Education, Fertility and Incomes in the States of India: Demographic Transition

Mandal, Abir and Regmi, Narendra and Tamura, Robert

University of Mount Olive, University of Wisconsin at Whitewater, Clemson University

11 October 2021

Education, Fertility and Incomes in the States of India: Demographic Transition*

Abir Mandal, Narendra Regmi, Robert Tamura†

October 2021

Abstract

Original decadal estimates of real output per worker, schooling per worker, mortality risk and total fertility rates for states of India covering 1951 to 2011 are produced. An intergenerational model with precautionary demand for fertility is used to fit the observations of fertility and schooling at the state level. The intergenerational human capital model is shown to explain about 75% of log level differences, 100% of average growth rates and 40% of the variation of growth rates across the states of India. These are all improvements relative to a standard Mincer human capital model of schooling and experience returns. The data covers the demographic transition of the states of India from high fertility, total fertility rates of 6, to low fertility, total fertility rates of 2.5.

1 Introduction

We present measures of decadal real output per worker at the state level for India, covering 1951-2011. In addition we provide years of schooling per worker in the states of India over the same period. Finally we produce decadal estimates of mortality risk and total fertility at the state level for India, covering 1951-2011. Using a model with precautionary demand for children, we fit the fertility and schooling data for these Indian states. The resulting state intergenerational human capital stocks are shown to explain about 75% of the log level differences in output per worker, 100% of average growth rates of output per worker and 40% of the variation in growth rates of output per worker in the states of India. Covering the demographic transition for most states of India, average fertility falls from 6.0 in 1951 to 2.3 in 2011.[1]

*We thank Aspen Gorry, Gerald P. Dwyer, Michal Jerzmanowski, Peter Klenow, Chad Jones, Kevin M. Murphy, Curtis Simon for helpful comments and suggestions. We also thank the macro workshop participants at Clemson University, University of Kansas, University of Mount Olive.

†University of Mount Olive, University of Wisconsin at Whitewater, Clemson University, corresponding author rtamura@clemson.edu

[1] While not examining the connection of international trade exposure for each state in India, we note that total fertility rates in 1981 in the states of India were all above 2.50, an average of 4.34 and five states above 5.0. This is the time period examined by the seminal paper by Galor and Mountford (2008). Clearly India was in transition due to economic reforms beginning in 1990, as well as in the middle of the demographic transition. We specifically encourage future research using state level data and penetration of international trade by state of India to extend Galor and Mountford (2008).

India is the world's third largest economy (on a purchasing power parity (PPP) basis), but houses the world's largest population of poor people. The enormity of these numbers implies that the story of India's economic transition to a modern regime bears examination if we want to describe the decline in world poverty levels. Since the time the country was granted Independence by the British, in 1947, many agencies responsible for the collection and archival of macroeconomic data have been formed and dissolved. However, there exists a lack of standardization and adequate time series, which is essential for the study of the relationship between long-run economic growth and the role of inputs.

This is especially true at the state-level, where a constitutionally mandated federal system has led to a multitude of bureaucracies, each of varying (and usually low) effectiveness and output. As a result, the few historical subnational data collated are not suitable for economic analysis, due to lack of standardization and lack of clarity on the methodology used. This paper makes the following contributions: (1) introduces original annual measures of average years of workforce schooling across the states of India, from 1951 (formation of the republic) through 2011, (2) constructs estimates of original real output per worker by state for the same period, (3) estimates the returns to schooling as a robustness check, (4) uses mortality driven fertility decline to explain rising educational levels using an endogenous growth and fertility model created by Tamura (2006) and Tamura, Simon and Murphy (2016), (5) conducts growth accounting, variance decomposition of growth and development accounting across the states of India. We find that the standard Mincer human capital augmented labor input, as in Klenow and Rodriguez-Clare (1997) and Hall and Jones (1999), explains less than half of the log level differences across states, less than half of the average growth in real output per worker across the states, and about one third of the variance of growth across the states. In contrast, the intergenerational human capital model explains over three fourths of the log level differences across the states, all of the average growth rate output per worker across the states, and about 40% of the variance of growth across the states.[2]

To accomplish all of these, we draw from a number of sources. We refer to multiple volumes of the Indian decennial census, starting from 1951 through 2011 to obtain population, workforce and fertility data. Mortality rates were obtained from actuarial reports created by government statisticians a few years after each census. Gross state product data came from the now-defunct Planning Commission of India's archives, which contain data collected as a part of the Soviet-style Five-Year plans, formulated in 1950 by the first prime-minister, Jawaharlal Nehru.

In theory, the administrative setup created by the British (and inherited by the first indige-

[2]Our development accounting results show that intergenerational human capital is quite important for explaining log level differences across the states. This is similar to that found in Erosa et al (2010). In the variance decomposition of growth, the intergenerational human capital model can explain over 50% of the cross sectional variation in growth rates three of eight cases, and 49% in a fourth case. A direct comparison with the Mincer model of human capital, the intergenerational human capital model outperforms in three of four cases.

nous government) should make quantitative delineation of the economic history of India and its constituents easy. However, in practice, political and bureaucratic hurdles make the task complicated. First, at the time of independence, the country consisted of fewer states and a large number of kingdoms of varying sizes. The administrative machineries and consequent data collection mechanisms were heterogeneous. Even after these kingdoms were brought into the union as states, political demarcations within India have changed several times, with the most recent state, Telangana, carved out of the erstwhile undivided Andhra Pradesh as late as 2014. As a result, the existing official data sources do not provide a common base across which to analyze the present-day states in the nation. Third, the data collection process has not been standardized or consistent across time and territories. Finally, large periods of political unrest and strife led to suspension of even these very imperfect data collection efforts. As examples of the last point, we cite the ongoing Islamic violence in the northernmost state of Jammu and Kashmir and the 1975 nationwide state of emergency declared by prime minister Indira Gandhi.

As a result of these challenges, a number of assumptions were made while constructing the dataset, while also interpolating for missing or implausible data. Despite these constraints, this is a much needed first step to come up with a dataset that is actually amenable to further econometric research. While the dataset we construct contains information on the newest and smallest states, in the interest of accuracy and due to the lack of adequate data, they are not included in our analysis. However, the 18 states that we observe output per worker contain approximately 93% of India's total population and covers all the regions of the country. Furthermore the 16 states that we observe with both output per worker and fertility contain approximately 90% of India's total population, and also comes from all the regions of the country. Therefore, our analysis should represent the subnational distribution of our measures reasonably well.

2 India in 1951 and in 2011

Here, we provide an overview of how the political demarcations within India have changed between 1951 and 2011 to give perspective into the creation of a uniform dataset that accurately describe over time the states as they exist today.

India gained its independence in August 1947, as a loosely knit amalgamation of kingdoms and princely states, which were given the option of ascension to India or Pakistan, or to exist as independent states. The union of the states acceding to India was formalized via the 1950 ratification of the Indian Constitution. Figure M1 provides the administrative make up of the country upon ratification. At that time, there were 29 administrative units in the country, out of which only 17 were governed via the legislature and bureaucracy. The remainder were semi-autonomous princely states, the royal lineage of which were given substantial powers. With the 1956 *States Reorganization Act*, this changed. The Act incorporated all the states under the constitutional authority of the central government, merging and carving up territories on linguistic and cultural lines. At that time, Andhra Pradesh was created by merger of the Hyderabad kingdom with Andhra State. Kerala and Karnataka (initially called Mysore) were also created. While the Bombay state was augmented with Saurashtra and Kutch, the *Bombay Reorganization Act* of 1960 split up the state into Maharashtra and Gujarat. Gujarat comprised mainly of the erstwhile Saurashtra and Kutch. The Punjab and East Punjab states were reorganized via the *Punjab Reorganization Act* of 1966, which created Haryana, formalized the modern Punjab state and enlarged Himachal Pradesh. Tamil Nadu was created out of the Madras State in 1968. Chhattisgarh (out of Madhya Pradesh), Uttaranchal (later named Uttarakhand, out of Uttar Pradesh) and Jharkhand (out of Bihar) were all created in 2000. The newest state, Telangana, had a violent creation, amidst riots and murders, in 2014, see Srikanth (2013) for background. The map in Figure M2 represents the present-day India. Table T1 in Appendix II lists all the present day states of India with their respective dates of formation.

For the purpose of data description, we have divided the country into five regions. Fortuitously, the states in each region share similar geography, so as to eliminate associated disparities within each group. Table 1 provides an overview of the states we have analyzed and their respective regions in the country. Note that, while we have listed Telangana as a separate state, since it was created in 2014, we have considered it only as a part of its parent state– Andhra Pradesh– for the purpose of our analysis. Table 1 includes whether we observe output per worker and fertility. For the empirical analysis of the states of India, we only examine those 16 states that have information both on output per worker and fertility.

The rest of the paper is structured as follows. The next section describes the creation of data describing average years of schooling in the work force. We also present an overview of our findings broken down by region graphically and in tabular form. Then we present our estimates of state output per worker. Subsequently, we test the created measures by estimating returns to schooling and comparing them with generally seen values. Then we use decreasing probability of young adult deaths as the explanation for the decrease in fertility and increase in average years of schooling in the workforce. We propose an endogenous fertility model as a framework for such a phenomenon. We calibrate this model to generate estimates of fertility and average education to fit our data. Lastly, we discuss the broader implications of our findings.

3 Education in India

Similar to Turner, et al. (2007), we believe that in order to study the link between human capital and income, measuring the average education in the labor force is more pertinent than the

Table 1: Regions of states of India

State	Region	Output	Fertility
Andhra Pradesh	South	yes	yes
Assam	East	yes	yes
Bihar	East	yes	yes
Delhi	North	yes	yes
Gujarat	West	yes	yes
Haryana	North	yes	yes
Himachal Pradesh	North	yes	no
Jammu and Kashmir	North	yes	no
Jharkhand	East	yes	no
Karnataka	South	yes	yes
Kerala	South	yes	yes
Madhya Pradesh	Central	yes	yes
Maharashtra	West	yes	yes
Odisha	East	yes	yes
Punjab	North	yes	yes
Rajasthan	North	yes	yes
Tamil Nadu	South	yes	yes
Telangana	South	no	no
Uttar Pradesh	North	yes	yes
Uttarakhand	North	yes	no
West Bengal	East	yes	yes

entire population. The Census Volumes entitled “Workers and Occupations,”introduced in 1961, enumerates workers based on their educational achievement.

With respect to the census of workers, the classifications of workforce education levels are not consistent year to year. For example, the 1971 Census lists workers with primary education (4 years and no higher), while the 2001 Census lists workers who are “Matric/secondary but below graduate.” The challenge in our study was to attribute an appropriate number of years of schooling to each classification. Furthermore, we had to interpret whether the term, “worker,” meant the entire labor force, or the currently employed, and if so, only those fully employed, or temporary workers as well, so as to construct a uniformly measured set of data. For 1951, worker data were not available; however, state populations were divided by educational levels. We assume that the labor force participation rate within each educational level in 1951 was the same as in 1961 to construct our average years of schooling for that year.

We constructed average years of schooling for state i in year t in the labor force for each census year using the following formula:

$$E_{it} = \sum_{i=1}^{n} P_{it}.yrs_t$$

where yrs_t is the number of years corresponding to a given level of education in year t, while P_{it} is the share of total workforce with yrs_t years of education in state i in year t.

There is a structural break in the way we handled the “literate” classification. Prior to 1991, the Indian government defined the term as someone able to read and write his or her own name. Subsequent to 1991, the government released a more rigorous definition of the term, asserting that a literate person “should be able to read and write and perform arithmetic operations sufficiently so as to function properly in society and execute contracts” (NLM).

Before 1991, we assign 0 years of schooling to people with only literacy and no other formal schooling, while for 1991 and beyond, we assign three years of education to workers classified as just literate. We have to caution here that the classification of literate or illiterate by the Indian government seems to be functionally irrelevant. Referring to five states– Uttar Pradesh, Madhya Pradesh, Bihar, Rajasthan, and Gujarat, Kothari, et al. (2004) found that while 68.2% claim to be literate in the sample, only 12% among them could read an assigned paragraph with conviction, while 36.3% had reading difficulties. A majority, 51.7%, could not read at all! In fact, only 37.5% could even write their full name correctly. A majority, 52%, could not read the bus schedule, critical to people moving around in the absence of private transport, 56% could not read a newspaper, 54.8% could not read letters, and 56.7% could not write a letter. Table T2 in the Appendix provides an overview of the illiterate proportion of population in each state reported in the Census.

