin America has not resulted in large enough changes to encourage

family size limitation. Fertility declines in Latin America in the future will be dependent on economic development, educational advancement for women, and a reduction in rural population. However, several studies mainly addressed the issue of sensitivity of poverty estimates by considering household size (Meenakshi and Ray, 2002; Coulter, Cowell, & Jenkins, 1992; Buhmann et al. 1988; Lancaster, Ray, & Valenzuela, 1999; Dreze& Srinivasan, 1997; Lanjouw&Ravallion, 1995). For instance, the empirical relationship between poverty and household size is fragile and especially sensitive to differences in the welfare indicator used in measuring poverty.

In case of India, Meenakshi and Ray (2002) analyzed the impact of household size and family composition on poverty in rural India. They found that the introduction of consumption economies of household size and of adult/child consumption relativities affect the poverty estimates but not the State poverty rankings. In contrast, female headed households display, in many States, higher poverty only in the presence of size economies and adult/child relativities. Nayak and Behera (2014) explored the trend and spatial pattern in the changes of household size in India using state level data. The study found that demographic parameters, in contrast to non-demographic parameters, have a stronger association with extremes of household size-either very small or very large. Smaller family size can spur economic development and reduce poverty**.** In this context, Kugler and Kumar (2016) found that family size indeed has a negative impact on schooling in India. The high fertility rate within households may therefore have caused the low level of human capital accumulation in India. Dommaraju (2015) using District Level Household and Facility Survey (DLHS-3) examines the demographic, social and economic characteristics of one-person households (OPH) in India. The study found that though the prevalence of OPH is low, the number of such households is large, and expected to grow in the next few decades. Elderly females and young migrants who live alone are potential vulnerable groups. The limited amount of research calls for serious attention on the effect of higher economic development on average household size in India. Allendorf (2013) examines whether young married women living in nuclear families have better health than those in patrilocal extended families in India. Results show that young married women living in nuclear families do not have better health than those in patrilocal extended families. From1992 to 2006, the percentage of young married women residing in nuclear families increased, although the majority remained in patrilocal extended families. This trend towards nuclear families will not benefit young women's health.

3. Model and Data

In order to analyze the effect of higher level of economic development on average household size in Indian state/UTs, we specify the following panel data model.

$$y_{it} = \beta_0 + \beta_1 X_{it} + \delta t + \eta_i + \varepsilon_{it} \quad (1)$$

Where y_{it} is the average household size of different state/UTs, X is a set of explanatory variables, η_i is the unobserved time-invariant specific effects; δt captures a common deterministic trend; ε_{it} is a random disturbance (assumed to be normal), and identically distributed with E (ε_{it}) = 0; Var (ε_{it}) = $\sigma^2 \succ 0$.

To choose between the panel data, models Breush and Pagan Lagrange Multiplier (LM) Test and the Hausman (H) Specification, diagnostic tests were conducted. The higher value (or significant) obtained in LM test indicated the advantages in choosing random effect or fixed effect model over pooled regression model. Further, the result of statistical significance of Hausman (h-test) specification test suggested that estimation by using FE model is advantageous over RE model. The panel data models are used as the models capable of capturing time invariant state/UTslevel characteristics such as geography and culture.

Though panel data model is efficient econometric model, but it suffers from the non-availability of data for various important variables in the different periods of time. Therefore, we consider the following OLS model for the same objective by considering several more important independent variables.

$$y_i = \alpha_0 + \sum_{i=1}^{22} \alpha_i x_i + e_i \text{---------- (2)}$$

In the above regression equation x_i represents the set of control variables, e as error term and α_0 stands as constant.

However, as aggregate level data can't depict the true picture at household level, we use the disaggregated level data at household level to investigate the same objective. Based on the nature of data which are used, we employ the following Tobit model as it fits perfectly. The Tobit model is represented by the following expression:

$$y_i^* = \mathbf{\beta X_i} + \varepsilon_i, i = 1,2,\ldots,N \quad \text{---------- (3)}$$

$$y_i = y_i^* \qquad if\, y_i^* > 0$$

$$= 0 \qquad if\ y_i^* \leq 0$$

where y_i^* is a latent variable that stands for average household size, Xi is a vector of independent variables,$\boldsymbol{\beta}$ is a vector of estimable parameters, ε_i is a normally and independently distributed error term with zero mean and constant variance $\sigma_{2.}$

Table 1 presents the detail of the control variables used for the estimation of above mentioned three econometric models. State/UTs level average household sizes are used as the dependent variables for all three models. The study period of the Panel data model is from 1991 to 2011 with 10 years interval. Due to unavailability of data for these specific Census years, the NSS data on "consumption expenditure" and "employment and unemployment" are used for the years of 1993-94, 1999-00 and 2011-12. The OLS model estimation is done by considering latest available Census data for the year of 2011. For the estimation of Panel data and OLS model, data is mainly collected from Census of India, Central statistical organization (CSO), and National Sample Survey Organization. More specifically, unemployment rate, poverty ratio, and Gini coefficient are calculated from NSS data whereas per capita income is collected from CSO. Rest of the data is taken from Census data. All these data are provided by Government of India. Finally, to run the Tobit model we use only NSS data on "consumption expenditure" for the year of 2011. Choice of independent variables is made based on the concept of development literature with the availability of data at our best.

Turning to our discussion on the measurement of the variable we first explain about the variables which are used in panel data model. Unemployment rate is measured by per 1000 population according to usual status. Poverty headcount ratio and Gini coefficient are used to measure the poverty rate in state/UTs level in India. All these three variables are sourced from NSS data on 'consumer expenditure' and 'employment and unemployment rate'. State level data are collected from Planning Commission, Government of India (GoI). Due to unavailability of one methodology for long term poverty measurement, Lakdawala methodology for the duration of 1993-94, 1999-00 and Tendulkar methodology for the duration of 2011-12 are used to measure the poverty in India. Therefore, the impact of poverty will be considered with some caution for the panel data model. Fertility rate, urbanization rate, literacy rate, and death rate are taken directly from Census of India. Finally, net state domestic products are collected from CSO. Again, as state net level domestic products are not available for one base year period in long

period of time, we have used per capita NSDP at current prices. The analysis is carried out by considering state/UTs aggregate level and with rural-urban distinction. Observations differ from different regression models due to missing information and because some of the states are recently formed.

