A District Level Analysis of the Total Fertility Rate Using Indian Census Data

S. Chandrasekhar[1]

Anuja Jayaraman[2*]

Abstract

India is characterized by regions of low fertility-high literacy and high fertility-low literacy rates. We analyze the district level determinants of total fertility rate (TFR) using 2001 Indian census data in order to explain the interregional variation. We find that higher levels of female literacy rate, female work force participation, degree of urbanization and spending by state governments on development lead to lower TFR. Women working as cultivators or agricultural laborers have lower fertility. Following this, we model the change in the TFR over the decade 1991-2001 as a function of change in the following variables: literacy rates, urbanization, occupation choice, availability of health and educational facilities and change in the state government spending on social sector. We control for the change in the population of minorities at the district level, construct a district level index of economic well being and include the same as controls.

[1] # 608 Kern Graduate Building, Department of Economics, The Pennsylvania State University, University Park, PA 16802. E Mail: sxc333@psu.edu

[2] #308 Armsby, Department of Agricultural Economics and Rural Sociology & Population Research Institute, The Pennsylvania State University, University Park, PA 16802.. E Mail: axj902@psu.edu

* Corresponding Author

Introduction

There are sizable variations in the demographic outcomes across the states and districts of India. India is characterized by regions of low fertility-high literacy and high fertility-low literacy rates. The second Indian National Family Health Survey (NFHS-2), conducted in 1998–99, sheds light on this diversity, "There are large variations in fertility among the states in India. States like Goa and Kerala have attained below replacement level fertility and Karnataka, Himachal Pradesh, Tamil Nadu, and Punjab are at or close to replacement level fertility. By contrast, fertility is 3.3 or more children per woman in Meghalaya, Uttar Pradesh, Rajasthan, Nagaland, Bihar, and Madhya Pradesh." For a snap shot of fertility and fertility preferences at the all India level see Table 1.

Hence given this diversity, we propose to undertake a district level analysis of the determinants of the total fertility rate (TFR). The proposed research is similar in spirit to the cross country study of fertility transition by Bongaarts and Watkins (1996). We will also be updating the work by Murthy, Guio and Dreze (1996) who undertook a district level analysis of mortality, fertility and gender bias in India. In addition to modeling the TFR, we also model the change in TFR over the decade 1991-2001 as a function of changes in economic and demographic variables at the district level.

This paper is structured as follows. In Section 1 we briefly discuss the various theories of fertility and also relate our contribution to the literature. In Section 2 we describe the data and in the following section we discuss the variables and the hypothesis. We provide the summary statistics in Section 4. There is a discussion of the methodology in Section 5, which is followed by the section discussing the preliminary results.

Section 1: Fertility Theories

Theories of fertility can be classified either at the macro (societal) level or at the micro (individual) level. One of the earliest macro level explanations is by Malthus. According to the Malthusian theory, population growth follows a geometric progression and the resources sustaining the population also grew at a geometric progression. The balance between resources and population was maintained through alterations in fertility and mortality rates. Another macro level explanation is the theory of demographic transition. This theory attempts to apply the demographic transition experience of the European countries in the 19^{th} century to developing countries. In the first phase of transition, countries experience high level of fertility and mortality rates. In the second phase, developments in technology and medicine lead to a steep decline in the mortality rates resulting in high rate of growth of population. The third and final phase of the transition is marked by low rates of both fertility and mortality rate. The European experience cannot be generalized to other developing countries because of differences in initial conditions.

Leibenstein formulated the first micro-economic theory relating income and fertility decisions. Under this framework parents make decisions on whether to have children by weighing the utility and disutility of having additional children. Becker extended this model by arguing that parents view children as consumer durables and parents might want to have better quality children when their income increase. Quality of children is defined in terms of investment on children. With an increase in income, demand for children decline. This is replaced by greater investment in children in terms of education.

Easterlin (1975) added the supply component (number of children parents would have if they did not use methods to limit fertility) and costs of fertility regulation to the economic theory of fertility theory. According to this theory there would be motivation to control fertility only if the supply exceeds the demand give costless fertility regulation.

Becker, Murphy and Tamura (1990) in their theoretical contribution have addressed the issue of why some regions have grown more rapidly than others. In their growth model, fertility choice is endogenous and so is investment in education. One conclusion of their paper is the following, "Societies with limited human capital choose large families and invest little in each member; those with abundant human capital do the opposite. This leads to two steady states. One has large families and little human capital and the other has small families and perhaps growing human and physical capital."