Table 2: Average years of education in workforce: leaders in **bold**

	1951	1961	1971	1981	1991	2001	2011
Central	0.66	0.74	1.58	1.88	2.61	4.70	6.27
East	1.16	1.18	2.29	2.84	3.45	5.07	6.82
North	0.91	**1.85**	**2.95**	**3.54**	**4.84**	**6.28**	**8.04**
South	**1.17**	1.34	2.68	3.14	3.86	5.69	7.75
West	1.07	1.71	2.67	3.36	4.36	6.19	7.98
		India					
Mandal, Regmi and Tamura	1.04	1.46	2.59	3.12	4.02	5.71	7.54
TDDB measure	1.56	2.10	3.33	4.44	5.60	6.79	7.85
UNDP measure				1.90	3.00	4.40	5.40

Table 2 provides our constructed measures of average education by region, weighted by the state workforce.[3] The country as a whole had about one year of schooling in the workforce at the time it gained independence. The northern and central regions of the country lagged behind the others by a significant margin. However by 1961, the North emerged as the highest schooled region, thanks to the capital city of New Delhi, which attracted educated workers. The Central region has remained the least schooled region for the entire 1951 to 2011 period. The workers of the West have generally been the second highest schooled region of India. Despite these trends, the country remains poorly educated, contrary to its image as a technology and outsourcing hub. We note that even in 2011, the average worker possesses less than a high school education. Figure M3 exhibits the average years of education in the labor force across our analysis period by the regions of the country.

For comparison, we include in Table 2, the average years of schooling for India from Tamura, et al. (2019) and *United Nations Development Programme's* own measures of average years of schooling in the entire population from 1980, which unsurprisingly shows the workforce to be significantly more educated than the population at large. The divergence between the average years of education in the workforce versus that in the population in general is about a year in the 1980s, expanding to almost two years by 2011. This suggests a higher skills requirements of the new jobs created in the modern Indian economy. Further evidence on higher schooling of workers comes from Tamura, et al. (2019). In this paper the authors assume higher participation rates for better educated workers. Thus workers exposed to higher education have greater labor force participation rates than those with only high school exposure, and high school exposed workers have higher labor force participation than those with only primary schooling exposure or less. However by 2011, the gap between the schooling measure of this paper with that of Tamura et al. (2019) is barely one third of a year. This is only one fifth of the gap observed in 1991 between our measure and Tamura, et al. (2019).

[3] As note before, we only focus on the 16 states with full economic and demographic information.

Table 3: Maximum schooling gaps between regions and states

	1951	1961	1971	1981	1991	2001	2011
State	1.84	4.13	4.87	5.07	5.77	5.26	5.04
Region	0.51	1.11	1.37	1.66	2.23	1.58	1.77

Table 3 shows the gaps between the states and regions with the highest average years of workforce schooling and those with the lowest. Since 1971, workers in Delhi have about five more years of education on the average than their counterparts in Bihar. Both gaps have increased since independence, owing to a vast differential in enrollment rates. For example, in 1951, Gujarat (West) had a primary enrollment rate of almost 55% and a secondary enrollment rate of around 15%. In comparison, Rajasthan and Haryana (North) had only 15% of its age appropriate citizens attending primary school. In 2011, however, primary enrollment rates are 90% or higher in all states of the country but one, so we should see some convergence in the future.

4 State output per worker

This section presents our estimates on state output per worker converted into real 2000 PPP dollars. The Indian government began estimating gross product data for each state in 1961. However, these data are not comparable across time for every state due to the change in political boundaries over time within the country and are not directly useful. In order to construct consistent estimates, we divided the originally provided GSPs by the populations as given in the Census books to obtain the per capita income of each state. We also obtain estimates for labor force participation using the total number of workers in the economics tables of the Census and dividing by the population given in the Census books. Fortunately, the Registrar General of India has released adjusted populations of states from 1951 through 2001 based on the revised boundaries, as of 2006. We multiply the per capita income obtained above to the revised population estimates to get revised GSP for each Census year. Then we obtain the number of workers using the labor force participation estimates and calculated the per worker output by dividing the revised GSP by our measure of state labor force. All amounts are converted to 2000 PPP dollar values using the estimates obtained from the Planning Commission and the Federal Reserve's FRED2 database. Data for 1951 are estimated using the national 1951-1961 real growth rate and assuming that the growth rates of individual states were the same number of standard deviations above or below the national rate in 1951 as in 1961. To obtain an estimate for standard deviation for 1951, we took the ratio of the national average for 1951 to that for 1961 and multiplied the 1961 standard deviation by the same figure.

Table 4: Estimates of Gross State Product per Worker 1951-2011 (PPP$ 2000): leaders in **bold**

	1951	1961	1971	1981	1991	2001	2011	%/year (1951-2011)	%year (2001-2011)
Central	674	1037	1454	2241	3278	2970	4306	3.09	3.72
East	1851	1926	1924	2459	3021	2908	4185	1.36	3.64
North	**2163**	**2509**	**3093**	**3940**	**5676**	**6265**	**11829**	2.83	6.36
South	1400	1587	1851	2032	3046	4164	9032	3.11	7.74
West	1904	2167	2468	3099	4136	5202	10835	2.90	7.34
India	1738	1968	2280	2829	3948	4516	8575	2.66	6.41
Region $\frac{Max}{Min}$	3.21	2.42	2.13	1.94	1.88	2.15	2.83		
State $\frac{Max}{Min}$	6.28	4.63	3.20	4.24	3.74	6.25	8.67		

Table 4 provides our estimates of real output per worker in the states of India aggregated into the five broad regions. The data show that following Independence, the workers in the North and industrial West were more productive than their counterparts in the other regions, a position they have maintained throughout our study period. While we see some evidence of convergence amongst the regions thereafter, the gap increased substantially in the 1991-2011 period. The ratio of the maximum to the minimum regional average declined to 1.9 by 1991, but increased to 2.15 and 2.83, respectively, in 2001 and 2011. While the East in 1951 was the third most productive region in India, by 2011 it was the poorest region in the entire country. This corresponds to six decades of rule of populist economic policies, see Pedersen (2001).

In 1951, the least productive state was the Central Indian tribal and rural Madhya Pradesh. Workers in the state of Delhi were the most productive, followed by those in West Bengal. In 2011, Delhi workers remained the most productive, with an output per worker almost twice that of the next two states– the newly created Uttarakhand and the services hub of Haryana, see Chatterji (2013) who details the rise of the latter comprehensively. On the other end of the scale, we have the eastern states of Odisha and Bihar. By 2011, the state gap is as wide, or wider, than at independence.

Reforms and their results

The reforms enacted by the Indian government, arising from the threat of sovereign default in 1991, under the aegis of then-finance minister and eventual Prime Minister, Dr. Manmohan Singh, are well known. However, less prominent, but nevertheless important reforms took place in the mid-1980s as well, under the prime minister Rajiv Gandhi. We provide a brief overview of the reform process.

Centrally planned industry was the hallmark of the Indian economy from the very beginning with the introduction of Industries Act of 1951, under which a license was needed to set up a new facility or to expand an existing one. Thus, while ostensibly the Indian government allowed the private sector to exist, it de facto controlled the quantity and location of all investment. When Rajiv Gandhi came into power on a wave of sympathy, after his mother, ex-Prime Minister Indira Gandhi, was assassinated in 1984, he began the deregulation process. His government eliminated licenses for one-third of all industries, Rodrik and Subramanian (2004). The second round of reforms occurred in 1991, under Prime Minister Narasimha Rao, after Rajiv Gandhi was assassinated in Sri Lanka. Some argue that his hand was forced by a balance of payments crisis, subsequent to which the IMF imposed the reforms as a precondition to allowing a line of foreign exchange credit to the country. Industrial licensing was no longer required, except for a small number of industries deemed strategic or polluting. At the same time, foreign direct investment was initially allowed in a few sectors, eventually giving way to broad based liberalization in this policy as well. As a result, FDI increased from USD 129 million in 1991 to a peak of USD 48 billion in 2008, before declining to USD 35 billion in 2014, see Dutta and Sarma (2008).[4] At the same time, tariffs, including those on capital goods and food grains, were reduced. This allowed industrial conglomerates to substitute between labor and capital and introduced competition in the large but already low-profit food market.

Ignoring causality concerns, our data show that not all parts of the country equally benefited from the reforms– in particular the East and Central regions saw a decline in labor productivity, between 1991 and 2001. This can perhaps be explained by the fact that the two biggest beneficiaries of reforms– trade and services– were concentrated mainly in the West and South. The South contains the dynamic IT outsourcing cities of Hyderabad and Bangalore, along with the automotive industrial hub of Chennai (earlier known as Madras). The West contains the financial and trade center of Mumbai (Bombay) and the business-friendly state of Gujarat. The Central and East regions, on the contrary, consist of low value and slow growth sectors of agriculture and mining. Further, we see that productivity growth, even in the regions benefiting most from the reforms, only accelerated in the period post 2001. Our findings are consistent with those of Dougherty et al. (2010). They found that a lack of labor mobility, due to transport and relocation costs, precluded the agricultural sector from experiencing a regime of labor productivity growth and that the majority of growth seen even in the services sector occurred after 2000.

The Central and South regions saw an average annual rate of growth of above 3% in per worker output over the 60 year period, with most of the gains coming in 2001-2011, in which they saw

[4] India's 2017 FDI inflows stood at an all time high of USD 60 billion.

annual growth rates of around 4% and 8%, respectively. By contrast, the Eastern states of Bihar, West Bengal, Orissa and Assam saw a growth rate of 1.4% in per worker output, growing 3.6% over 2001-2011. Our findings are consistent with those of Bhattacharya and Sakthivel (2004), which estimated that the overwhelming share of post-reform economic growth could be attributed to industrial and service-based states, while the states with large primary sectors languished. Figure M4 shows the change in per worker output by region over 1951-2011.

Part of the reason for this disparity is the high growth-states' ability to attract service sector investments. This can be traced back to the empowerment of state administrations after the reforms, prior to which the central government's Planning Commission dictated where investments would go. After the reforms, some states, like Andhra Pradesh and Gujarat, openly embraced free market policies, while others, such as Kerala and Bihar chose the political safety of populist regimes, see Kennedy et al. (2013). The results can be seen in the growth rate data (Table 4). Bhattacharya and Sakthivel (2004) agree with this explanation, as do Aghion et al. (2008), who show that states that enacted pro-industry rather than pro-worker regulations post-reforms benefited from reforms in terms of faster economic growth. In Table T4 we present the data for each state. The final two columns present the growth rate in real output per worker for the entire period, 1951 to 2011, and the last decade, 2001 to 2011. While there is considerable heterogeneity in both the long term growth rates and the last decade growth rates, it is true that every state experienced an acceleration in growth rate of output per worker in the final decade compared with their long term average growth rate.

5 Returns to education

In order to examine the validity of our estimates of education and output per worker, we estimate returns to education to see if we obtain figures reasonable for a country like India. We make some assumptions since we do not have data on other inputs, such as physical capital per worker at the state level. We assume free factor mobility across states, and perfect competition in factor markets. While these assumptions are questionable for a country like India, especially pre-liberalization, there has never been a restriction on cross-state migration of labor. Further, while the economy was planned, the planners were not restricted from freely deciding the allocation of new investment or reallocating existing investment to a new state.

In what follows, we mimic Turner et al. (2007). Consider a model with two factors of production, human capital and all other inputs which we call physical capital. We assume production of a single final output is Cobb-Douglas. Output per worker in state i is given by:

$$y_{it} = A_{it} k_{it}^{\alpha} (\text{human capital})_{it}^{1-\alpha}, \tag{1}$$

where k_{it} is the physical capital per worker and h_{it} is the human capital per worker. To simplify, assuming perfect competition in output market, with final output as numeraire, the representative firm solves:

$$\max_{k_{it},h_{it}} \{A_{it}k_{it}^{\alpha}(\text{human capital})_{it})^{1-\alpha} - r_t k_{it} - w_t(\text{human capital})_{it}\} \tag{2}$$

where r_t and w_t are rental rate per unit of capital and human capital, respectively. The firm chooses physical capital in proportion to human capital,

$$k_{it} = \left(\frac{w_t}{r_t}\left(\frac{\alpha}{1-\alpha}\right)\right) \text{human capital}_{it} \tag{3}$$

Substituting (3) back into (1),

$$y_{it} = A_{it}\left(\frac{w_t}{r_t}\left(\frac{\alpha}{1-\alpha}\right)\right)^{\alpha} \text{human capital}_{it} \tag{4}$$

We assume that human capital can be specified in a Mincerian fashion:[1]

$$human\ capital_{it} = \exp(\beta E_{it} + \gamma x_{it}), \tag{5}$$

where E_{it} is the average number of years of education in the workforce of state i in year t and x_{it} is the average number of years of experience for the same.