Table 1: Details of the control variables used in Equations 1, 2 and 3

Major development indicators	Panel data model	OLS model	Tobit Model
Income	Per capita state domestic product	Per capita state net domestic product, owned house, housing condition, use of computer/laptop, mobile, two wheelers, and car.	Monthly per capita expenditure, total land possessed, regular salary earning member of a household, ration card holding status, and dwelling unit
Labour	Unemployment rate	Unemployment rate	Self-employed, regular wage/salary earning, casual labour, and others
Education	Literacy rate	Literacy rate and Gross enrollment ratio	
Health	Death Rate and fertility rate	Fertility rate and sex ratio	
Infrastructure		Road, electricity, tape water, Bore water, latrines, and banking	
Poverty	Poverty ratio	Poverty ratio	
Inequality	Gini coefficient	Gini coefficient	
Urbanization	Rate of urbanization	Rate of urbanization and density of population	
Religion			Hinduism, Islam, Christianity, Sikhism, Jainism, Buddhism, Zoroastrianism, and others
Social group			Scheduled Tribes, Scheduled Castes, Other Backward Classes, and others
Primary source of energy for cooking			Liquid petroleum gas, electricity, and others
Primary source of energy for lighting			Electricity, and others

Source: Author's compilation

For the OLS estimation, variables are collected from different sources. The analysis is done at state/UTs level. The common variables those are used in panel data model are sourced with same measurements. Percentage of people that have good house, own house, computer, mobile, 2-wheelers, car, literacy rate, sex ratio, urbanization rate, population density, road length, electricity generation, information on drinking water, latrine facility, and baking information all

are collected from Census of India. Gross enrollment for ration in primary school is taken from Education Ministry of Human Resource Development, GoI. Road length is measured by adding the roads of national, state highway, and public works department.

Tobit model estimation is done by considering unit level data from national sample survey organization on 'consumer expenditure survey' at household level by rural urban distinction. Total land possessed has been calculated by adding total land owned, leased-in, neither owned nor leased-in, and leased out. Monthly per capita consumption expenditures estimates are based on Mixed Recall Period (MRP) estimates of consumption expenditure. Dummy variable for regular salary earning member has been considered 1 if any member of the household a regular salary earner or 0 otherwise. If households possess ration card, the dummy variable takes the value 1 or 0 otherwise. The value of dummy variable for primary source of energy for cooking is 1, if households use Liquefied Petroleum Gas (LPG) and electricity, or, otherwise 0.[1] Finally, if households are using primary source of lighting as electricity then dummy variable takes value 1, or, otherwise 0.[2]

We expect that higher level of income, employment situation, education level, health situation, infrastructure facilities, urbanization rate, extent of inequality, better source of cooking and lighting have a negative effect on average household size, as improvement of all these factors leads to higher level of economic development and reduce average family size. On the other hand, higher level of poverty increases family size, as it stands as lower level of economic development. As India is a religious and caste based country, we always see public policies are made differently for different religious groups and castes. Keeping in mind, we try to see whether religion and caste system have any impact on average family size or not.

[1]NSS data provides primary source of energy for cooking from coke, coal, firewood and chips, LPG, gobar gas, dung cake, charcoal, kerosene, electricity, others, and no cooking arrangement.
[2] NSS data provides primary source of energy from lighting, kerosene, other oil, gas, candle, electricity, others-9, and no lighting arrangement.

4. Results

4.1 Main results

Estimated results from different regression model are presented in the section. Appendix table 1-3, presents the summery statistics of the variables used in panel data model, whereas, Appendix table 4 for OLS estimation. Summary statistics present the coefficient of variations as it provides relative variability and unit free measurement. Pair-wise correlation coefficients for panel data models are presented in Table5 -7. Appendix table 8 shows that pair-wise correlation coefficient for OLS model. The correlation coefficients show that average household size is negatively correlated with urbanization and literacy rate in the panel data model at the aggregate state/UTs level analysis. On the other hand, it is positively correlated with unemployment rate.

Table 2 presents the estimated results of equation (1) for the panel data model of the regression model 1 to 7. Estimations are done not only at overall state level, but also for rural and urban separately. Separate estimations for rural and urban areas are done to see whether there is any regional variation effect on state level average household size in India.

The significant value of $chi^2$of the LM test for all the regression models validates the use of estimation of panel model. The significant value of chi^2 of the Hausman test validated the choice of the fixed effect (FE) model for regression models 1 and 6 over random effect model for the regression estimation for regression models 2, 3, 4, 5 and 7. The significant level (1 % level) of F statistics for FE models and Wald $chi^2$for RE models show the overall significant level of the model. Overall correlation coefficients are also quite high except regression models 4 and 7. For instance, the regression model 1 explains 81 per cent of the total variation in the average household size.

The results of the estimated fixed and random effect models show that unemployment rate has a positive and statistically significant effect on the state level average household size in India. Poverty headcount ratio which is used to see the impact of poverty situation on average household size, show a statistically significant positive effect on average urban household size. In contrast, it does not show any statistically significant effect in rural area. The result indicates that though higher poverty situation in urban area encourages higher average urban family size but it does not encourage the same in rural areas. Total fertility rate also shows a positive effect on average household size. In contrary, inequality level has a statistically significant negative

effect on average household size. Generally, results are consistent for overall state/UTs level as well as for rural and urban analysis done separately. In particular, regression model 7 shows that a 100 % increase in unemployment rate increases average state level household size by 3.8 % of average household size. Regression model 6 shows that with a 100 % increase in poverty head count ratio and total fertility ratio, average household size increases by 4.4 % and 5.6 %, respectively. The results indicate that higher level of unemployment rate, total fertility rate and poverty rate have a positive effect on average household size. On the other hand, level of inequality has a negative effect on household size. Higher level of development as evidenced by lower level of poverty, unemployment rate, and total fertility rate, all have negative effect on average household size. In other words, higher level of poverty, unemployment rate and total fertility rate situation represents the lower level of development, which encourages higher average family size and helps financially through sharing and investing with more number of children.

As our estimated results show a positive impact of these variables on average household size, we can infer that lowering of the value of these indicators reduces the family sizes which are evidenced in the current Indian economy. Results also support that higher level of inequality, which is a byproduct of India's current development process has a negative effect on average household size. The results clearly indicate that higher level of development has a negative effect on average household size. The results also confirm the negative effect of urbanization rate and literacy rate, which also increases by the higher level of economic development, on average household size at overall state level in India.

Table 2: Dependent variable: Log of average household size,1991-2011

Independent variable	Rural		Urban		State/UTs level		
	FE	RE	RE	RE	RE	FE	RE
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Log of unemployment rate	0.033	0.039**	0.062***			0.021	0.038**
	(0.025)	(0.013)	(0.019)			(0.012)	(0.015)
Log of poverty head count rate	0.053		0.036	0.054*	-0.015	0.044***	
	(0.0322)		(0.022)	(0.025)	(0.027)	(.0143)	
Log of inequality	-0.363***	-0.201**	-0.219***	-0.232**		-0.014*	-0.018**
	(0.121)	(0.09)	(0.07)	(0.084)		(0.007)	(0.008)
Log of total fertility rate	0.325***	0.325***	0.09	0.133**		0.056*	
	(0.038)	(0.028)	(0.055)	(0.062)		(0.031)	
Log of per capita net state domestic product					-0.003		
					(0.009)		
Log of urbanization rate					-0.069**		
					(0.024)		
Log of literacy rate					-0.91***		
					(0.153)		
Log Death Rate by Residence					-0.093		
					(0.068)		
Intercept	0.583**	0.961***	0.975***	1.09***	6.112***	1.34***	1.56***
	(0.264)	(0.159)	(0.071)	(0.072)	(0.879)	(0.078)	(0.057)
LM(chi^2)	14.31***	24.62***	16.78***	12.44***	29.01***	22.44***	28.02***
H(chi^2)	16.67***	0.118	6.67	1.83	1.75	9.55**	2.55
F stat/ Wald chi^2	37.16***	143.02***	103.64***	84.69***	54.13***	18.39***	17.42***
Overall R^2	0.81	0.58	0.22	0.21	0.30	0.49	0.17
Number of observation	62	63	60	62	84	72	94

Robust standard errors in parentheses

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

Table 3 presents the estimated regression results from Equation (2). Regressions 8–10report OLS results, with controlling heteroskedasticity and multicollinearity problem. The dependent variable is logarithm of state/UTs level average household size. Regression models are different for the different number of available observations. Regression (8) shows the estimates of the full model which includes all variables. Regression (9) and (10) report results for a parsimonious model, excluding controls that are not found to be statistically significant or matched with the expected sign of the regression parameters. More specifically, due to paucity of data, we ran regression (9) to (10) and have presented the results of the best fitted models in terms of predicted signs, significance level of the variables and goodness of fit of the regressions, according to available different number of observations of the variables.