The impact of economic development and modernization on fertility behavior came to be referred as classical demographic transition theory. Bongaarts and Watkins (1996) in their research find a weak relation between development and reproductive behavior. After analyzing the trends in fertility and indicators of social and development for 69 developing countries between 1969 and 1990, they suggest that diffusion of information about birth control methods is an important mechanism of fertility change. According to Bongaarts and Watkins, diffusion refers to the mechanism by which innovation spreads among regions, social groups and individuals. This spread of information is independent of social and economic changes and is through social interaction. Social interaction includes exchange of information, joint evaluation of the meaning of the idea and finally social influence that would lead to action or no action on part of the individuals. Channels of interactions could be on a local or day to day

individual basis, a national level in terms of migration or better transportation system or on a global basis via world level organizations or multinational organizations.

The diversity in Indian demographic features compares with the results of Becker Murphy and Tamura. At the one end of the spectrum is Kerala while at the other end of the spectrum are the states of Bihar, Madhya Pradesh, Rajasthan and Uttar Pradesh.

Hence given this diversity, we propose to undertake a district level analysis of the determinants of the total fertility rate. The proposed research is similar in spirit to Bongaarts and Watkins' (1996) cross country study. They regressed the TFR on among others the gross domestic product, infant mortality, life expectancy, literacy rate, percent urban, percent labor force in agriculture. They found that the variables that were significant were infant mortality, life expectancy and percent labor force in agriculture. As mentioned earlier, we will also be updating the work by Murthy, Guio and Dreze (1996) who undertook a district level analysis of mortality, fertility and gender bias in India. They found that the only variables having "a significant effect on fertility are female literacy and female labor force participation".

Our contribution will complement the research output of the studies using the Indian National Family Health Survey. On the one hand, while the analysis at the household level helps to uncover the household level characteristics affecting fertility rate, there are many community level unobservables that these studies fail to capture. On the other hand, an analysis of the total fertility rate at the district level, while missing the nuances captured by the household studies, would shed light on the key determinants of interregional variation in fertility outcomes.

In this proposal we model the TFR calculated from the Census 2001 data as a function of the following district level variables: male and female literacy rates, degree of urbanization, and the distribution of women according to their primary occupation. We find that higher levels of female literacy rate, female work force participation, degree of urbanization and spending by state governments on development lead to lower TFR. Women working as cultivators or agricultural laborers have lower fertility.

In the final version of the paper we will merge the Census 1991 and Census 2001 data in order to model the change in TFR over this decade. We will model the change in TFR as a function of changes in the following district level variables: male and female literacy rate, urbanization, presence of minority groups, occupation choice of working women, availability of schools and health clinics per 1000 population and each state governments spending on education and health. In addition we will use the following variables to control for characteristics of the state: the percentage of population having a bank account, the percentage of population having a separate kitchen in the house, the percentage of population having a bicycle and state dummies.

Over the years theories in fertility have shifted emphasis from attributing the fall in TFR solely to economic development to including social and cultural factors. Our analysis will help us identify the key economic, social and cultural factors that have contributed to the decline in TFR and there by help identify potential policy interventions.

Section 2: Data

We use the data available as part of the census 2001 on 594 Indian districts. We use the state and district level estimates of total fertility rate constructed by Guilmoto and

Rajan (2002). As part of Census 2001, district level information is available on the literacy rates, female work force participation rates, female and male occupation choice and urbanization. Presently district level data is not available on social heterogeneity (i.e. the proportion of population belonging to the minority groups viz. the scheduled castes and scheduled tribes) and percentage of population owning different assets.

The Registrar General of India is scheduled to release data on these variables at the district level by the end of this year. We will include the information on minorities and also construct an index of economic well being using principal component analysis based on the percentage of population owning each assets district wise. These will be used as control variables.

In order to compare the change in TFR, we will use the 1991 census data and the fertility rates calculated by Bhat (1996) who used the reverse survival technique method for calculating TFR.

Section 3: The Variables and the Hypotheses

Based on the past studies and the reports of the Indian National Family Health Survey one can argue that the following variables impact the fertility rate

Literacy: The male literacy rate and the female literacy rate are key determinants of TFR. Of these female literacy rate has been deemed to be more important given the findings of earlier studies.

Urbanization: The higher the level of urbanization (the percentage of population in the district living the urban areas) the lower should be the TFR.