Workforce survey data were not available to obtain average experience for the states. In order to construct our measure of relative experience, we calculate the average age of the population not enrolled in school and below the age of retirement in India, 60. From that, we subtract the average years of education in the workforce for the state and year and the 6 years it takes children to begin enrollment.

Using this definition of human capital, the earnings regression is:

$$\ln y_{it} = lnA_{it} + \alpha \ln\left(\frac{w_t}{r_t}\left(\frac{\alpha}{1-\alpha}\right)\right) + \beta E_{it} + \gamma x_{it} \tag{6}$$

Assuming that all states have a common level of total factor productivity, we can estimate β by using time dummies in a pooled panel dataset. The results of the regression are given in column 1 of Table 5.

[1]We are not including the quadratic term in synthetic experience for the same reason as in Turner, et al. (2007). There is not enough variation in the data to identify the curvature parameter off of aggregate data.

Under the hypothesis that TFP does not differ across states, i.e., $A_{it} = A_t$ for all i, differencing each states log output per worker from the labor force weighted log national output per worker, years of schooling, and average experience from the labor force weighted averages allows for the estimation of the earnings equation without any time controls. The differenced regression is reported in column two.

We may also consider the case that TFP may vary from state to state or every state can have some traits or institutions unique to it that would affect its labor product, all else being equal. To account for this, we can rewrite equation (6) as:

$$\ln y_{it} = c_i + b_t + \beta E_{it} + \gamma x_{it} \tag{7}$$

In this case, we could use a state fixed effects model, with time dummies, as long as there is no feedback from income to education in a future period. We check whether present levels of income affect future education using a fixed effects panel regression. The results of the regression (see Table T3) suggest the parameter on the future value of education, with respect to present income, is not significantly different than 0, allowing us to consider the state fixed effects model with time dummies. Due to the fact that we only have a maximum of 7 observations per state (T<N), we estimated a dynamic panel regression instead of using a two-way, fixed-effects panel.

The measured return to experience was never significantly different from zero, and in the cases of the dynamic panel and year dummies models, was insignificant. This was likely because the majority of the labor force in India was in the agricultural or other primary occupations, and it is unlikely its productivity would be enhanced by experience greater than one or two years. For those interested, Table T6 in the appendix contains our constructed measures of the average experience across states and time. The results of the dynamic regression are given in column 3.

Table 5: Earnings regressions: Decadal data (Standard error)

E	0.2229***	0.2484***	0.2237***	0.2311***	0.2483***	0.2484***	0.2310***	0.2311***
	(0.0134)	(0.0245)	(0.0121)	(0.0225)	(0.0238)	(0.0245)	(0.0219)	(0.0225)
exp	-0.0138	-0.0075	0.0029	0.0280	-0.0074	-0.0075	0.0275	0.0280
	(0.0141)	(0.0180)	(0.0134)	(0.0180)	(0.0175)	(0.0180)	(0.0174)	(0.0180)
N	112	112	112	112	112	112	112	112
R^2	.7988	.8152	.8440	.8557	0.5466	.5483	.6835	.6854
Year dummies	No	Yes	No	Yes	No	Yes	No	Yes
Region dummies	No	No	Yes	Yes	No	No	Yes	Yes
Differenced	No	No	No	No	Yes	Yes	Yes	Yes

Our results suggest a return of between 22% to 25% of real per-worker output to every year of education. This is the total return to education which also includes the return to physical capital. If labor share is 2/3, then the return to the typical worker is between 15%-16.5% per year of schooling. These are comparable, albeit a bit higher, than those for India in Annex Table 1 in Montenegro and Patrinos (2014). For single years between 1983 and 2009, they have three estimates of about 12.5%, two around 8.5% and one at 7% in returns to additional years of schooling. However for their estimates of returns to an additional year of primary schooling, they report one over 20.5%, one of 15.7%, three around 11.5% and the final one at 5.9%. These are more similar to that found here, as almost all of the states have schooling under 8 years for the bulk of the observations. In fact the number of state-year observations with schooling of less than 7 years is 98 out 112. Of the 14 observations with more than 7 years of schooling 5 have less than 8 years of schooling, and only three have more than 9 years of schooling, Delhi in 2001 and 2011, and Kerala in 2011. This compares with the overall 15% return found in Turner, et al. (2007), or 10% per year of schooling for the typical worker.[5]

Previous research on India mostly examine returns to education on wage income or consumption. Duraisamy (2002) finds, for 1993 wage survey data, private rates of return to education in India increase up to the secondary level and diminish afterwards. At the primary, middle, secondary, higher-secondary and post-secondary levels, he asserts that each additional year of education leads to 7.9, 7.4, 17.3, 9.3 and 11.7%, respectively, of higher income. Vasudeva Dutta (2006) finds the wage premium of post-secondary education to that of primary education widening between 1983 and 1999. He asserts that this is due to the higher skill nature of the jobs created by trade liberalization and reforms of the 1990s at the expense of demand for low skilled workers. In line with this, Agrawal (2012) finds no evidence of declining returns to education in India and suggests higher dividends to post-secondary level education. We cannot comment on such differences between levels since our expected levels of education across all states are so low, and even by 2011, no state had a human capital level equivalent to a high school diploma (12th grade).

[5] In Montenegro and Patrinos (2014), there estimates of returns to additional schooling for the United States for 4 years, 1990, 2000, 2005, and 2010, range from 11.8% to 13.8%, averaging 12.8%

Table 6: Fixed Effects with Lead of Education (Standard error)

E_t	0.2192***	0.1889***	0.1548*	0.0272	0.0382	0.0302	0.0527	0.0310
	(0.0147)	(0.0180)	(0.0826)	(0.0569)	(0.0452)	(0.0395)	(0.0738)	(0.0587)
E_{t+1}					0.1350***	0.1131***	0.1191**	0.0238
					(0.0375)	(0.0328)	(0.0519)	(0.0625)
exp	-0.0025	-0.0290**	0.0264	-0.0015	0.0150	0.0003	0.0150	-0.0078
	(0.0111)	(0.0111)	(0.0192)	(0.0131)	(0.0104)	(0.0150)	(0.0186)	(0.0147)
N	112	96	112	96	96	80	96	80
R^2	.7976	.7965	.7694	.4879	.7221	.7249	.7364	.6206
Year dummies	No	No	Yes	Yes	No	No	Yes	Yes
AR errors	No	Yes	No	Yes	No	Yes	No	Yes

The first column of Table 6 reports the results of standard fixed effects regression. With fixed effects we find roughly 22% return to a year of schooling, and no significant returns to experience. It is possible that the errors are serially correlated, so column 2 in Table 6 presents the results of fixed effects with autocorrelated errors. Returns to schooling are still 19% per additional year, and experience returns are negative 3% per year. Columns 3 and 4 of Table 6 repeat the first two fixed effects regressions, but add decade dummies. Without correction for autocorrelated errors, the return to an additional year of schooling is 15%, but once serially correlated errors are take into account, the return falls to 3% and is insignificant. Accumulation of schooling could respond to expected future income growth, and because income growth can raise schooling attainment, we need to be concerned about fixed effects in the presence of feedback effects, cf. Wooldridge (2002). To test for possible feedback effects, we follow Wooldridge (2002) and add the lead of years of schooling. If future schooling years is significant, we consider this as evidence that contemporaneous innovations in income lead to future schooling attainment. The last four columns of Table 6 repeat the first four model estimates, but with the lead of schooling attainment as an additional regressor. In three of the four regressions, future schooling returns are about 12% and are statistically significant. Only the final regression with decade dummies and autocorrelated errors fails to produce significant returns to schooling, both contemporaneous and future. In all four fixed effects regressions with the lead of education, the return to contemporaneous schooling is never significant, but always positive around 4%.

Since it appears that feedback effects are significant, we again follow Turner, et al. (2007) and use Bond and Blundell (1998, 1999) to reestimate returns to schooling. We rewrite (6) as:

$$lny_{it} = c_i + b_t + \beta E_{it} + \gamma x_{it} + u_{it}$$
$$u_{it} = \rho u_{it-1} + e_{it}$$

This can be rewritten as:

$$lny_{it} = (1-\rho)c_i + (1-\rho L)b_t + \rho lny_{it-1} + \beta(1-\rho L)E_{it} + \gamma(1-\rho L)x_{it} + e_{it} \tag{8}$$

where L is the lag operator. Thus the estimating equation becomes:

$$lny_{it} = (1-\rho)c_i + (1-\rho L)b_t + \pi_1 lny_{it-1} + \pi_2 E_{it} + \pi_3 E_{it-1} + \pi_4 x_{it} + \pi_5 x_{it-1} + e_{it} \tag{9}$$

As in Bond and Blundell (1998), we use differenced and lagged values of the data as instruments in the levels regression. As additional instruments, we experimented with lags of the difference between state $i's$ average educational attainment and the average educational attainment of the other states in the region - this variable may capture the changes in educational attainment related to regional convergence. More specifically, we create the variable:

$$E^c_{it} = \left[E_{it} - \frac{1}{N^R - 1} \sum_{j \neq i}^{N^R} E_{jt} \right] \tag{10}$$

Table 7: System GMM dynamic panel estimates (Standard error)

E_t	0.1594**	0.3235***
	(0.0737)	(0.1000)
E_{t-1}	0.0089	0.0103
	(0.0393)	(0.0653)
E^c_t		-0.1802**
		(0.0733)
E^c_{t-1}		-0.0295
		(0.0607)
lny_{t-1}	0.5844***	0.5585***
	(0.0915)	(0.1036)
exp	0.0362	0.0301
	(0.0254)	(0.0230)
Instruments	1 lag	1 lag, diff-ed
N	96	96
Year dummies	Yes	Yes

Because we only have 16 states and only 7 observations per state, we only used one lag as instruments.[6] Table 7 contains the results of our dynamic panel estimation. Without the region deviation in schooling, the return to schooling is estimated at 16%. When we control for differences between state schooling and the region schooling, the returns are dramatically different, depending on whether the state is below the region average or above the region average. A worker in a state with 1 year more schooling than the regional average schooling produces about 14% in additional output per additional year of schooling. A worker in the same region, but in a different state with 1 year less schooling than the regional average schooling produces 51% more output per additional year of schooling. Of course the average of these is the estimated 33% return to an additional year of schooling. Again these values are similar to the returns to schooling from micro-econometric studies, e.g. Montenegro and Patrinos (2014).

The implication of our results is that a worker with primary (8 years) of education is expected to be about 4.5 times as productive as person with no education. Further, workers with approximately 11 years of education, the highest average amongst the states (Delhi), are expected to be 3 times as productive as those with about 6 years of education, the lowest among the states (Bihar). The data show that workers in Delhi have a per worker output of almost 9 times that of those from Bihar in 2011.

6 Human capital and secular fertility decline

Noting that present income is not a significant predictor of future education, c.f. Table T3, we now turn to the question of why the levels of education in Indian states increased, and why we see the varied time paths that we do. In this section, we present a model of secular fertility decline and increase in average education of workforce in the states of India as an identification exercise. Before we present the model of fertility decline, we present the data on state fertility from 1951 - 2011. These are contained in T7.[7] The demographic transition in India is quite stunning; fertility declined from 6 in 1951 and 1961 to 2.3 in 2011. This decline in fertility coincides the the dramatic rise in schooling in the labor force. Schooling rose from 1 year in 1951 to over 7.5 years of schooling by 2011.

Young adult mortality risk, δ

In order to calculate our measure of young adult mortality risk, we used actual life tables for each state released by the government in the Census reports from 1971 through 2011. For 1951 and

[6] Unlike Turner, et al. (1997) which had 50 states, and roughly 15 observations per state, we only have 112 total state - year observations.

[7] See our Appendix for Indian state fertility.

Table 8: Fertility by Region

	1951	1961	1971	1981	1991	2001	2011
Central	6.47	6.68	5.68	5.24	4.67	4.00	3.20
East	6.41	6.35	5.48	4.67	3.67	3.07	2.50
North	6.06	6.08	5.71	4.67	3.95	2.98	2.43
South	5.20	5.40	4.37	3.60	2.66	2.15	1.85
West	5.98	6.13	5.18	4.02	3.11	2.70	2.19
India	5.92	6.00	5.21	4.32	3.46	2.80	2.31

1961, the government only released life tables by region and, thus, we estimate these probabilities for the states, based on their relative mortalities as compared to the respective regional averages in 1971. We understand that such a method maybe noisy, but we believe the estimates are reasonable, given the smoothness of the actual data for the later years, and the best available in the literature thus far. Using the data available and those that we estimated, we calculated δ as the risk of death before the age of 35 utilizing the Kaplan-Meier method to first calculate the probability of surviving till the age of 35, and then subtracting that value from 1.