The significant values of F statistics for Regressions 8–10 indicate that the overall model is statistically significant. The insignificant Chi^2 values of Breusch-Pagan / Cook-Weisberg test indicate the absence of heteroskedasticity problem in our models. The insignificant F statistics of Ramsey RESET test for omitted-variable bias show that we do not need more variables. Though the mean variance inflation factor (VIF) values are less than 10 for Regression models 8-10, but regression Model 8 has some multicollinearity problem, as the individual VIF values of some of the variables are more than 10. For that reason, we rely more on Regression models 9 and 10 than Regression model 8. The test of normality, i.e., that the residuals are normally distributed, is confirmed by kernel density estimates, which are presented in Appendix Figures A1, A2, and A3. Two non-graphical tests are also done by considering the Shapiro–Wilk test for normality and Jarque-Bera normality test. The statistically insignificant Z values of the Shapiro–Wilk test values do not reject the null hypothesis that the distribution of the residuals is normal. The insignificant p values of Jarque-Bera normality test also support the same. The higher values of R^2 indicate that Regressions 8– 10 explain a good percentage of total variation in the dependent variable. We also calculate the adjusted R^2, as it adjusts for the number of explanatory terms in a model, i.e., it incorporates the model's degrees of freedom. In regression 8-10, among the proxy variables of income, we find that per capita net state domestic product, the percentage of people that have good housing and the percentage of people that are using computers have a significant negative effect on state/UTs level average household size in India. In particular, a 10 percent increase in per capita net state domestic product is associated with2 per cent decrease in average household size. The result is as per our expectation, that is, the higher level-

Table 3: OLS and IV - estimates dependent variable: log of average household size, 2011

Independent variables	*OLS*			*IV*
	Model 8	***Model 9***	***Model 10***	***Model 11***
Income				
Log of per capita net state domestic product	-0.222 (0.0850)**	-0.200(0.0662)**		-0.187 (0.0388)***
Log of percentage of people have good housing	-0.344 (0.155)*	-0.333(0.140)**		-0.165 (0.112)
Log of percentage of people live in own house	-0.106 (0.233)	-0.0929 (0.155)		0.211 (0.354)
Log of percentage of people are using computer	0.0350 (0.132)		-0.224(0.0928)**	0.0831 (0.0836)
Log of percentage of people are using mobile	-0.0797 (0.175)		0.249 (0.110)**	-0.122 (0.154)
Log of percentage of people are using 2-wheelers	0.0980 (0.0439)*	0.0920 (0.0374)**		-0.000954 (0.0638)
Log of percentage of people are using car	0.109 (0.0725)	0.104 (0.0465)**	0.0926 (0.0511)*	0.173 (0.0384)***
Unemployment rate				
Log of unemployment rates	-0.0232 (0.0281)	-0.0202 (0.0235)	0.0653 (0.0228)***	-0.0271 (0.0127)**
Education				
Log of literacy rate	-0.108 (0.282)	-0.123 (0.256)	-0.710 (0.295)**	-0.209 (0.120)*
Log of gross enrollment ratio	-0.134 (0.0886)	-0.150 (0.0560)**	-0.116 (0.0475)**	-0.0429 (0.1000)
Health				
Log of total fertility rate	0.0621 (0.0310)*	0.0559 (0.0262)*		0.0382 (0.0266)
Log of sex ratio	-0.858 (0.404)*	-0.816 (0.363)**		-0.676 (0.397)*
Inequality				
Log of Gini coefficient	0.0118 (0.183)	0.0318 (0.148)		-0.0692 (0.211)
Poverty				
Log of poverty head count ratio	-0.154 (0.0488)**	-0.136 (0.0305)***		-0.00236 (0.00828)
Urbanization				
Log of percentage of people living in urban area	-0.0105 (0.0545)	-0.00947 (0.0366)	0.00998 (0.0443)	0.0189 (0.0400)
Log of population density	-0.0162 (0.0218)	-0.0132 (0.0188)		-0.0158 (0.0111)
Infrastructure				
Log of total road length	0.0338 (0.0204)	0.0301 (0.0155)*		0.0207 (0.0243)
Log of Electricity Generation	-0.0917 (0.0405)**	-0.0843 0.0318)**		-0.0607 (0.0328)*
Log of percentage people use drinking water from tape	0.117 (0.0813)	0.112 (0.0497)**		0.0687 (0.0450)
Log of percentage of people use drinking water from bore	-0.0473 (0.0158)**	-0.0468 (0.0144)***	-0.0395 (0.0130)***	-0.0187 (0.0314)

Log of percentage of people have latrine facility	-0.0922 (0.0997)	-0.0787 (0.0791)	0.00521 (0.0756)	-0.103 (0.0571)*
Log of percentage of people use bank	-0.254 (0.198)	-0.257 (0.130)*		-0.0679 (0.174)
Constant	13.82 (2.847)***	13.04 (2.186)***	4.136 (1.123)***	9.415 (5.962)
F statistics	4.79***	6.24***	4.34***	34.8849**
Breusch-Pagan / Cook-Weisberg (chi2)	0.04	0.03	0.01	
Ramsey RESET test (F value)	0.7659	0.62	1.26	
Mean VIF	8.74	5.03	2.77	
Shapiro-Wilk test for normality (Prob>z)	0.81230	0.97389	0.69839	
Jarque-Bera normality test (p value)	0.6795	0.8993	0.6928	
R-squared	0.921	0.919	0.619	0.934
Adjusted R-square	0.7291	0.7718	0.4763	
Observations	32	32	34	27