Women in the Work Force: Ideally one should include female work force participation rates. In a recent paper based on a study of a village in Tamil Nadu, India, Nakkeeran (2003) based on a recent study in Tamil Nadu states, "Women's work plays a significant role in reducing gender inequality and is also seen to affect levels of fertility and child mortality". Sarangi (1998) dwells on issues relating to female autonomy and fertility. Higher the employment opportunities for women, lower would be the fertility rate. However this could be an endogenous variable since family size and the decision to participate in the work force could be jointly determined. Hence this variable should ideally not be included in the regression analysis. One can argue that the percentage of population in the district living the urban areas partially accounts for this variable since job opportunities are higher for literate women in urban areas. In the next version of the paper the change in TFR will be modeled also as a function of the change in the female work participation rates during the decade. The percentage change in work force participation rates is a good instrument for current level of female work force participation. In this version, we instrument for female work force participation rates by the distribution of women workers according to occupation.

Poverty: As an independent variable we include the percentage of population below the poverty line in the year 1991. It can be expected that the higher this percentage is, the higher would be the TFR. This variable is available as part of the India Human Development Report.

State Government Initiatives: In order to capture the impact of each state government's initiative in the social sector, we include as an explanatory variable the share of developmental spending[1] in total public expenditure. We have information for the year

1991 on availability of health and educational facilities at the district level. Soon data for the year 2001 will be released at the district level on these variables and we can include these in alternative specifications.

Social Heterogeneity: We need to include the percentage of households from the minority groups as an explanatory variable. The scheduled caste and scheduled tribe households are documented to have higher fertility levels. The NFHS Report states, "In India, rural women and women from scheduled tribes and scheduled castes have somewhat higher fertility than other women, but fertility is particularly high for illiterate women, poor women, and Muslim women." Once this data is available we will include this variable in the regression analysis.

Infant Mortality Rate: We do not include the infant mortality rate (IMR) as an explanatory variable since it could potentially cause problems relating to endogeneity. If one has to include the IMR as an explanatory variable we need to estimate the TFR and IMR jointly. Alternatively one can include the change in the infant mortality rate as an instrument in the regression. We will include this variable in the final version of the paper.

State Level Variables: We can potentially include the following variables at the state level: the human development index, the human poverty index and the gender disparity index[2]. While the gender disparity index will capture the opportunities available to women it will also be highly correlated with the female literacy rate. So using these variables at the same time in a specification is not warranted. Similarly the human poverty index is highly correlated with the percentage of population below the poverty line there by mitigating the need to use both these variables in the same specification.

District Level Indices: As mentioned earlier, once the district level information on asset ownership is released we will construct an index of economic well being using principal component analysis based on the percentage of population owning each assets district wise.

Section 4: Summary Statistics

We construct measures of correlation between the state level estimates of the TFR and the percentage of households owning different kinds of assets and having some basic facilities. We find evidence that there is evidence of negative correlation between the percentage of households in the states possessing assets or having access to basic amenities and the TFR. (Table 2).

The Indian Government has for long been using electronic and other mass media to promote family planning. Thus the ownership of television and radio in a large segment of the population can possibly have a salutary impact on the TFR. For instance we find evidence for the following. The correlations between TFR and percentage of households having a bank account, owning a radio or television are negative and significant. Similarly, the correlations between TFR and percentage of households having tap as the source for water or having electricity are negative and significant. These correlations suggest that we need to control for district level ownership of assets in our regression analysis.

At the district level we find that the correlations between TFR and both the male and female literacy rate are negative and significant. Similarly, the correlation between TFR and the proportion of main workers in the female population is negative and significant.

The correlation between TFR and the proportion of population living in urban are is negative and significant (Table 3). These correlations are similar to the results of the papers by Bongaarts and Watkins (1996) and Murthy, Guio and Dreze (1996).

Section 5: Methodology

As a first step we undertake an analysis of the district level estimates of the TFR. We estimate a reduced form equation where we regress the TFR on a set of variables including male and female literacy, urbanization, occupation choice of women and men and state dummies (See Table 4).

Model 1a (Determinants of District Level TFR) Estimated

TFR = F (male literacy, female literacy, % women working, % population living in urban areas, state governments developmental spending as percentage of total expenditure, % households below poverty line in the state, state dummies, constant)

Model 1b (Determinants of District Level TFR) Estimated

TFR = F (male literacy, female literacy, urbanization, % population living in urban areas, state governments developmental spending as percentage of total expenditure, % households below poverty line in the state, occupation choices of working women, state dummies, constant)

In Model 1b, we instrument the percentage of women who are part of the work force (since this is potentially an endogenous variable) by the occupation profile of working women.