For 1971 onwards, we reduced the 0-4 year probability of dying by .17 to account for the fact that an infant death is less costly to replace than the death of older children [8]. This is because with each passing year, the woman loses another year of her child bearing ability as well as the human capital investment made in the child, if he/she dies. The life tables have mortality data for only the entire age range of 35 to 39. To get the probability of death at just 35, we assume a uniform distribution and calculate $P(\text{death at } 35) = \dfrac{P(\text{Death between 35 and 39})}{5}$.

Mathematically, to clarify, for 1971 through 2011, our measure of young adult mortality is calculated in this manner:

$$\delta = 1 - [1 - 0.83P(0-4)][1 - P(5-9)][1 - P(10-14)] \ldots [1 - P(30-34)].[1 - P(35)] \quad (11)$$

For 1951 and 1961, we have data at every single age and so we calculated,

$$\delta = 1 - [1 - .33P(0)] \prod_{i=1}^{35} [1 - P(i)], \quad (12)$$

[8] If we assume that infant deaths are 1/4 of all deaths before 5, and if we wish to weight infant deaths by only 1/3, consistent with Tamura, et al. (2016), Tamura & Simon (2017), Regmi and Tamura (2021) and Tamura and Witham (2021), then we would downweight probability of dying before 5, p(0-4) by 1/6. This is a rough approximation, but data constraints compelled us to do so.

Table 9: Young Adult Mortality by Region

	1951	1961	1971	1981	1991	2001	2011
Central	.6000	.3401	.2626	.2503	.2231	.1776	.1225
East	.5887	.3494	.2725	.2244	.1994	.1517	.1043
North	.4775	.2485	.2071	.1938	.1550	.1227	.0944
South	.3967	.3801	.2247	.1684	.1416	.1067	.0802
West	.5028	.2912	.2169	.1866	.1281	.1193	.0918
avg India	.4926	.3203	.2324	.1965	.1625	.1284	.0944
From Regmi and Tamura (2021)							
India	.4682	.3677	.2742	.2014	.1383	.0985	.0845
USA	.0849	.0688	.0588	.0453	.0367	.0298	.0211

where, again, we weighted the infant mortality rate by a third.

Table 9 contains our estimates of young adult mortality. We report the results by geographic region, where each state's mortality risk is weighted by their labor force size. For the national Indian data row, we aggregated up the states, again weighting by the state's labor force. We also report the separate measures of young adult mortality risk for India, using only national data, and the United States. These are both from Regmi and Tamura (2021). Differences in the national India row from this paper, and that from Regmi and Tamura (2021) arise mostly because we have only 16 Indian states with both output per worker information and mortality and birth information. In contrast Regmi and Tamura (2021) contains national data, including all states that are included here, plus 4 more.

Principally, we identify declining mortality risk as the cause of declining fertility, similar to Tamura (2006), Tamura and Simon (2017), Tamura, Simon and Murphy (2016), Regmi and Tamura (2021), Tamura and Witham (2021). In addition, similar to Tamura, Simon and Murphy (2016), Tamura and Witham (2021), we also add to the model relative differences between states of changes in opportunity cost of land (rent), as different parts of India transitioned to modern economies at different rates.

The model focuses on the parental choices of generation t in country i of fertility, x_{it}, and human capital of their children, h_{it+1}; a composite consumption good, c_{it}; and space per child, S_{it}. Parents respond to the probability of young adult mortality, δ_{it}, in making their choices.

We use the same preferences as in Tamura, Simon and Murphy (2016), Regmi and Tamura (2021), Tamura and Witham (2021) denoted by:

$$\alpha(c_{it}^{\psi}S_{it}^{1-\psi})^{\varphi}[(1-\delta_{it})x_{it}-a]^{1-\varphi}+\Lambda h_{it+1}^{\varphi}\{1-\frac{\beta_{it}\delta_{it}^{\nu_{it}}}{[(1-\delta_{it})x_{it}-a](1-\delta_{it})}\}, \tag{13}$$

where β_{it} and ν_{it} are time varying preference parameters. The parameter ψ determines the split of expenditures between adult consumption and space. The young adult mortality rate is given by δ. We assume that parents only care about the net fertility, given by $(1-\delta)x - a, a \geq 0$. This implies that the elasticity of substitution of net expected fertility with human capital investments is greater than 1, which is the elasticity of substitution between net expected fertility and space as well.[9] The last term, captures precautionary demand for fertility, identical to Tamura, Simon and Murphy (2016), Regmi and Tamura (2021) and Tamura and Witham (2021).[10] The precautionary demand becomes increasingly small as the probability of young mortality falls, and is essentially 0 for developed countries.

The parents face the following budget constraint:

$$c_{it} + r_{it}x_{it}S_{it} = h_{it}[1 - x_{it}(\theta + \kappa_{it}\tau_{it})] \tag{14}$$

where θ is the time cost of raising children, τ_{it} is the time spent educating children, κ_{it} is the inverse efficiency of education time–implying efficiency decreases when κ increases and r_{it} is the price of a unit of space.

As in Tamura, Simon and Murphy (2016), Regmi and Tamura (2021), Tamura and Witham (2021), human capital accumulates via the following technology:

$$h_{it+1} = A\overline{h}_t^{\rho_{it}} h_{it}^{1-\rho_{it}} \tau_{it}^{\mu} \tag{15}$$

$$\rho_{it} = min\{.5, \frac{.5\overline{\tau}_{it}}{.38125}\} \tag{16}$$

where $\overline{\tau}_{it}$ is the average time spent in education in the state and is an external effect of schooling. The fact that $\rho_{it} > 0$ signifies that while schooling is positive, the children can benefit from the existence of higher levels of human capital in the world. The more education society provides on average to its children, the more it can benefit from learning as opposed to innovating and discovering by itself. This effect is maximized at $\overline{\tau} = .38125$, which for a 40 year period occurs at 15.25 years of schooling. This accumulation technology is identical to Tamura, Simon and Murphy (2016), Regmi and Tamura (2021), Tamura and Witham (2021) and is similar to Tamura (1991, 1996, 2006). $\overline{h}_t$ is the frontier human capital and is assumed to be the human capital of

[9] This is similar to Jones (2001).

[10] Precautionary demand for children was pioneered by Kalemli-Ozcan (2002,2003), and used in Tamura (2006), Tamura and Simon (2017), and Mandal (2017). Tamura and Simon (2017) and Mandal (2017) use preferences similar to those here, which are asymptotically identical to those here when $\delta = 0$. However Tamura (2006), Tamura and Simon (2017) and Mandal (2017), Regmi (2019) use a different accumulation technology for human capital. That accumulation technology produces too rapid convergence across regions, in fact a regression of the growth rate of human capital on the growth rate of income is strongly negative in those models.

the United States for all the states in our work.[11]

Substituting (15)-(16) and (14) into (13), ignoring the country subscript i we get the Euler equations determining optimal choices of human capital investments, fertility and space,

$$\begin{aligned} \frac{\partial}{\partial \tau} &: \psi\varphi\alpha c_t^{\psi\varphi-1} S_t^{(1-\psi)\varphi} \left[(1-\delta_t) x_t - a\right]^{1-\varphi} h_t x_t \kappa_t \\ &= \Lambda\mu\varphi A^{\varphi} (\overline{h}_t^{\rho} h_t^{1-\rho})^{\varphi} \tau_t^{\mu\varphi-1} \left(1 - \frac{\beta_t \delta_t^{\nu_t}}{\left[(1-\delta_t)x_t - a\right](1-\delta_t)^{\varepsilon}}\right) \qquad (17) \\ \frac{\partial}{\partial x} &: \psi\varphi\alpha c_t^{\psi\varphi-1} S_t^{(1-\psi)\varphi} \left[(1-\delta_t) x_t - a\right]^{1-\varphi} \left[h_t \left[\theta + \kappa_t \tau_t\right] + r_t S_t\right] \\ &= (1-\varphi)\, \alpha c_t^{\psi\varphi} S_t^{(1-\psi)\varphi} \left[(1-\delta_t) x_t - a\right]^{-\varphi} (1-\delta_t) + \Lambda h_{t+1}^{\varphi} \frac{\beta \delta_t^{\nu_t}}{\left[(1-\delta_t)x_t - a\right]^2} \qquad (18) \\ \frac{\partial}{\partial S} &= \psi\varphi\alpha c_t^{\psi\varphi-1} S_t^{(1-\psi)\varphi} \left[(1-\delta_t) x_t - a\right]^{1-\varphi} r_t x_t \\ &= \alpha (1-\psi)\, \varphi c_t^{\psi\varphi} S_t^{(1-\psi)\varphi-1} \left[(1-\delta_t) x_t - a\right]^{1-\varphi}. \qquad (19) \end{aligned}$$

We can then solve for c_t as a function of S_t and x_t. This produces

$$c_t = \left(\frac{\psi}{1-\psi}\right) r_t x_t S_t. \qquad (20)$$

Substituting this into the budget constraint produces

$$r_t x_t S_t = (1-\psi)\, h_t \left[1 - x_t \left(\theta + \kappa_t \tau_t\right)\right]. \qquad (21)$$

Substituting this back into the objective function yields

$$\max_{x_t,\tau_t} \left\{ \begin{array}{c} \alpha (\psi)^{\psi\varphi} \left(\frac{1-\psi}{r_t x_t}\right)^{(1-\psi)\varphi} \left(h_t \left[1 - x_t \left(\theta + \kappa_t \tau_t\right)\right]\right)^{\varphi} \left[(1-\delta_t) x_t - a\right]^{1-\varphi} \\ +\Lambda h_{t+1}^{\varphi} \left(1 - \frac{\beta_t \delta_t^{\nu_t}}{\left[(1-\delta_t)x_t - a\right](1-\delta_t)^{\varepsilon}}\right) \end{array} \right\}. \qquad (22)$$

The equation shows that we have fertility rate, x, decreasing with a decline in young adult mortality, δ. Due to the interaction of fertility with space cost, r and human capital investments, τ_t, the budget constraint is not convex– implying that we may not have an interior solution and require a numerical method to solve the problem. We solve the problem by taking into consideration that for a given level of fertility, the problem is concave in (c, S, τ). Therefore, we use a grid over possible values of fertility to solve the household's problem, subsequently choosing the fertility that maximizes utility. Our parameters chosen for preferences (α, ψ, ϕ, a)

[11]This accumulation technology differs from the intergenerational human capital accumulation technology in Tamura, Dwyer, Devereux and Baier (2019). In that version, ignoring the life cycle accumulation of human capital of each birth cohort, the next generation's human capital is produced via a decreasing returns to scale technology in the two inputs, external human capital and parental human capital. Asymptotically $h_{t+1} = A h_t^{.69} \exp(.1 E_{t+1})$.

and technology (ρ, μ) produce interior solutions for (x, c, S, τ).[12]

We solve the model annually for each state from 1951 through 2011, producing fertility, schooling choice and human capital for that birth cohort, interpolating δ for the intervening years between censuses. This allows us to compute the stock of human capital in the population. This is used to judge the goodness of fit of the model. Right after independence, with widespread poverty and lack of access to healthcare, the population of India was afflicted with high infant mortality and consequent high young adult mortality rates, defined as the probability of death before the age of 35. This made parents choose high precautionary fertility. However, high fertility implied that imparting their children with education was expensive, and parents chose lower levels of schooling for them. As mortality fell, the precautionary demand declined, leading to lower fertility. At the same time, the increasing opportunity cost of space, as India's population exploded and the country developed, making it more profitable to undertake commerce, industry and agriculture on land instead of utilizing it for dwelling, also made raising children more expensive, leading to lower fertility. However, this meant that the opportunity cost of children quality declined, leading to parents choosing to educate their children more, aided by public policy measures, see Kumar (2004), Ghosh (2000). Our numerical solutions show that the decline in young adult mortality and a dramatic rise in the price of space were the major reasons for the decline in fertility rates across the country. We are able to replicate the pattern of fertility and years of education in the labor force observed over our study period in the states of India.