Standard errors in parentheses

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

of economic development associated with increase in per capita income, use of good house and computer, lead to reduction in average household size. However, surprisingly, the percentage of people using mobile, 2-wheelers, and car have a positive effect on household size which run counter to our expectation. Results indicate that though overall family size is declining in India, but richer families show an increase in average household size, as richer families generally possess cars, two wheelers and more number of mobile phones. Unemployment rate has a positive and statistically significant (at 1 % level) effect on log of average household size. Results indicate that a 10 % increase of unemployment rate increases average household size by 0.6 %. Educational achievements which are measured by literacy rate and gross enrollment ratio have a negative effect on log of average household size. This implies that higher achievements in education which lead to higher economic development reduce average family size. Measurement of health outcome, total fertility rate has a positive effect on average household size in India whereas higher sex ratio decreases average household size. It is important to note that, currently, India is experiencing an increase in sex ratio from 927 in 1991 to 943 in 2011. In the same time, average family size is also decreasing. Our results, which show that increasing sex ratio leads to decline in average household size, support this phenomenon. Impacts of inequality and urbanization have statistically insignificant effect on average household size in Indian state/UTs. These results are quite justifiable, as India's current rate of urbanization (i.e., 31% in 2011) and inequality (i.e., consumption inequality is 35.9 in 2011-12) are moderate. Therefore, the impact of these variables on urbanization is insignificant. Surprisingly, poverty rate has a negative effect on the average household size which was not expected. Finally, provision of infrastructure which is measured by availability of total road length, electricity generation, percentage of people sourcing drinking water from tap and bore, use of latrine facility, and banking systems have a mix effect on average household size in India. More particularly, total road length and percentage of people drinking water from tap have a positive effect on average household size. In contrast, state/UTs wise electricity generation, percentage of people drinking water from bore, and percentage of people using banking systems have a negative effect on the dependent variable. The results indicate that level of infrastructure does not have any strong effect on the average household size, as India is suffering from severe deficiency in infrastructure. However, as the important infrastructure variables have a negative effect we can argue that by and large development in infrastructure affects the average household size in India negatively.

Table 4: Tobit regression model estimates - Dependent variable: Average household size, 2011-12

VARIABLES			
Rural	***Urban***	***Rural***	***Urban***
		Model 12	**Model 13**
Income			
Total land possessed		0.195 (0.005)***	0.18727 (0.009)***
Monthly per capita consumption expenditure		-0.273 (0.006)***	-0.204 (0.004)***
Dummy variable for households have regular salary earning member		0.959 (0.0444)***	0.748 (0.0460)***
Dummy variable for households have ration card		0.338 (0.0248)***	0.563 (0.0226)***
Reference category: Owned			
Hired		-0.930 (0.0489)***	-0.733 (0.0231)***
No dwelling unit		-0.460 (0.315)	-1.029 (0.317)***
Others		-0.867 (0.0646)***	-1.061 (0.0484)***
Labour			
Reference category: self-employed in: agriculture	Reference category: self-employed		
Self-employed in non-agriculture	Regular wage/salary earning	-0.149 (0.0248)***	-1.006 (0.0475)***
Regular wage/salary earning	Casual labour	-0.979 (0.0503)***	-0.358 (0.0319)***
Casual labour in: agriculture	Others	-0.734 (0.0355)***	-1.595 (0.0334)***
Casual labour in non-agriculture		-0.423 (0.0294)***	
Casual labour in others		-1.675 (0.0414)***	
Religion			
Reference category: Hinduism			
Islam		0.781 (0.0278)***	0.799 (0.0278)***
Christianity		0.352 (0.0382)***	0.0732 (0.0445)*
Sikhism		0.396 (0.0592)***	0.304 (0.0739)***
Jainism		-0.343 (0.272)	0.315 (0.120)***
Buddhism		-0.202 (0.0795)**	-0.158 (0.107)
Zoroastrianism		1.138 (2.086)	9.196 (1.372)***
others		-0.0103 (0.0872)	0.253 (0.118)**
Social group			
Reference category : Scheduled Tribes			
Scheduled Castes		0.137 (0.0329)***	-0.141 (0.0462)***
Other Backward Classes		0.141 (0.0288)***	-0.332 (0.0409)***
others		0.0578 (0.0309)*	-0.357 (0.0412)***
Dummy variable for Primary source of energy for cooking		-0.0751 (0.0221)***	0.261 (0.0228)***
Dummy variable for Primary source of lighting Electricity		-0.0512 (0.0227)**	0.0542 (0.0482)
Constant		4.888 (0.0405)***	4.848 (0.0632)***
LR chi^2		9168.07***	11215.09***
Pseudo R^2		0.0345	0.0610
Log likelihood		-128222.2	-86335.635
Observations		59523	41633

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4 presents the estimated results from Tobit model. The Tobit model or censored regression model is used to estimate the linear relationship between variables as we find the evidence of the left-censoring in the dependent variable of average family size. Histogram in the Appendix of Figures A4 and A5 confirm that. As NSS unit level data on consumption expenditure is available separately for rural and urban distinction, we run the Tobit model for rural and urban separately. The statistically significant values of the likelihood ratio (LR) chi-square tell us that our model, as a whole, fits well. Regression models 12 and 13 present the estimated results for rural and urban area, respectively.

Results indicate that total land possessed, dummy variable for households that have regular salary earning member and dummy variable for households that have ration card, statistically have a positive effect on log of average household size in India. The results are consistent for rural and urban separate estimations. This indicates that poor households that have ration cards or rich families that have regular salary earning members, or possess bigger areas of land have higher number of family members then those who do not have these assets. However, higher level of monthly per capita consumption expenditure (MPCE) significantly and negatively impact on average household size. Dwelling status, those living in the rented house and 'others' have a negative effect on average household size compared to reference category household living in own dwelling unit. This is obvious as development process is evidenced by increase in households with a rented house. National Multifamily Housing Council data shows that 37% of American households were living in rented houses in 2014. In case of India, Census data shows that 53% households in Chennai district lived in the rented house in 2011. However, urban household member those do not have any dwelling unit have a negative effect on average household size whereas the result is not significant for rural household in India.

Employment status in rural and urban areas shows that increase in any of the categories of employment has a negative effect on average household size. For instance, a 10 percent increase in self-employed in non-agricultural rural areas and regular wage/salary earning in urban areas decrease average household size by 1.5 percent in rural and 10 percent in urban areas, compared to reference category, self-employed in agriculture in rural areas and self-employed in urban areas.

The impact of different categories of religious groups show that Islam, Christianity and Sikhism have a positive impact on average household size in both rural and urban areas compared to the reference category, Hinduism. Jainism, Zoroastrianism and others in urban areas have a positive effect on average household size whereas the results are not statistically significant for rural areas. On the other hand, though Buddhism has a negative effect on average household in rural areas but it does not have any impact in urban areas when we compare with Hinduism.

The results show that though scheduled castes, other backward classes and other religious groups have a positive impact on average household size in rural areas, but they have negative effect in urban areas compared to the reference category Scheduled Tribes. Finally, the results show that dummy variable of primary source of energy for cooking and dummy variable of primary source of lighting have positive (or negative) effect on average household size in urban (or rural) areas in India.

4.2 Comparison of major regression results from different estimates

Though previous section clearly explained the details about the results, but for an overall assessment, it is important to compare the main results those are obtained from various regression models. Our income variable i.e., per capita net state domestic product in OLS model and monthly per capita consumption expenditure in the OLS model has negative effect on the average household size. This clearly indicates that higher level of development which is evidenced by higher income and consumption reduces average family size. Higher level of employments which includes self-employed, regular wage/salary earning, and casual labour have a negative effect on average family size in the Tobit model. Panel and OLS regression models show that higher level of unemployment increases it. This also indicates in the same direction, that higher level development through increasing employment level decreases household size. Higher achievement of education which is measured by literacy rate and gross enrollment ratio have a negative effect on the dependent variable. Health outcomes, such as lower level of total fertility rate leads to lower average household size as we find positive relationship between them. Urbanization rate, which is measured by percentage of people living the urban area reduces average family size as we estimate the negative relationship between them in the panel data regression models. Level of inequality measured by Gini coefficient has a negative effect on

average household size. This clearly says that higher economic development which is evidenced by increasing inequality level has a negative effect on average family size. Poverty head count ratio has a positive effect on average household size in panel data model but it has a negative effect in OLS regression model. The results contradict to each other, however, as panel data model is more superior estimation than OLS model, we rely upon the panel data model and infer that the lower the poverty level, the more is the decrease in average household size. Finally, Tobit model which uses household level data shows that different religious groups and social groups also have a strong effect on average household size in India.