Model 2a (Determinants of District Level TFR) Still to be Estimated

TFR = F (male literacy, female literacy, urbanization, % population living in urban areas, state governments developmental spending as percentage of total expenditure, % households below poverty line in the state, % women working, index of asset ownership, presence of minority groups, state dummies, constant)

Model 2b (Determinants of District Level TFR) Still to be Estimated

TFR = F (male literacy, female literacy, urbanization, % population living in urban areas, state governments developmental spending as percentage of total expenditure, % households below poverty line in the state, occupation choices of working women, index of asset ownership, presence of minority groups, state dummies, constant)

Except for the state governments spending, all explanatory variables used in Models 2a and 2b will be at the district level. We will replace the percentage of households below poverty line by a district level index of asset ownership. The percentage of minorities in the district population will be included as an independent variable.

Model 3 (Determinants of Changes in District Level TFR) Still to be Estimated

As mentioned earlier the final version of the paper will estimate not only the determinants of the TFR but also the determinants of changes in TFR.

TFR = F (Δmale literacy, Δfemale literacy, Δurbanization, Δ% population living in urban areas, Δstate governments developmental spending as percentage of total expenditure, Δ% households below poverty line in the state, Δ% women working, Δoccupation choices of working women, Δindex of asset ownership, Δpresence of minority groups, Δinfant mortality rate, Δ in health and educational infrastructure facilities (primary health center, schools etc), constant)

We use the symbol Δ to denote change over the period 1991-2001. We will instrument for the percentage of women who are part of the work force by the percentage change in the women being part of the work force during the decade 1991-2001 and the change in the occupation profile of women. We will also include the change in the district level infant mortality rate.

The objective of this specification is to find the key variables that contribute to the decline in fertility. In particular we will examine the impact of improvements in availability and access to health and education facilities.

Section 6: Preliminary Results

The preliminary results of our regression analysis (Model 1) reveal the following (Table 5). While we tried many different specifications we report the results on only two of them. The female literacy rate comes out significant; higher levels of female literacy lead to lower TFR. Male literacy is not significant. We find the coefficient on the variable URBAN (% of population living in urban areas) is negative and significant. The higher the percentage of population below the poverty line in a state, the higher is the fertility rate in each of the states districts. We also find that higher levels of spending by state governments on development lead to lower TFR. We find that higher rates of female work force participation lead to lower TFR. In order to control for the possibility of endogeneity we instrument this by the distribution of working women according to occupation. We find that women working as cultivators and agricultural laborers having lower fertility rates. When we include female work force participation rates and the occupation choice as independent variables, we find that occupation choice does not

mater. In the final version of the paper we will be estimating Models 2 and 3. These specifications will help us untangle what drives changes in TFR. We will be able to answer the question on what percentage of the change is attributable to changes in each of the following variables: literacy, social heterogeneity, level of economic development of the district, urbanization, occupation choice, availability of medical and educational facilities and spending by the state governments on health and education. We will run various specifications in order to test the robustness of our findings. The conclusions of the paper will help arriving at some stylized facts, which could potentially complement the studies using household data. By identifying the variables that contribute most towards decline in fertility we can come up with some policy prescriptions for state governments.

References

Becker Gary S.; Kevin M. Murphy; Robert Tamura (1990) Human Capital, Fertility, and Economic Growth, The Journal of Political Economy, Vol. 98, No. 5, Part 2: The Problem of Development: A Conference of the Institute for the Study of Free Enterprise Systems, Oct., 1990 pp. S12-S37

Bhat, P N Mari (1996): 'Contours of Fertility Decline in India: A District Level Study on the 1991 Census' in K Srinivasan (ed), Population Policy and Reproductive Health, Hindustan Publishing Corporation, New Delhi, pp. 96-179.

Bongaarts, John and Watkins Susan Cotts. 1996. Social Interactions and Contemporary Fertility Transitions. Population and Development Review, Vol. 22, No. 4. (Dec., 1996), pp. 639-682.

Easterlin, Richard A. 1975. An Economic Framework for Fertility Analysis. Studies in Family Planning. 6: pp 54-63.

International Institute for Population Sciences and ORC Macro (2000): National Family Health Survey (NFHS -2), 1998-99, India, Mumbai, IIPS.

Leibenstein, Harvey. 1957. Economic Development and Economic Backwardness. London: Chapman Hall.