We also use the parameter κ to produce the appropriate secular rise in human capital investment time. We use our previous-section estimates of years of schooling in the labor force for each state as a measure of τ. We assume that a period is 40 years– therefore, $40\tau_t$ is the years of schooling for typical individual workers born in the year t.

We seed the initial level of human capital, as per data, and then we use the taste parameters, κ and r, to calibrate the model and match the observed rise in human capital across the country.

Numerical Solutions

Here, we analyze the stationary solution to our model and present the numerical solutions. We assume that the stationary fertility rate is 1. With our Euler equation for fertility when mortality risk is 0, we can impose the following restriction on a as a function of parameters and the stationary human capital investment, $\overline{\tau}$:

[12] This is the same solution methodology used in Tamura (2006), Tamura, Simon and Murphy (2016), Tamura and Simon (2017), Mandal (2017), Regmi (2019), Regmi and Tamura (2021) and Tamura and Witham (2021).

$$a = 1 - \frac{(1-\varphi)(1-[\theta+\overline{\tau}])}{\varphi(1-\psi(1-[\theta+\overline{\tau}]))} \tag{23}$$

We also have an implicit function determining the stationary human capital investment rate, $\overline{\tau}$:

$$1 = \frac{\Lambda\mu[Ar^{1-\psi}]^{\varphi}(1-\theta-\overline{\tau})^{1-\varphi}}{\alpha[\psi^{\psi}(1-\psi)^{1-\psi}]^{\varphi}(1-a)^{1-\varphi}\overline{\tau}^{1-\mu\varphi}}, \tag{24}$$

where, under the balanced growth path, $\overline{h}_t = h_t$ and the right hand side value is constant. Under these restrictions and the convergence of mortality risk to 0, the long run fertility rate, x, will be 1 and human capital investment, $\tau = \overline{\tau}$. Table T8 gives the parameter values common to all the states. Most of our parameter choices are standard. The time cost of rearing a child, $\theta = 0.125$, implies a biological maximum fertility of 8 in an asexual model, or 16 in a model with both sexes. Our choice of $\overline{\tau} = .38125$ implies a steady state value of 15.25 years of schooling, consistent with the developed country measures of schooling in Tamura (2006), Tamura et al. (2016), Tamura et al. (2019).

Our choice for (A, μ) is consistent with an annualized balanced-growth path growth rate of 1.8%. Our choice of parameters (θ, ψ), in combination with our calibrated long-run values of fertility and schooling, $x = 1, \overline{\tau} = 0.38125$, and our assumed stationary value of $\kappa = 1$, implies a stationary budget share of housing, S, of 19%. This is the US housing budget share reported in the *OECD Better Life Index*. India has an estimated budget share for housing of between 20% to 21%, see Suisse Group (2017). All other consumption expenditure comes from (20), and is about 37%. Total consumption in our model is therefore around 56% with the remainder towards education. This is obviously much higher than data, but in this model there is no physical capital, so all investment for the next generation is via human capital investment. Consider India's physical investment rate of about 32% (2000-2014 average of WDI and PWT), educational spending of about 4%, *World Development Indicators 2014*, combined public and private health expenditure of 5% and R&D expenses of 1%, *World Development Indicators 2014*, we get a total of 42% of GDP being spent towards the next generation, not considering the opportunity cost of foregone earnings incurred by the student. Also, comparing it to the US next generation share of 48% as calculated in Tamura, et al. (2016), we see that our model steady state rate expenditure on the next generation is in between present data from India and the US. Table T8 summarizes our calibration measures.

Results

Using these data and our solution, we generate data on Total Fertility Rate (TFR) and average years of schooling. We then plot these generated series with plots of actual data for these two measures. The model solutions are generated by solving the Euler equation for τ_t for each fertility value, selecting the fertility and schooling pair generating the highest utility. This method is used so as to allow for the possibility of corner solutions, since the budget set is not convex. For each state i, we allow $\nu_{it}, r_{it}, \kappa_{it}$ to vary in year t to fit the observed fertility and average schooling data as closely as possible.

TFR data were retrieved from the *Fertility Volumes* of the Census Reports for 1961 through 2011. Values were imputed for reconstructed states as well as for the year 1951, for which the government did not collect or release state level data. By 2011, many states had TFR similar to or lower than that of the United States, as a comparison. These include Andhra Pradesh, Delhi, Punjab and West Bengal. Himachal Pradesh had the lowest fertility in that year (1.14). On the other hand, states like Bihar, Madhya Pradesh and Rajasthan had TFR of around 3 or above. The average fertility rate in the country declined from around 6 in 1951 to 2.3 in 2011, see Table T7.

With respect to "rental rate," r, or the price of space, a couple of factors are at play. To clarify, the rental rate is not the absolute price of space, but rather a figure more related to the relative share of income that is allocated to space costs. India is a very rural country even today, but was overwhelmingly rural in 1951. In 1951, 83% of Indians lived in rural areas, Datta (2006), as compared to 67% in 2011 *World Development Indicators*. A rural area would typically have lower population densities and would have a majority of its land mass devoted to low return agricultural and other primary activities (low scale forestry, non-commercial scale mineral gathering etc.). Means of communication, inter-region transport of goods and people and alternate sources of employment were non-existent in 1951 and remain restricted even today. As such, the marginal space cost of another child was negligible on average. This started changing as India industrialized and the service industry became more prominent. The severe land scarcity in India's biggest cities forced the new foreign (owing to a more liberal FDI policy) and domestic companies to expand to the hinterland and communications started improving substantially due to the proliferation of cellphones. As a result, it became more profitable to use land for purposes other than habitation, increasing the opportunity cost of the marginal child. Thus, two forces are at play. On one hand, the increasing opportunity cost of already occupied lands would exert an upwards pressure on r. At the same time, since the fertility rate was significantly above the replacement rate of 2.1 (and remains, on average, slightly higher even today), per capita availability of these lands decreased. On the other hand, better transport infrastructure and affordability would allow land previously unfit for use to come into play, increasing supply and decreasing r. A prime example of this phenomenon are the states of Haryana, overwhelmingly rural and agricultural at India's founding, and Delhi, the relatively very urban capitol. Space constraints in Delhi forced companies to relocate to the erstwhile village of Gurgaon, Haryana, around 40 km away. From its origins as a dusty, semi-arable land mass, Gurgaon is one of the largest information technology and financial services hub in Asia, with some of the priciest real estate in the country. This is reflected via an unconstrained increase in r for Haryana, while the space costs for Delhi reflect both the opposing forces. Such stories have been repeated in almost all states of the country, in varying degree.

Unlike Tamura, Simon and Murphy (2016), while we do not have such state level data across time to illustrate these phenomenon, our calibrated rental rates point to the domination of the increasing alternate use and population density effects, especially subsequent to the first and second round of economic reforms during the 80s and 90s. The exceptions to the exponential increases in rents are Bihar, Rajasthan, and Uttar Pradesh. These three are also among the most poor states of the country. Bihar saw a sharp increase in space costs in 1971, but costs rapidly declined immediately afterwards, with rents ending up essentially at the same level as at the start of our analysis. Rajasthan, as per our calibration, experienced a rise in rent through 1981, followed by rapid decline over the next decade before experiencing growth again subsequently. Uttar Pradesh experienced rental rate growth similar to other states through 2001, when Uttarakhand was carved out of the state. Subsequently, our calibration points to a sharp decline in rent, although it ended up higher than at the beginning. Graphs in figures C1 through C16 show the data and our model solutions of fertility and average years of schooling. After seeding the initial level of human capital, and calibrating values for the parameters, we were able to fit our model to the time paths of fertility and human capital very well.

Goodness of fit and returns to model human capital

In order to verify the estimates for fertility and average education levels generated by our model, we first regress the model estimates on the actual data. If the model is correct, we expect to see a slope coefficient not significantly different than 1. Table 10 provides the regression estimates for goodness of fit. The OLS coefficient for model years of education on data was 0.9871, significant at the 1% level. Further, we were unable to reject the null hypothesis that the coefficient on our model estimate of schooling was equal to 1 at the usual significance levels. While we rejected the null hypothesis in the case of TFR, the coefficient on the model estimate was very close to 1. The final two columns contains the log-log regressions. In the case of education, the log - log specification produces similar results to the levels regression. In particular we cannot reject the null hypothesis that $\beta = 1$ & $\alpha = 0$ at the 5% level. Finally the log - log specification in fertility is again similar to the levels regression. We reject the null hypothesis that $\beta = 1$ & $\alpha = 0$ at the 1% level. However similar to the levels regression, the slope coefficient is pretty close to 1, 1.14.

Table 10: Pooled Regressions of Actual Observations on Model Data

	E	TFR	$ln(E)$	$ln(TFR)$
model solution, β	0.9871***	1.0920***		
	(0.0249)	(0.0188)		
ln(model solution), β			1.0716***	1.1405***
			(0.0310)	(0.0242)
Intercept, α	0.1088	−0.5209***	-0.0921**	-0.2468***
	(0.1082)	(0.0877)	(0.0392)	(0.0356)
N	112	112	112	112
R^2	0.9348	0.9684	.9159	.9529
p($\beta = 1$ & $\alpha = 0$)	.5114	.0000	.0559	.0000
Significant at: ***-1% **-5% *-10% levels				

7 Development Accounting, Growth Accounting and the Role of Human Capital

How well does human capital explain cross state differences in log output per worker, and how well does it explain variations in state growth of output per worker? We show that while a standard Mincer definition of human capital explains some of the cross state income differences, the intergenerational human capital model does a better job of capturing cross state differences in living standards and growth of living standards. Let schooling of generation t be given by $E_t = 40\tau_{t-1}$; x_t is the average experience of generation t, and our intergenerational human capital is given by h_{t+1}, then Mincer human capital of the next generation, and the intergenerational human capital of the next generation are given by:[13]

$$H_{t+1}^{Mincer} = \exp(.1E_{t+1} + .0495x_{t+1} - .0007x_{t+1}^2) \tag{25}$$

$$h_{t+1} = A\overline{h}_t^{\rho_t} h_t^{1-\rho_t} \tau_t^{\mu} \tag{26}$$

For each state we construct the human capital in two ways: (1) using the standard Bils and Klenow (1997) and Hall and Jones (1999) method in (25) where we use the average schooling in the state and the average experience in the state,[14] (2) we solve the model annually and construct annual measures of new generation human capital using (15)-(16), then we average over these solutions for the existing population 15 to 64, controlling for survival probability. We use Tamura, Dwyer,

[13] By convention, parents choose the fraction of their children's time spent in school, τ_t, but this is schooling for the next generation given by: $E_{t+1} = 40\tau_t$.

[14] For each state we estimate the average age of the population 15 to 64 not enrolled in school and subtract from that the sum of the estimate of schooling in the state and 6.

Table 11: Development Accounting: Share Explained by Inputs

Mincer human capital & physical capital	Intergenerational human capital	Intergenerational human capital & physical capital
.4250	.7889	.7641

Devereux and Baier (2019) for estimates of physical capital per worker in India for years 1951-2010. We assume that physical capital is freely mobile across states, and that each state uses the same technology. Under these assumptions, then physical capital per worker will be proportional to the human capital per worker in each state. We thus evaluate the following models of output per worker in state i:

$$y_{it} = Z_{it}k_{it}^{\alpha}(H_{it}^{Mincer})^{1-\alpha} \tag{27}$$

$$y_{it} = \hat{Z}_{it}\hat{h}_{it} \tag{28}$$

$$y_{it} = \hat{Z}_{it}^{k}\hat{k}_{it}^{\alpha}\hat{h}_{it}^{1-\alpha} \tag{29}$$

where $\hat{h}_{it}$ is the average intergenerational human capital in state i in year t.[15] We first conduct development accounting. That is we ask what proportion of the variation in log output per worker is explained by variation in inputs and what proportion of the variation is explained by variation in TFP. We follow Tamura, Dwyer, Devereux and Baier (2019) by using the two extreme assignments of the correlated component (the covariance between log inputs and log TFP) based on theory. The standard neoclassical growth model with exogenous technological change and the endogenous technological change theories imply that all capital deepening arises from the advancement of technology. The endogenous growth theories of Romer (1986), Lucas (1988) and Tamura (2002,2006) have TFP growth arising from input accumulation. Thus we compute the average contribution arising from inputs and TFP as:

$$\overline{S}_{\ln x} = \frac{\sigma^2_{\ln x}}{\sigma^2_{\ln y}} + \frac{1}{2}\frac{\rho^2_{\ln x,\ln z}\left(\sigma^2_{\ln z} - \sigma^2_{\ln x}\right)}{\sigma^2_{\ln y}} + \frac{\sigma_{\ln x}\sigma_{\ln z}\rho_{\ln x,\ln z}}{\sigma^2_{\ln y}} \tag{30}$$

$$\overline{S}_{\ln z} = \frac{\sigma^2_{\ln z}}{\sigma^2_{\ln y}} + \frac{1}{2}\frac{\rho^2_{\ln x,\ln z}\left(\sigma^2_{\ln x} - \sigma^2_{\ln z}\right)}{\sigma^2_{\ln y}} + \frac{\sigma_{\ln x}\sigma_{\ln z}\rho_{\ln x,\ln z}}{\sigma^2_{\ln y}} \tag{31}$$

Table 11 presents the results of these three different measures of inputs. Both the intergenerational human capital without physical capital, and intergenerational human capital with physical capital explain more than the traditional Mincer human capital and physical capital model. The basic Mincer human capital model captures about 43% of the log output per worker variation, quite similar to that found in Hall & Jones (1999), Caselli (2005), Manuelli and Seshadri (2014) The improvement from about 42% to something more than 75% indicates that the intergenera-

[15] In both the Mincer case and the final input case, we assign physical capital per worker in each state in proportion to their human capital, which would occur if physical capital markets were fully developed in India.