4.3 Robustness check

A serious concern about the relationship between poverty and average household size is the question of reverse causality. Does the poverty increase heighten average household size, or, does higher average household size increases poverty? The reality is possibly a mix of both. Poor households may be less concerned about the family planning or on the other hand, higher level of average household size may have positive impact on poverty; due to lesser opportunities in developing country like India. By considering this phenomenon, we use instrumental variable (IV) regression model to check the robustness of our regression results.

The suitable instruments are considered in such a way that they have very strong relationship with poverty but are exogenous, i.e., not anyway associated with average household size. Keeping in mind all these issues we find the following instruments for poverty; first, state wise yield of food grains which have relationship with poverty; as higher agricultural production reduces poverty by supplying higher amount of food. Second, number of land holdings family in states have, as it reduces poverty by providing opportunity to work in their own land and increases food supply. Third, state wise total Scheduled Castes (SCs) and Scheduled Tribes (STs) population as they are historically disadvantaged people in India and have strong relationship with poverty abject. It is important to note here, that though these instruments are time variant but due to limitation of data we could only run it for single period of time by considering the variables those are used to run the OLS model. In the context of exogeneity, we find that all these instruments do not have any strong effect on average household size. The simple correlation between state wise average household size and state wise yield of food grains (or number of land holdings family or SC and ST population) is -0.06 (0.12 or 0.09). We run the

full regression model 8 by keeping all control variables same. We estimate the model using 2SLS estimator. Regression model 11 in Table 3 presented the estimated results. Our instruments work well as first stage F statistics pass comfortably the rule of thumb threshold for strong instrument (Staiger and Stock, 1997) and exceed the Hausman et al. (2005) threshold values.

The result shows that poverty does not have any statistically significant effect on average household size. Therefore, we rely on the estimation of panel data model than OLS regression and conclude that lower level of poverty reduces state level average household size in India. The results are consistent for unemployment rate, net state domestic product, electricity generation, literacy rate, sex ratio, availability of latrine facility, and car owned. All these variables have statistically significant effect on average household size at state/UTs level in India. The sign of the estimated coefficient of these variables are matched with the sign which are obtained in OLS estimation. However, unemployment rate has a negative effect on average household size which is positive in the OLS and Panel data model. This result is very important to explain the impact of unemployment on average household size. Rural male (or female) unemployment rates per 1000 persons in the labour force according to usual status show to have increased from 14 (or 9) to 17 in 1993-94 to 17 in 2011-12. On the other hand, urban male (or female) unemployment rates per 1000 persons in the labour force according to usual status show to have decreased from 41 (or 61) to 30 (or 52) in 1993-94 to 17 in 2011-12. Therefore, we rely more on IV and panel regression results than OLS results and argue that though in short run higher level unemployment rate reduces average household size but in long run lower level of unemployment rate, which is the part of higher economic development, reduces average household size. All these results confirm our main claim that higher level of economic development reduces average household size in India.

5. Conclusions

In this paper, we have analyzed the impact of higher level of economic development on average household size in Indian state/UTs. We use aggregate and disaggregate level data to analyze this relationship in cross sectional and panel data models. Higher economic development is measured by increase in income, employment status, educational achievement, health outcome, poverty, inequality, urbanization, infrastructural improvement, primary source of energy for cooking and lighting. In addition to that as India's policy makers consider religious and social groups in various decisions, we also consider them in our analysis. Average household size is considered

as a dependent variable for the analysis. We use panel data i.e., fixed effect and random effect models from 1991-2011 to assess the relationship in long run, whereas, OLS, Tobit and IV regression models are used for cross sectional study for the period of 2011. We consider different regression models due to availability of different variables in different time periods and use of different units of analysis. Aggregate level study which considers 28 states and 7 UTs for the unit of analysis are done by sourcing data mainly from Census of India, Central Statistical Organization, and National Sample Survey data on 'consumption expenditure' and 'employment and unemployment'. Disaggregated level study considers unit/household level data from the NSS survey on 'consumer expenditure' in 2011-12.

Estimated Panel data models show that unemployment rate according to usual status, health outcomes measured by death rate and total fertility rate, poverty head count ratio have a positive effect on state/UTs level average household size in India. Educational achievement, measured by literacy rate, level of inequality (Gini coefficient), and rate of urbanization have a negative effect on average household size in India. Rural, urban and overall level separate analysis provides similar results except that rural poverty does not have any significant effect on the dependent variable.

OLS regression results show that the income variable which is measured by per capita net state domestic product and housing condition, and use of computer/laptop have a negative effect whereas, use of mobile, two wheelers, and car have positive effect on average household size. Unemployment and literacy rate show the same result which are obtained in panel data model. Gross enrollment ratio, poverty head count ratio and sex ratio show the negative effect. Infrastructural variables, such as road length, use of tap water, show the positive effect whereas, use of bore water and electricity have negative effect on average household size.

Tobit model shows that monthly per capita expenditure and status of dwelling unit of a household have a negative effect on average household size. On the other hand, dummy variable for household with regular salary earner and ration card holder have positive effect. Employment statuses such as self-employed, regular wage/salary earning and casual labour have a negative effect on average household size. Various religious groups such as Islam, Christianity, Sikhism, and Jainism have positive and Buddhism has negative effect on average household size, compared to the reference category Hinduism. Results also show that though different social

groups have positive effect in rural areas but they have negative effect in urban areas. Finally, it shows that use of LPG and electricity for cooking and use of electricity for lighting have negative (or positive) effect on rural (or urban) average household size.

As poverty may have impact on household size and household size may have impact on poverty, we estimate IV regression model to solve this endogeneity problem. We use following three instrumental variables of poverty, i.e., state wise yield of food grains, number of land holding families, and state wise total SCs and STs population. Instruments pass the required statistical tests. The results show that the results are consistent for major variables which we have estimated in OLS and Panel data model. Most importantly, it shows that though unemployment level has negative effect in cross sectional study but it has positive effect on average household size in the long run. In addition to that, lower poverty head count ratio decreases average household size at state/UTs level in India. The results confirm that higher level of development has a negative effect on average household size in India.

6. Discussions

Our results confirm that higher economic development leads to reduce the family size. As we have discussed in the introduction part that small household sizes also have some shortcomings such as, problems in child-rearing, help during illness, divorce, separation, depression and anxiety among the people. It is also important to note that lowering the family size may increase the land distribution dispute between old and new separated households.