Guilmoto C Z , Rajan S I (2002) District Level Estimates of Fertility from India's 2001 Census Economic and Political Weekly February 16, 2002 pp. 665-72

Murthy M, Guio Anne-Catherine and Dreze J (1996). Mortality, Fertility and Gender Bias in India: A District-Level Analysis Population and Development Review, Vol. 21, No. 4, pp. 745 - 782.

Nakkeeran N (2003) Women's Work, Status and Fertility Land, Caste and Gender in a South Indian Village. Economic and Political Weekly September 13, 2003 pp. 3931-39

Sarangi, Mrutyunjay. 1998. Women's Empowerment and Fertility Behavior in Rural India. Ph.D. dissertation. The Pennsylvania State University.

Table 1: Fertility and Fertility Preferences	
Total fertility rate (for the past 3 years)	2.85
Mean number of children ever born to women 40–49	4.45
Median age at first birth among women age 20–49	19.6
Mean ideal number of children	2.7
Percent of women with 2 living children wanting another child	23
Source: Fact Sheet - India National Family Health Survey, 1998–99	

Table 2: Correlation between the Total Fertility Rate and the proportion of households having the following facilities (State Wise)	
Bank	-0.39
Radio	-0.45
Telephone	-0.46
TV	-0.6
Bicycle	-0.04
Scooter	-0.49
Car	-0.34
Water Source Tap	-0.46
Bathroom Within House	-0.43
Electricity	-0.6
LPG	-0.47
Open Drainage	0.16
Kitchen	-0.32
Permanent House	-0.33
Roof Concrete	-0.45
Source: TFR - Guilmoto and Rajan (2002) Other Variables - Census 2001 N=36, 5% Cut off Values, 1% Cut off Values	

Table 3: Correlation Between the Total Fertility Rate and Select Indicators (District Wise)	
% Population Literate	-0.6785
% Males Literate	-0.5986
% Females Literate	-0.7172
% Females Main Workers	-0.2345
% Population in District Living in Urban Areas	-0.3791
% Male Workers as Cultivators	0.4985
% Male Workers as Agricultural Laborers	0.2063
% Male Workers as Household Industry Workers	0.0676
% Male Workers in Other Occupations	-0.5528
% Female Workers as Cultivators	0.148
% Female Workers as Agricultural Laborers	-0.0735
% Female Workers as Household Industry Workers	-0.1877
% Female Workers in Other Occupations	0.0151
Source: TFR - Guilmoto and Rajan (2002) Other Variables - Census 2001	

Table 4: Description of Variables

Variable	Source	Mean	Standard Deviation
Dependent Variable			
Total Fertility Rate (TFR)	Guilmoto and Rajan (2002)	3.29	1.02
Independent Variable			
Male Literacy Rate	Census 2001	75.13	11.25
Female Literacy Rate	Census 2001	53.10	15.53
% Population in District Living in Urban Areas (URBAN)	Census 2001	0.237	0.197
Development Expenditure (1998-99)	India Human Development Report	45.48	9.46
% Population in State Below Poverty Line (BPL)	India Human Development Report	26.43	12.08
% Cultivator of Female Workers	Census 2001	.221	.146
% Agricultural Laborers of Female Workers	Census 2001	.217	.146
% Household Industry Workers of Female Workers	Census 2001	.0439	.071

Table 5: OLS Estimates - Dependent Variable TFR

Independent Variables	Coefficient	Robust Standard Errors	Coefficient	Robust Standard Errors
Male Literacy Rate	0.004	0.005	0.005	0.005
Female Literacy Rate	-0.033[a]	0.004	-0.034[a]	0.004
Percent Women Working	-1.375[a]	0.370		
Urban	-0.468[a]	0.133	-0.459[a]	0.140
Development Expenditure Ratio	-0.021[a]	0.003	-0.023[a]	0.004
% Below Poverty Line	0.011[a]	0.005	0.010[a]	0.005
% Cultivators of Female Workers			-0.628[a]	0.241
% Agricultural Laborers of Female Workers			-0.345[a]	0.164
% Household Industrial Workers of Female Workers			-0.350	0.320

N = 580, [a] Significant at 1 percent. Constant, State Dummies Not Reported
For definitions of variables see Table 3

[1] Total development expenditure (including social services and economic services) includes social sector expenditure, including expenditure on education, health, amenities (i.e. water supply & sanitation, housing and urban development) and other social services (including welfare of SC, ST & OBC; social security and welfare etc.).

[2] The National Human Development Report 2001 provides a detailed discussion on the construction of these indices. This report is available at http://www.planningcommission.nic.in/reports/genrep/reportsf.htm