Table 12: Growth Accounting: Annualized Growth Rates of Output, g_y, & Share Explained by Inputs, S_x

	Growth Horizon							
	60 years		30 years		20 years		10 years	
	g_y	S_x	g_y	S_x	g_y	S_x	g_y	S_x
Mincer human capital & physical capital	2.56	0.4612	2.59	0.4655	2.58	0.4620	2.74	0.4577
Intergenerational human capital	2.56	1.1866	2.59	1.1847	2.58	1.1822	2.74	1.1372
Intergenerational human capital & physical capital	2.56	1.0545	2.59	1.0588	2.58	1.0533	2.74	1.0323
N	16	16	32	32	48	48	96	96

tional human capital model captures much more of the cross state differences, which is consistent with the augmented human capital model of Manuelli and Seshadri (2014). In both cases with the intergenerational human capital, the development accounting results are similar to the cross country development accounting results in Tamura, Dwyer, Devereux and Baier (2019). Using all years 1790-2010, TDDB (2019) finds that about 76% of the cross country log output per worker variation is explained by the cross country log input variation, nearly identical to the share of Indian cross state log output per worker variation explained by cross state log input variation.

Before we examine the variance decomposition of state growth rates, we first conduct growth accounting. Table 12 contains the results of this exercise. We break up the time period into four different time samples: a single growth rate over 60 years for each state, 2 observations per state over 30 year growth horizons, 3 observations per state over 20 year growth horizons, and finally 6 observations per state over 10 year growth horizons. We can see that the standard Mincer human capital model only explains about 46% of the mean growth in output per worker.[16] Thus the majority of growth arises from TFP growth. By contrast the intergenerational human capital model explains too much, that is mean growth of inputs varies from 114% to 119% of growth, without physical capital, and 103% to 106% of growth with physical capital.[17]

Next we examine the cross sectional variation in state output per worker growth rates. As with development accounting, we use Tamura, Dwyer, Devereux, Baier (2019) to compute the share explained by inputs and TFP by appealing to the two theories of growth. Thus the share

[16] This result is similar to what is found in Tamura, Dwyer, Devereux and Baier (2019) for Asia. TDDB found only about 56% of growth was explained by input growth.

[17] Again by comparison with TDDB, they find that about 83% of mean labor force weighted growth in output per worker arises from input growth. Their results for the Asian countries is similar, with 86% of growth accounted for by input growth.

explained by variation in growth of inputs and the variation in the growth of TFP are given by:

$$\overline{S}_{g_x} = \frac{\sigma^2_{g_x}}{\sigma^2_{g_y}} + \frac{1}{2}\frac{\rho^2_{g_x,g_z}\left(\sigma^2_{g_z} - \sigma^2_{g_x}\right)}{\sigma^2_{g_y}} + \frac{\sigma_{g_x}\sigma_{g_z}\rho_{g_x,g_z}}{\sigma^2_{g_y}} \tag{32}$$

$$\overline{S}_{g_z} = \frac{\sigma^2_{g_z}}{\sigma^2_{g_y}} + \frac{1}{2}\frac{\rho^2_{g_x,g_z}\left(\sigma^2_{g_x} - \sigma^2_{g_z}\right)}{\sigma^2_{g_y}} + \frac{\sigma_{g_x}\sigma_{g_z}\rho_{g_x,g_z}}{\sigma^2_{g_y}} \tag{33}$$

We examine four different data sets of growth in state output per worker. The first is the standard annualized growth rate between 1951 and 2011 for each state. The second is annualized growth rate between 1951 and 1981, and 1981 and 2011 for each state. Thirdly the annualized growth rates from 1951 to 1971, from 1971 to 1991 and from 1991 to 2011. Finally the annualized growth rates for each decade are examined. The results of the three human capital specifications for each of these four data sets are contained in Table 13. There are interesting differences between the standard Mincer human capital model and the two intergenerational human capital models (with and without physical capital). For the Mincer human capital model, the range is between 18% and 44% explained by input variation, and an average of 29%. These results are below that contained in TDDB (2019), where 46% of the variation in the growth rates of output per worker are explained by variations in input per worker growth rates. For the Asian region, TDDB (2019) reports 48% of the variation in growth rates are explained by variation in input per worker growth rates.

By contrast, the intergenerational human capital models (with and without physical capital) do surprisingly poorly in the longest horizon, explaining only 20% of the growth variations. At the 30 year, 20 year and 10 year horizons, the intergenerational models always outperform the Mincer human capital model, explaining between 30% to 65% of the variation in growth rates. While a major improvement, relative to the Mincer human capital model, the 10, 20 and 30 year horizon results are still significantly lower than found in TDDB (2019). The average is 38%, without physical capital, and 44% with physical capital. The results with physical capital, however are quite similar to the variance decomposition results using Mincer human capital in Turner, Tamura, and Mulholland (2013) of 44% for US states 1840-2000, and 46% for US states 1900-2000. It is also similar to the 46% and 48% over the 1800-2010 horizon for countries and Asian countries using the Mincer human capital in TDDB (2019). However in TDDB (2019) the intergenerational human capital model explains 95% of long run variation in output per worker growth, and 94% of long run Asian output per worker growth. On the other hand, for the intergenerational human capital model, variation in state output per worker growth is better captured by variation in input per worker growth than found in Klenow and Rodriguez-Clare (1997).

Table 13: Variance Decomposition of Growth Rates: Share Explained by Inputs

	Growth Horizon				
	60 years	30 years	20 years	10 years	Average
Mincer human capital & physical capital	.2895	.4354	.2710	.1799	.2940
Intergenerational human capital	.1967	.5677	.4895	.2717	.3814
Intergenerational human capital & physical capital	.1967	.6463	.5372	.3736	.4384
N	16	32	48	96	

8 Conclusions and future extension

Utilizing data collated from the decennial censuses, Planning Commission reports and the Ministry of Human Resource Development, we construct estimates of years of education in the workforce, from 1951 through 2011, for the states of India in their present form, correcting for several intra-national changes of boundaries. A time series spanning these many years did not exist before. We also construct real output per worker across the states, and provide estimates of returns to education consistent with existing literature. Further, we fit an endogenous fertility choice model to the time paths of fertility and education in the states, accounting for the distribution of declining mortality rates across the states and increase in opportunity cost of space. Using national estimates of real physical capital per worker from Tamura, Dwyer, Devereux and Baier (2019), we assume competitive capital markets across the states of India. Thus our estimates of state physical capital per worker are proportional to our estimates of human capital per worker in each state.

We show that the intergenerational human capital model outperforms the standard Mincer human capital model in explaining log level differences in output per worker across the states. Our development accounting results show that between 75% and 80% of the log level differences are explained by log level differences in inputs and only 20% to 25% are explained by log differences in TFP. The Mincer human capital model performs less well, where log input differences only explain about 45% of log differences in output per worker.

In variance decomposition of growth rates of the states, our intergenerational human capital model explains between 38% and 44% of the variation in growth rates of output per worker. Again this is an improvement over the standard Mincer human capital model, which explains slightly less than 30% of output per worker growth variations. The remaining importance of TFP growth variations to explain variation in growth rates suggests that state heterogeneity in adoption of economic reforms are quite important, and not fully contained in the variation in input accumulation.

The seminal work of Galor and Mountford (2008) is also relevant. The varying rates of demographic transition among the Indian states, as well as the varying rates of economic growth suggest that international trade was also unequal in its importance. An important extension would be to examine the degree to which each state of India was open to international trade, and if the Galor and Mountford theory of differential impacts holds for the heterogeneous states of India. An interesting extension is an examination into whether distributional of educational expenditures and land use and infrastructural developments match up to our calibrated costs of education and rental rates, respectively. We delegate that task to a future body of work.

References

Aghion, P., Burgess, R., Redding, S.J., and Zilibotti, F. "The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India, *American Economic Review* 98, 2004: 1397-1412.

Agrawal, T. "Returns to Education in India: Some Recent Evidence," *Journal of Quantitative Economics* 20, 2012: 131-151.

Bhattacharya, B. B., Sakthivel, S. "Regional Growth and Disparity in India: Comparison of Pre- and Post-Reform Decades, *Economic and Political Weekly* 39, 2004: 1071-1077.

Blundell, R., Bond, S.R. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics* 87, 1998: 115-143.

Blundell, R., Bond, S.R. "GMM Estimation with Persistent Panel Data: An Application to Production Functions,*Econometric Review* 19, 2000: 321-340.

Caselli, F. "Accounting for Cross-Country Income Differences,"in Philippe Aghion and Steven Durlauf (eds.) *Handbook of Economic Growth* Amsterdam: North-Holland, 2005.

Chatterji, T. "The Micro-politics of Urban Transformation in the Context of Globalisation: A Case Study of Gurgaon, India," *South Asia: Journal of South Asian Studies* 36, 2013: 273-287.

Datta, P. "Urbanisation in India," *Regional and Sub-Regional Population Dynamic Population Process in Urban Areas European Population Conference* 2006, http://library.isical.ac.in:8080/jspui/bitstream/10263/2460/1/urbanisation

Dougherty, S., Chalaux, T., and Herd, R. What is Holding Back Productivity Growth in India?, *OECD Journal: Economic Studies* 2009: 1-22.

Duraisamy, P. "Changes in Returns to Education in India, 1983-1994: by Gender, Age-Cohort and Location," *Economics of Education Review* 21, 2002: 609-622.

Dutta, M. K., Sarma, G. K. "Foreign Direct Investment in India Since 1991: Trends, Challenges and Prospects," *SSRN Electronic Journal* 2008.

Easterly, W., Levine, R. "Its Not Factor Accumulation: Stylized Facts & Growth Models," *World Bank Economic Review* 15, 2001.

Erosa, A., Koreshchova, T.A., Restuccia, D. "How Important is Human Capital? A Quantitative Theory Assessment of World Income Inequality," *Review of Economic Studies* 77, 2010: 1421-1449.

Galor, O., Mountford, A. "Trading Population for Productivity: Theory and Evidence," *Review of Economic Studies* 75, 2008: 1143-1179.

Ghosh, S.C. *The History of Education in Modern India, 1757-1998* Orient Longman, New Dehli 2000.

Hall, R.E., Jones, C.I. “Why Do Some Countries Produce More Output Per Worker Than Others?” *Quarterly Journal of Economics* 114, 1999: 83-116.

Hsieh, C. T., Klenow, P. J. “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics* 124, 2009: 1403-1448.

Jones, C. I. “Was An Industrial Revolution Inevitable? Economic Growth Over the Very Long Run,” *Advances in Macroeconomics* 1, 2001: 1-43.

Kalemli-Ozcan, S. “Does Mortality Decline Promote Economic Growth?” *Journal of Economic Growth* 7, 2002: 411-439.

Kalemli-Ozcan, S. “A Stochastic Model of Mortality, Fertility and Human Capital Investment,” *Journal of Development Economics* 62, 2003: 103-118.

Kennedy, L., Robin, K., Zamuner, D. “Comparing State-Level Policy Responses to Economic Reforms in India,” *Revue de la Régulation* 13, 2013: https://journals.openedition.org/regulation/10247.

Klenow, P.J., Rodriguez-Clare, A. “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?,” *NBER Macroeconomics Annual 1997* Cambridge: MIT Press, 1997.