Urban, and even in rural busy life have widened the communication gap between the parents and children in the sense that how much space should be given to their children and how much sense of discipline shall be inculcated. Children are missing love, affection and care from their busy workaholic parents. Most of the urban working parents use day care centres for their babies, but it is important to note that nobody can take the place of parents in a child's life. The consequence of this phenomenon is that children are lacking their moral development which is affecting their lives. Children are using social media and internet inappropriately as it is tough to separate them from these things currently.

Recently, developed countries have experienced a reduction of average family size and divorce rate. For example, the divorce rate in United Kingdom was 42 percent in 2015 whereas, 46

percent in United States in 2014. The divorce rate in India is about 13 per 1000 which is much lower than 500 in 1000 marriages in UK, but recently we have seen a rise in this ratio in India, in the metro cities and developed states such as Punjab. For example, from the period of 2003 to 2011, Kolkata has seen a 350 percent increase in divorce rate. This indicates that freedom of decision making, which generally a small size family offers may lead to this problem. However, the problem is that the system and structure of Indian society must change a lot to accommodate the separated couple, as it may lead to mental problems. Rognmo (2013) found that mental problems in couples after divorce are increasing in Norway. Most importantly, single women face many problems such as physical disorder and limited scope of social interaction which leads to serious problems.

Finally, it is also important to note that dispute related to land and property make up to about two-thirds civil cases in India.[3] This could be the reason that large sized families parted out to small family size households but splitting the property has either not been done yet, or, not been done properly. This has serious problem in using the land or conversion of land for different uses, such as from agriculture to industrial use.

Therefore, we suggest that there is a serious need for the research on the issue of higher level of economic development with increasing number of families of smaller sizes. Parents have to provide quality time to their children so that children get moral support and can believe in themselves to handle the pressure which is thrown out by the higher level of developmental aspects, such as education and economic pressure. To avoid divorce and other problems between the couples, they need to have a good balance between work and family. This has to be supported not only by the employer but also by the employees. In most of the cases it is the employer than the employees as they get less freedom in the unorganized sector which accommodates a major(i.e., 82.7% of workforce in 2011-12) share of workers in India. Finally, land reform is essential to reduce the property dispute in India for a healthy family life which will appreciate the higher economic development with a wide spread success.

[3] The results are obtained from a recent study Daksh, a civil society organisation that undertakes research and activities to promote accountability and better governance in India.

Reference

Adams, B. N. (2010). Themes and threads of family theories: A brief history. *Journal of Comparative Family Studies, 41,* 499–505.

Allendorf, K., (2013). Going Nuclear? Family Structure and Young Women's Health in India, 1992-2006",*Demography* 50 (3), 853-880.

Barro, R. J.,Becker, G. S., "Fertility Choice in a Model of Economic Growth" *Econometrica*57(2): 481-501.

Bongaarts, J. (2001) "Household Size and Composition in the Developing World in the 1990s", *Population Studies* 55(3), 263-279.

Buhmann, B., Rainwater, L., Schmaus, G., Smeeding, T. (1988). "Equivalence scales, well-being, inequality and poverty: Sensitivity across ten countries using the Luxembourg Income Study (LIS) Database",. *Review of Income and Wealth*, *94*, 115–142.

Coulter, F. A. E., Cowell, F. A., Jenkins, S. P. (1992). "Equivalence scale relativities and the extent of inequality and poverty. *Economic Journal102*, 1067–1082.

Dommaraju, P. (2015). "One-person households in India", *Demographic Research* 32 (45), 1239−1266.

Dreze, J., Srinivasan, P. V. (1997). "Widowhood and poverty in rural India: Some inferences fromHousehold Survey Data", *Journal of Development Economics*54, 217–234.

Hausman, J., Stock, J. H., Yogo, M. (2005). "Asymptotic properties of the Hahn-Hausman test for weak-instruments",*Economics Letters* 89, 333-342.

Kugler, A., Kumar, S. (2016). "The fewer the merrier: Family size and education in India", Published on *VOX, CEPR's Policy Portal. Web address:* http://voxeu.org/print/60268 *(Accessed on 28th November, 2017).*

Lancaster, G., Ray, R., Valenzuela, R. (1999). A cross country study of household poverty and inequality on Unit Record Household Budget Data. *Economic Development and Cultural Change*48, 177–208.

Lanjouw, P., Ravallion, M. (1995). "Poverty and household size". *Economic Journal*, *105*, 1415–1434.

Malthus T. (1798). An Essay on the Principle of Population. London: Printed for J. Johnson, in St. Paul's Church-Yard.

Meenakshi, J.V., Ray, R. (2002). "Impact of household size and family composition on poverty in rural India", *Journal of Policy Modeling* 24, 539–559

Nayak, D. K,Behera, R. N. (2014). "Changing household size in India: An inter-state comparison",Transactions, 36(1), 1-18.

Rios, R.J. (1991)."Economic Development and Family Size", *American Economist* 35, 81-85.

Rognmo, K., Torvik, F.A., Idstad, M., Tambs, K. (2013). "*More mental health problems after divorce in couples with high pre-divorce alcohol consumption than in other divorced couples: results from the HUNT-study*", MC Public Health, doi: 10.1186/1471-2458-13-852.

Staiger, D., Stock, J. H. (1997). "Instrumental variables regression with weak instruments", *Econometrica* 65(3), 557-586.

Tripathi, S. (2013), "Do large agglomerations lead to economic growth? Evidence from urban India", *Review of Urban & Regional Development Studies* 25(3), 176-200.

Tripathi, S., Mahey, K. (2017). "Urbanization and economic growth in Punjab (India): an empirical analysis", *Urban Research & Practice* 10 (4), 379-402.

Web articles

https://timesofindia.indiatimes.com/india/supplemented-nuclear-families-make-16-of-indian-households/articleshow/59449874.cms
http://bestofme.in/nuclear-family-vs-extended-family-in-india/
https://scroll.in/article/669053/the-nuclear-family-is-on-the-decline-in-india
https://blog.id.com.au/2012/population/australian-census/2011-census-our-expanding-households/
https://timesofindia.indiatimes.com/india/Median-household-size-drops-below-4-in-cities/articleshow/12397117.cms
http://www.thehindu.com/news/cities/chennai/half-of-city-lives-in-rented-homes/article3251937.ece
http://www.livemint.com/Opinion/a5jnMOHQsHEk47Rr9mUWPI/Five-charts-on-the-state-of-Indias-housing-sector.html
http://www.thehindu.com/opinion/open-page/The-right-balance-in-child-rearing/article16907501.ece
https://en.wikipedia.org/wiki/Divorce_demography
https://www.thequint.com/news/india/not-so-happily-ever-after-yet-india-has-a-low-divorce-rate
http://www.deccanchronicle.com/nation/current-affairs/090816/millions-of-land-property-cases-stuck-in-indian-courts.html
http://www.un.org/en/development/desa/population/publications/pdf/popfacts/PopFacts_2017-2.pdf

Appendix Table 1: Summary statistics of the variables used in panel data model for aggregate state/UTs level

Variable	Observation	Mean	Standard deviation	Minimum	Maximum	Coefficient of variation
Average household size (HH)	101	5.22	0.73	3.70	7.90	13.98
Unemployment rates (per 1000) according to usual status (UNP)	105	36.04	35.17	0.00	184.50	97.59
Total fertility rate (TFR)	78	2.69	1.16	0.7	5.8	43.00
Gini coefficient in percentage (GINI)	71	0.281	0.038	0.198	0.38	13.49
Per capita net state domestic productin Rs. (SDP)	91	35436.49	38810.97	2868	211569	109.52
Poverty head count ratio in percentage (PHR)	98	23.91	13.09	1.00	54.96	54.75
Percentage of urbanization (UN)	101	33.06	20.13	8.47	97.50	60.89
Literacy rate in percentage (LIT)	101	82.18	6.91	61.00	97.60	8.41
Death rate (DR)	98	7.20	2.07	3.30	13.80	28.75

Source: Author's calculation

Appendix Table 1: Summary statistics of the variable used in panel data model for rural state/UTs level

Variable	Observation	Mean	Standard deviation	Minimum	Maximum	Coefficient of variation
Average household size (HH)	101	5.22	0.67	3.63	7.00	12.85
Unemployment rates (per 1000) according to usual status (UNP)	98	26.52	36.54	0.00	194.00	137.78
Poverty head count ratio in percentage (PHR)	97	25.00	15.74	0.00	62.60	62.97
Gini coefficient in percentage (GINI)	70	0.25	0.04	0.16	0.36	15.09
Total fertility rate (TFR)	91	2.52	0.94	0.70	5.10	37.29

Source: Author's calculation

Appendix Table 3:Summary statistics of the variable used in panel data model for urban state/UTs level

	Observation	Mean	Standard deviation	Minimum	Maximum	Coefficient of variation
Average household size (HH)	101	5.11	0.73	3.80	7.00	14.31
Unemployment rates (per 1000) according to usual status (UNP)	99	45.65	35.29	0.00	200.00	77.32
Poverty head count ratio in percentage (PHR)	99	17.31	11.81	0.00	48.38	68.27
Gini coefficient in percentage (GINI)	71	0.32	0.05	0.20	0.44	15.73
Total fertility rate (TFR)	90	1.99	0.60	1.10	3.70	29.99

Source: Author's calculation

Appendix Table 4: Summary statistics at state level OLS regression model

Variable	Observation	Mean	Standard deviation	Minimum	Maximum	Coefficient of variation
Average household size (HH)	4.93	0.74	3.70	7.30	14.94	500
Percentage of people have good house condition (GH)	57.55	12.67	29.53	78.65	22.01	233
Percentage of people owned house (OH)	79.54	14.49	38.26	96.79	18.22	264
Unemployment rates (per 1000) according to usual status (UNP)	36.40	38.81	0.00	177.00	106.63	0
Total fertility rate (TFR)	9.26	5.14	1.00	20.00	55.54	19
Gini coefficient in percentage (GINI)	0.29	0.04	0.20	0.38	15.42	500
Per capita net state domestic product in Rs. (SDP)	74792.48	40960.18	22581.75	211569.80	54.77	55
Poverty head count ratio in percentage (PHR)	18.49	11.58	1.00	39.90	62.63	9
Percentage of people living in urban area (UN)	36.38	20.24	10.03	97.50	55.62	50
Total road length in kilometer (ROAD)	35433.09	52811.96	88.00	237725.00	149.05	0
Total electricity generation in megawatt (ELE)	18.00	10.25	1.00	35.00	56.93	10
Literacy rate in percentage (LIT)	77.85	8.59	61.80	94.00	11.04	719
Density (DEN)	16.74	9.48	1.00	33.00	56.61	11
Sex ratio (SEX)	931.03	79.74	618.00	1084.00	8.56	775
Gross enrollment ratio in percentage (GER)	21.18	9.57	3.90	42.20	45.21	41
Percentage of people use drinking water from tape (TAPE)	53.03	27.22	4.40	96.72	51.34	16
Percentage of people use drinking water from bore (BORE)	7.09	6.79	0.04	21.87	95.77	1
Percentage of people have latrine facility (LATR)	62.38	22.50	22.03	97.80	36.07	98
Percentage of people are using banking facility (BANK)	62.10	16.01	28.64	89.28	25.78	179
Percentage of people are using computer or laptop (LAP)	11.44	6.75	4.65	33.23	58.99	69
Percentage of people are using mobile facility (MOBILE)	63.04	15.17	29.17	85.63	24.07	192
Percentage of people are using two-wheelers (TW)	23.00	13.44	2.83	56.94	58.44	21
Percentage of people are using car (CAR)	7.11	5.84	1.71	25.74	82.21	29

Note: The results are based on 35 observations. Source: Author's calculation

Appendix Table 5: Pair-wise correlation coefficient of the variable used in panel data model for aggregate state/UTs level

	HH	UNP	TFR	GINI	SDP	PHR	UN	LIT	DR
HH	1								
UNP	0.24*	1							
TFR	0.10	-0.21	1						
GINI	-0.17	-0.09	-0.21	1					
SDP	0.01	0.04	0.04	0.32*	1				
PHR	0.18	-0.12	0.41*	-0.04	-0.04	1			
UN	-0.30*	0.12	-0.19	0.02	0.03	-0.52*	1		
LIT	-0.35*	0.24*	-0.40*	0.01	0.01	-0.41*	0.26*	1	
DR	0.11	-0.25*	0.27*	0.15	-0.002	0.46*	-0.41*	-0.57*	1

Note: See appendix table 1 for variable definitions. * indicates that correlation coefficients are significant at 5% level.
Source: Author's calculation

Appendix Table 6: Pair-wise correlation coefficient of the variable used in panel data model for rural state/UTs level

	HH	UNP	PHR	GINI	TFR
HH	1				
UNP	-0.02	1			
PHR	0.27*	-0.23*	1		
GINI	-0.25*	-0.04	0.01	1	
TFR	0.62*	-0.28*	0.46*	-0.10	1

Note: See appendix table 2 for variable definitions. * indicates that correlation coefficients are significant at 5% level.
Source: Author's calculation

Appendix Table 7: Pair-wise correlation coefficient of the variable used in panel data model for urban state/UTs level

	HH	UNP	PHR	GINI	TFR
HH	1				
UNP	0.22*	1			
PHR	0.23*	0.09	1		
GINI	-0.19	-0.08	0.09	1	
TFR	0.29*	-0.02	0.10	-0.09	1

Note: See Appendix table 3 for variable definitions. * indicates that correlation coefficients are significant at 5% level.
Source: Author's calculation

Appendix Table 8: Pair-wise correlation coefficient of the variable used in OLS model for aggregate state/UTs level