Kothari, B., Pandey, A., and Chudgar, A. R. “Reading Out the “Idiot Box”: Same-Language Subtitling on Television in India,” *Information Technologies and International Development* 2, 2004: 23-44.

Kumar, C. R. “International Human Rights Perspectives on the Fundamental Right of Education - Integration of Human Rights and Human Development in the Indian Constitutional,” *Tulane Journal of International and Comparative Law* 12, 2004: 237-285.

Mandal, A. “Rich State, Poor State: Essays on Economic Disparities in India,” 2017: `https://tigerprints.clemson.edu/all_dissertations/1925`.

Manuelli, R. E., Seshadri, A. Human Capital and the Wealth of Nations,” *American Economic Review* 104, 2014: 2736-2762.

Montenegro, C. E., Patrinos, H. A. “Comparable Estimates of Returns to Schooling Around the World,” *Policy Research Working Paper 7020* Education Global Practice Group, World Bank Group, 2014.

Patrinos, H. A., Psacharopoulos, G. “Returns to Education in Developing Countries,” in *Economics of Education* Elsevier, 2020.

Pederson, J.D. “India’s Industrial Dilemmas in West Bengal,” *Asian Surveys* 41, 2001: 646-668.

Regmi, N. “Essays in Economic Growth and International Trade,” 2019, `https://tigerprints.clemson.edu/all_dissertations/2359`.

Regmi, N., and Tamura, R. "International Trends in Fertility and Educational Achievement,"Clemson University working paper, 2021.

Rodrik, D., Subramanian, A. "From "Hindu Growth" to Productivity Surge: The Mystery of the Indian Growth Transition,"*IMF Staff Papers* 2004, 52: 193-228.

Srikanth, H. "Construction and Consolidation of the Telangana Identity,"*Economic and Political Weekly* 2013: 45-46.

Suisse Group, C. A. *Emerging Consumer Survey 2017* https://www.credit-suisse.com/media/assets/corporate/docs/about-us/research/publications/emerging-consumer-survey-2017.pdf.

Tamura, R. "Income Convergence in and Endogenous Growth Model,"*Journal of Political Economy* 99, 1991: 522-540.

Tamura, R. "From Decay to Growth: A Demographic Transition to Economic Growth,"*Journal of Economic Dynamics & Control* 20, 1996: 1237-1261.

Tamura, R. "Human Capital and Economic Development,"*Journal of Development Economics* 79, 2006: 26-72.

Tamura, R., Simon, C.J., and Murphy, K.M. "Black and White Fertility, Differential Baby Booms: The Value of Equal Education Opportunity,"*Journal of Demographic Economics* 82, 2016: 27-109.

Tamura, R., Simon, C.J. "Secular Fertility Declines, Baby Booms and Economic Growth: International Evidence,"*Macroeconomic Dynamics* 21, 2017: 1601-1672.

Tamura, R, Dwyer, G., Devereux, J., and Baier, S.L. "Economic Growth in the Long Run,"*Journal of Development Economics* 137, 2019: 1-35.

Tamura, R, Witham, A. "Fertility, Demographic Change, and Economic Growth: Urban and Rural Evidence,"Clemson University working paper, 2021.

Turner, C, Tamura, R., Mulholland, S., and Baier, S.L. "Education and Income of the States of the United States: 1840-2000,"*Journal of Economic Growth* 12, 2007: 101-158.

Turner, C., Tamura, R., and Mulholland, S. "How Important Are Human Capital, Physical Capital, Total Factor Productivity for Determining State Economic Growth in the United States, 1840-2000?"*Journal of Economic Growth* 18, 2013: 319-371.

Vasudeva Dutta, P. "Returns to Education, New Evidence from India: 1983-1999,"*Education Economics* 14, 2006: 431-451.

Wooldbridge, J. *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA., M.I.T. Press, 2002.

World Bank, *World Development Indicators* 2016.

9 Appendix: Data

Here, we provide some context to the state level dataset we created for India from 1951 through 2011 in the background of existing work. Specifically, we compare the national-level data provided by Tamura, Dwyer, Devereux and Baier (TDDB) (2016) to aggregates for each year obtained from our data. TDDB created original estimates for data on real output per worker and schooling for 168 countries, including India, spanning the period 1820-2010. To compare our estimates to those of TDDB, we create labor force-weighted national aggregates for each year. The TDDB dataset has data on India for 1951, 1961, 1971, 1980, 1990, 2000 and 2010; as compared to our dataset, which covers the decadal census years falling between 1951-2011. Thus, while the years of data differ a tad between the two datasets, they are sufficiently close to warrant a legitimate comparison.

Population by age group are available from the population volumes of each census.

Labor force was calculated as the sum of main workers, marginal workers and unemployed as given in the economics volumes of the individual censuses.

Our 1951 real gross state products come from "Estimates of State Domestic Product,"by States Directorate of Economics and Statistics. For 1961 - 2011, real gross state products come from "Economic and Political Weekly time series: Estimate of State Domestic Product."The 1951, 1961, 1971 and 1981 values were in constant 1981 prices. The 1991 and 2001 values were reported in constant 1994 prices. The 2011 values are in 2004 prices. See `http://mospi.nic.in/data`. We used the OECD PPP conversion table to produce 2000 PPP$, `https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm`.

For young adult mortality we used "Age Tables Censuses of India 1961 through 2011."We used projections for 1951.

For total fertility rates we used "Fertility Tables Censuses of India 1951 through 2011."

For 1951 through 1981 primary, secondary and tertiary school enrollment rates we used data from *Ministry of Human Resource Development*. For 1991-2011 we used information from *Selected Statistics of School Education*.

While constructing our aggregates, for sake of completeness, we included our estimates for the new states that were excluded from our calibration exercise as well. This makes for a more complete estimate of the national level data.

Our data trend that of TDDB to a substantial extent, even though there was some discrepancy. Our estimates for income per worker are lower than those of TDDB by between 12% to 30% most

Table A1: Comparison with TDDB estimates

Mandal-Regmi-Tamura					TDDB				
Year	Real GDP per worker	Labor force	Real GDP	Avg. years of schooling	Year	Real GDP per worker	Labor force	Real GDP	Avg. years of schooling
1951	1692	114	192965	0.88	1951	1952	140	272384	1.56
1961	1901	138	262464	1.34	1961	2146	189	404718	2.10
1971	2161	168	362213	2.37	1971	2872	200	575070	3.33
1981	2709	223	604193	2.89	1980	2585	302	780751	4.44
1991	3708	268	995289	3.82	1990	4258	320	1362343	5.60
2001	4244	449	1905849	5.52	2000	5911	396	2338738	6.79
2011	8489	511	4336805	7.36	2010	10725	455	4883831	7.85
	PPP 2000	million	million, PPP 2000			PPP 2000	million	million, PPP 2000	

Figure A1

years, except for that in 1981, where ours was slightly higher. Our estimates for GDP are also lower for most years, while the average years of schooling is lower by a little less than a year, although our estimate nearly matches the one of TDDB for 2010/2011.

Upon examination of the Census data files, it seems that the discrepancy may be due to the way the state census documents presented estimates of labor force. The 2001 and 2011 censuses had greater detail for labor force by educational attainment, and thus we could include the main workers, marginally employed and the unemployed within our estimate for the work force, in line with its economic description. The TDDB estimates seem to not consider the unemployed within the labor force, but do include the marginal workers. We were not able to include either the marginal workers or the unemployed workers while constructing the 1951-1991 estimates because the state level census volumes did not include these figures. Our examination of the proportion of discrepancy between our estimates and TDDB's seem to suggest that the latter include the marginally employed for all years. Further, the World Development indicators, upon which the TDDB estimates are based, consider only people of ages 15 and above as part of the labor force. The census estimates, however, consider younger workers as well.

Our aggregate data also exclude a number of territories for which data were not avalable for one or more of the years of our analysis period. None of the union territories, which generally are richer, are included. Many of the numerous smaller kingdoms decided to join the Indian union

Figure A2

Figure A3

only much later and we do not have data compiled for those either, since they were likely not subject to India's census laws. A2 provides an overview of the difference in coverage between our dataset and that from TDDB.

Table A2: Population covered with output per worker and output per worker and fertility vs TDDB

	Population, in millions, & (share) covered		
	output per worker	output per worker & fertility	TDDB
1951	335 (93.8)	326 (91.3)	357
1961	429 (98.4)	395 (90.6)	436
1971	536 (97.8)	492 (89.8)	548
1981	636 (93.7)	613 (90.3)	679
1991	826 (97.2)	758 (89.2)	850
2001	945 (93.2)	922 (90.9)	1014
2011	1075 (90.9)	1059 (89.6)	1182

10 Creation of state data used in calibration

Andhra Pradesh

The state of Andhra Pradesh was carved out from the states of Hyderabad and Madras in 1953. Data for the 1951 Census were created by averaging the numbers for the two parent states. Data were collected specifically for the state from the 1961 Census onwards and could be used as they were. In 2014, the state was divided into Telangana and Andhra Pradesh but since our dataset ended in 2011, it did not figure into our calculations.

Assam

While Assam was divided several times into smaller states (Meghalaya, Nagaland, Mizoram and Arunachal Pradesh) and Assam between the years 1960 and 1986, the state retained the majority of its population and data were consistently collected throughout the Census years. Thus, we used data from Assam as reported in the Census state volumes.

Bihar

Bihar has been a state within the Republic of India since Independence and so we used data for the state as provided in the state Census volumes.

Delhi

While Delhi was a union territory at independence, owing to its status as the capital, the Census collected detailed data and we were able to use the numbers as provided in the volumes.

Gujarat

Gujarat was carved out of the region of Saurashtra and the state of Bombay in 1960. For 1951, we used weighted average data from the two regions to construct our estimate for the state.

Haryana

Haryana was formed from the state of Punjab. We assigned data from the Punjab state to both the modern Punjab and Haryana for 1951 and 1961, since their per capita incomes were roughly the same in 1971, the first year in which official data were available.

Karnataka

Karnataka was formed mostly from the Mysore state in 1973, with minor parts of Madras and Bombay forming the northern and southern extremes respectively. We attributed data from only the Mysore state to Karnataka for 1951, 1961 and 1971.

Kerala

The majority of Kerala was created from the state of Travancore-Cochin in 1956, while a minor bit of Madras was also added to it. For 1951, we considered just data given for Travancore-Cochin in the Census.

Madhya Pradesh

The state was created by combination of Madhya Bharat, Vindhya Pradesh and Bhopal states in 1956. For 1951, thus, we used data from all three states as given in the Census for creating estimates for Madhya Pradesh.

Maharashtra

Maharashtra was formed taking the majority of the Bombay state in 1960, and, thus, we attribute data from Bombay state in 1951 to Maharashtra.

Odisha

Odisha comprised of 12 princely states that acceded to form the unified state under Indian rule right at independence. Data for the state, thus, is available directly from the very first Census.

Punjab

*See Haryana

Rajasthan

Rajasthan state was formed by the 1956 amalgamation of the Ajmer province into the area known during independence as greater Rajasthan. Thus, for 1951, we consider data collected for both these areas to form our estimates for the present day state.

Tamil Nadu

A majority of the erstwhile Madras state went to the creation of Tamil Nadu in 1969. Therefore, we used data for the Madras state to form our 1951 and 1961 estimates for Tamil Nadu.

Uttar Pradesh

Uttar Pradesh joined the Indian union as a state in 1950- thus, the Census has data on the state from 1951 and we used them directly to create our estimates.

West Bengal

West Bengal created as a state of India in 1947 via the partition of Bengal– half of which went to Pakistan as East Bengal (later renamed East Pakistan). Thus, data are directly available for the state from 1951 onwards.