	HH	GH	OH	UNP	TFR	GINI	SDP	PHR	UN	ROAD	ELE	LIT	DEN	SEX	GER	TAPE	BORE	LATR	BANK	LAP	MOBILE	TWO	CAR
HH	1.00																						
GH	-0.19	1.00																					
OH	0.23	-0.52*	1.00																				
UNP	0.16	0.10	0.05	1.00																			
TFR	-0.11	-0.06	0.05	-0.27	1.00																		
GINI	-0.07	0.11	0.10	-0.15	-0.10	1.00																	
SDP	-0.33	0.67*	-0.57*	0.04	-0.11	0.18	1.00																
PHR	-0.02	-0.62*	0.20	-0.23	0.15	0.05	-0.59*	1.00															
UN	-0.28	0.61*	-0.41*	0.09	-0.02	0.03	0.59*	-0.45*	1.00														
ROAD	0.05	-0.08	0.29	-0.27	0.12	0.26	-0.18	0.09	-0.13	1.00													
ELE	-0.39*	-0.07	0.14	-0.32	0.06	0.00	0.14	0.13	-0.05	0.32	1.00												
LIT	-0.24	0.67*	-0.46*	0.39*	-0.08	-0.01	0.59*	-0.59*	0.59*	-0.26	-0.20	1.00											
DEN	0.11	0.13	-0.16	-0.11	0.02	0.29	0.21	-0.11	-0.07	0.12	-0.21	0.16	1.00										
SEX	-0.06	-0.09	0.59*	0.15	-0.05	0.16	-0.22	0.00	-0.20	0.11	0.24	-0.03	-0.12	1.00									
GER	-0.34*	0.38*	-0.12	-0.11	-0.04	0.19	0.57*	-0.31	0.13	0.06	0.39*	0.19	0.14	0.27	1.00								
TAPE	-0.20	0.72*	-0.62*	-0.16	-0.05	0.04	0.74*	-0.57*	0.33*	-0.07	0.08	0.43*	0.19	-0.23	0.62*	1.00							
BORE	-0.34*	-0.29	-0.01	-0.18	0.07	0.14	-0.20	0.16	0.06	0.04	0.09	-0.15	0.02	-0.34*	-0.32	-0.23	1.00						
LATR	-0.10	0.54*	-0.40*	0.45*	-0.13	-0.17	0.53*	-0.56*	0.46*	-0.39*	-0.29	0.80*	0.11	-0.09	0.27	0.44*	-0.21	1.00					
BANK	0.12	0.59*	-0.21	0.13	-0.09	0.39*	0.58*	-0.64*	0.32	0.03	-0.39*	0.52*	0.31	-0.17	0.18	0.49*	-0.18	0.39*	1.00				
LAP	-0.29	0.49*	-0.39*	0.16	-0.13	0.24	0.84*	-0.36*	0.45*	-0.17	0.10	0.55*	0.23	-0.11	0.55*	0.51*	-0.24	0.55*	0.48*	1.00			
MOBILE	0.02	0.73*	-0.45*	0.07	-0.01	0.14	0.67*	-0.72*	0.59*	-0.08	-0.26	0.62*	0.31	-0.31	0.40*	0.65*	-0.19	0.60*	0.73*	0.62*	1.00		
TWO	-0.18	0.61*	-0.31	-0.05	-0.20	0.32	0.69*	-0.38*	0.59*	-0.04	0.08	0.42*	0.23	-0.17	0.39*	0.48*	0.06	0.30	0.49*	0.68*	0.67*	1.00	
CAR	-0.20	0.41*	-0.38*	0.07	-0.24	0.22	0.83*	-0.33	0.27	-0.23	0.07	0.42*	0.25	-0.21	0.52*	0.58*	-0.19	0.51*	0.45*	0.92*	0.56*	0.63*	1.00

Note: See Appendix table 4 for variable definitions. * indicates that correlation coefficients are significant at 5% level.
Source: Author's calculation

Appendix

Figure A1: Regression Model 8

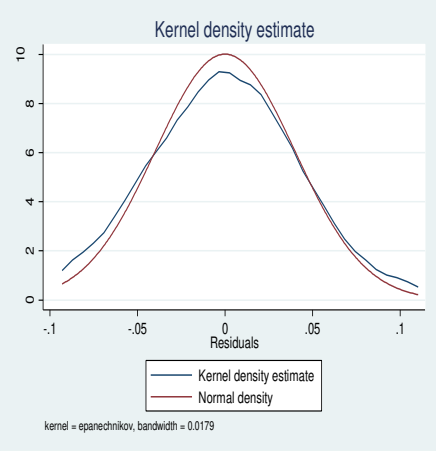


Figure A2: Regression Model 9

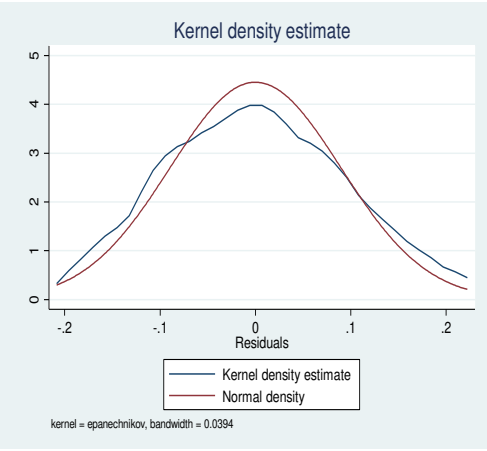
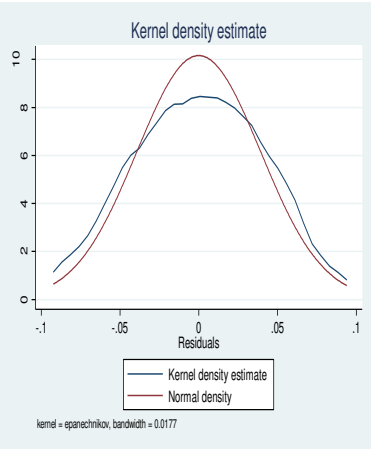


Figure A3: Regression Model 10



Source: Author's estimation

Figure A4:
Histogram for average rural household size

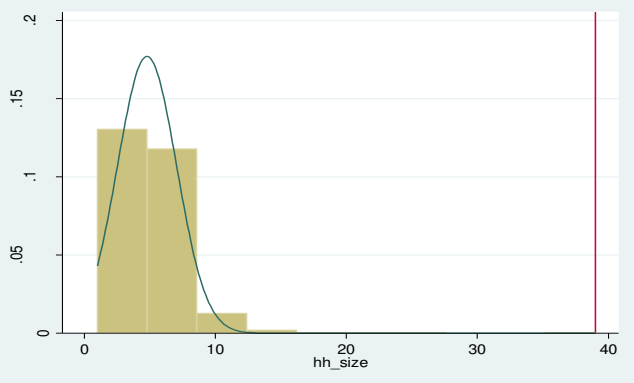
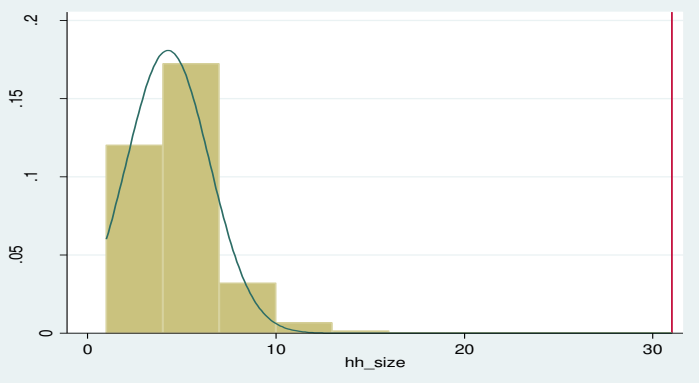


Figure A5:
Histogram for urban average household size



Source: Author's estimation