11 Tables

Table T1: Modern States of India and their creation

State	**Formation**
Assam	*January 26, 1950*
Bihar	*January 26, 1950*
Jammu and Kashmir	*January 26, 1950*
Odisha	*January 26, 1950*
Tamil Nadu	*January 26, 1950*
Uttar Pradesh	*January 26, 1950*
West Bengal	*January 26, 1950*
Andhra Pradesh	*October 1, 1953*
Karnataka	*November 1, 1956*
Kerala	*November 1, 1956*
Madhya Pradesh	*November 1, 1956*
Rajasthan	*November 1, 1956*
Gujarat	*May 1, 1960*
Maharashtra	*May 1, 1960*
Nagaland	*December 1, 1963*
Haryana	*November 1, 1966*
Punjab	*November 1, 1966*
Himachal Pradesh	*January 25, 1971*
Manipur	*January 21, 1972*
Meghalaya	*January 21, 1972*
Tripura	*January 21, 1972*
Sikkim	*May 16, 1975*
Arunachal Pradesh	*February 20, 1987*
Mizoram	*February 20, 1987*
Goa	*May 30, 1987*
Chhattisgarh	*November 1, 2000*
Uttarakhand	*November 9, 2000*
Jharkhand	*November 15, 2000*
Telangana	*June 2, 2014*

Table T2: Proportion of population deemed illiterate in Census Years
(1-Proportion literate of any education level)

States/Union Territory	1951	1961	1971	1981	1991	2001	2011
Andaman & Nicobar Islands	0.70	0.60	0.49	0.37	0.27	0.19	0.14
Andhra Pradesh		0.79	0.75	0.64	0.56	0.40	0.32
Arunachal Pradesh		0.93	0.89	0.74	0.58	0.46	0.33
Assam	0.81	0.67	0.66		0.47	0.37	0.27
Bihar	0.87	0.78	0.77	0.68	0.63	0.53	0.36
Chandigarh			0.30	0.25	0.22	0.18	0.14
Chhattisgarh	0.91	0.82	0.76	0.67	0.57	0.35	0.29
Dadra & Naqar Haveli			0.82	0.67	0.59	0.42	0.22
Daman & Diu					0.29	0.22	0.13
Delhi		0.38	0.35	0.28	0.25	0.18	0.13
Goa	0.77	0.65	0.48	0.34	0.24	0.18	0.13
Gujarat	0.78	0.69	0.63	0.55	0.39	0.31	0.21
Haryana			0.74	0.63	0.44	0.32	0.23
Himachal Pradesh					0.36	0.24	0.16
Jammu & Kashmir		0.87	0.78	0.69		0.44	0.31
Jharkhand	0.87	0.79	0.76	0.65	0.59	0.46	0.32
Karnataka		0.70	0.63	0.54	0.44	0.33	0.24
Kerala	0.53	0.45	0.30	0.21	0.10	0.09	0.06
Lakshadweep	0.85	0.73	0.48	0.32	0.18	0.13	0.08
Madhya Pradesh	0.87	0.79	0.73	0.61	0.55	0.36	0.29
Maharashtra	0.72	0.65	0.54	0.43	0.35	0.23	0.17
Manipur	0.87	0.64	0.62	0.50	0.40	0.29	0.20
Meghalaya		0.73	0.71	0.58	0.51	0.37	0.25
Mizoram	0.69	0.56	0.46	0.40	0.18	0.11	0.08
Nagaland	0.89	0.78	0.66	0.50	0.38	0.33	0.20
Orissa	0.84	0.78	0.74	0.66	0.51	0.37	0.27
Puducherry		0.56	0.47	0.35	0.25	0.19	0.13
Punjab			0.66	0.57	0.41	0.30	0.23
Rajasthan	0.92	0.82	0.77	0.70	0.61	0.40	0.33
Sikkim			0.82	0.66	0.43	0.31	0.18
Tamil Nadu		0.64	0.55	0.46	0.37	0.27	0.20
Tripura		0.80	0.69	0.50	0.40	0.27	0.12
Uttar Pradesh	0.88	0.79	0.76	0.67	0.59	0.44	0.30
Uttarakhand	0.81	0.82	0.67	0.54	0.42	0.28	0.20
West Bengal	0.75	0.66	0.61	0.51	0.42	0.31	0.23
INDIA	0.82	0.72	0.66	0.56	0.48	0.35	0.26

Table T3: Two-way fixed-effects regression for feedback from income to future education

	with decade dummies	
	Estimate	Standard Error
$log(y_t)$	0.2445	0.2403
E_t	0.3436***	0.0791
$log(y_{t+1})$	0.2274	0.2289

Table T4: Output per worker 2000 PPP dollars

	1951	1961	1971	1981	1991	2001	2011	%/year (1951-2011)	%year (2001-2011)
Andhra Pradesh	1125	1311	1574	1698	2592	3747	6760	2.99	5.90
Assam	2354	2224	1996	2559	3139	3087	4998	1.25	4.82
Bihar	1305	1458	1355	2197	2924	1827	2967	1.37	4.85
Chhattisgarh	NA	NA	NA	NA	NA	NA	NA	NA	NA
Delhi	4227	4805	4336	7196	9218	11415	25715	3.01	8.12
Gujarat	1972	2316	2694	3370	3987	4843	10206	2.74	7.45
Haryana	1655	2308	3714	4432	6358	6670	12645	3.39	6.40
Himachal Pradesh	1210	1627	2249	2652	3337	4792	9247	3.39	6.57
Jammu and Kashmir	NA	NA	1612	3164	2915	3694	6664	3.55*	5.90
Jharkhand	NA	NA	NA	NA	NA	2636	5422	-	7.21
Karnataka	1545	1642	1943	2018	3141	4000	9415	3.01	8.56
Kerala	1544	1882	2083	2757	3838	4815	10748	3.23	8.03
Madhya Pradesh	674	1037	1454	2241	3278	2970	4306	3.09	3.71
Maharashtra	1845	2042	2272	2884	4264	5554	11463	3.04	7.25
Odisha	1141	1376	1547	1989	2465	2653	3191	1.71	1.85
Punjab	1452	1896	3980	4034	6704	7243	10692	3.33	3.89
Rajasthan	1928	1954	1971	2086	3025	3320	6576	2.04	6.83
Tamil Nadu	1462	1631	1890	1937	2895	4156	9344	3.09	8.10
Telangana	NA	NA	NA	NA	NA	NA	NA	NA	NA
Uttarakhand	NA	NA	NA	NA	NA	3050	13294	-	14.7
Uttar Pradesh	1535	1700	1653	2219	3034	2971	4750	1.88	4.69
West Bengal	2728	2745	2903	3194	3609	3835	5311	1.11	3.26

NA: Not available, * (1971-2011)

Table T5: Average years of education in labor force

	1951	1961	1971	1981	1991	2001	2011
Andhra Pradesh	0.44	0.91	1.60	1.77	2.33	4.30	6.01
Assam	1.21	0.92	2.43	3.14	3.50	5.21	6.66
Bihar	0.96	0.88	1.71	2.34	3.00	4.15	5.78
Chhattisgarh	0.60	0.81	1.63	1.91	2.77	4.65	6.60
Delhi	2.21	4.68	6.34	6.84	8.10	9.41	10.83
Gujarat	0.83	1.83	2.46	3.21	4.27	5.95	7.70
Haryana	0.56	1.41	2.37	3.45	4.70	6.28	8.57
Himachal Pradesh	0.40	0.80	2.40	3.17	4.49	6.71	9.15
Jammu and Kashmir	0.50	0.66	1.81	2.69	3.86*	5.54	8.35
Jharkhand	0.88	0.94	1.79	2.40	3.09	4.87	6.81
Karnataka	1.33	1.02	2.92	2.94	3.54	5.45	7.38
Kerala	2.29	2.37	3.92	5.68	6.53	7.96	10.00
Madhya Pradesh	0.66	0.74	1.58	1.88	2.61	4.70	6.27
Maharashtra	1.28	1.61	2.86	3.48	4.44	6.42	8.26
Odisha	1.11	0.91	2.21	2.50	3.11	5.19	7.31
Punjab	0.80	1.63	2.75	3.06	4.86	6.46	8.19
Rajasthan	0.46	0.56	1.47	2.15	2.95	4.47	6.06
Tamil Nadu	1.02	1.40	2.67	3.18	3.96	5.30	7.76
Telangana	0.45	0.99	1.33*	1.78	2.47	4.52	6.07
Uttarakhand	0.54	1.09	1.88	2.62	3.67	6.42	8.64
Uttar Pradesh	0.51	1.02	1.80	2.54	3.54	5.06	6.87
West Bengal	1.37	2.10	2.88	3.44	4.22	5.54	7.28

*Interpolated

Table T6: Average years of experience

	1951	1961	1971	1981	1991	2001	2011
Andhra Pradesh	18.06	17.29	17.32	18.96	18.46	19.78	14.44
Assam	15.83	15.51	14.77	15.36	15.57	16.18	11.68
Bihar	17.20	15.83	17.02	16.88	15.24	15.16	10.34
Chhattisgarh	NA	NA	NA	NA	NA	NA	NA
Delhi	15.81	14.58	13.67	14.57	13.85	14.52	9.20
Gujarat	17.11	15.37	16.80	17.92	17.05	18.22	12.11
Haryana	16.40	14.74	14.97	15.10	14.45	15.45	10.31
Himachal Pradesh	19.04	22.26	17.05	18.56	18.35	18.94	12.00
Jammu and Kashmir	16.00	16.13	14.99	16.49	16.34	17.17	9.83
Jharkhand	NA	NA	NA	NA	NA	15.76	10.78
Karnataka	15.97	16.41	15.92	17.44	18.13	19.26	13.02
Kerala	16.76	17.01	16.83	15.68	17.29	8.96	12.87
Madhya Pradesh	17.67	16.12	17.12	17.75	16.83	16.97	11.75
Maharashtra	16.52	16.41	16.64	19.20	17.97	18.33	11.97
Odisha	17.89	17.19	16.32	17.71	17.98	18.52	12.55
Punjab	15.86	14.52	16.00	18.53	16.34	17.16	12.33
Rajasthan	17.27	15.47	15.81	15.89	15.53	16.59	11.45
Tamil Nadu	18.95	17.99	18.23	20.33	20.53	20.96	14.27
Telangana	NA	NA	NA	NA	NA	NA	NA
Uttarakhand	NA	NA	NA	NA	NA	17.42	10.16
Uttar Pradesh	18.12	16.22	16.52	15.81	13.91	15.56	10.10
West Bengal	19.26	15.10	15.68	16.04	16.71	18.34	13.16

NA: Not available

Table T7: Total Fertility Rates

State	1951	1961	1971	1981	1991	2001	2011
Andhra Pradesh	5.39	5.57	4.68	4.07	3.10	2.35	1.85
Assam	7.02	6.06	5.80	4.24	3.53	3.06	2.46
Bihar	6.27	6.32	6.12	5.79	4.49	4.38	3.81
Chhattisgarh	NA	NA	NA	NA	NA	3.30	2.70
Delhi	4.17	4.47	3.49	2.93	2.64	2.15	1.80
Gujarat	6.44	6.55	5.71	4.44	3.17	2.93	2.53
Haryana	7.23	7.20	6.82	5.16	4.08	3.19	2.36
Himachal Pradesh	NA	NA	NA	NA	1.54	1.31*	1.12
Jammu and Kashmir	NA	NA	NA	NA	NA	2.40	1.90
Jharkhand	NA	NA	NA	NA	NA	3.56	2.96
Karnataka	5.73	6.01	4.54	3.65	3.16	2.47	1.92
Kerala	4.94	5.05	4.22	2.93	1.86	1.78	1.84
Madhya Pradesh	6.47	6.68	5.68	5.24	4.67	4.00	3.20
Maharashtra	5.60	5.77	4.72	3.68	3.05	2.46	1.86
Odisha	5.91	6.22	4.80	4.35	3.40	2.65	2.24
Punjab	6.04	6.03	5.31	4.11	3.17	2.50	1.87
Rajasthan	6.62	6.60	6.37	5.30	4.65	4.07	3.12
Tamil Nadu	4.68	4.88	3.97	3.48	2.29	1.99	1.79
Uttarakhand	NA	NA	NA	NA	NA	4.50	2.09
Uttar Pradesh	6.40	6.25	6.63	5.87	5.26	2.77	2.95
West Bengal	6.51	6.83	5.21	4.28	3.30	2.47	1.77
Average	6.01	6.03	5.25	4.34	3.38	2.87	2.29
Max	7.23	7.20	6.82	5.87	5.26	4.74	3.81
Min	4.17	4.47	3.49	2.93	1.56	1.35	1.14

NA: Not available, *Interpolated

Table T8: Parameter values used in calibration

Parameter	Value	Parameter	Value
α	0.275	μ	0.085
ψ	0.66	$\overline{\tau}$	0.3825
φ	0.55	a	0.4007383
θ	0.125	Λ	2.014584
A	1.55	p	1
r	1.52		

Calibration

	model	United States	India	
fertility	2	2	2.4	
schooling	15.25	15.5	7.2	
annualized growth rate	1.8	1.8	2.4	
housing share	0.19	0.19	0.20	
next generation share	0.44	0.48	0.42	
		0.21	0.32	India investment rate
		0.08	0.04	education spending
		0.05	0.05	health expenditure
		0.02	0.01	R&D
		0.12	0.00	foregone earnings: schooling beyond 12 years

These two tables provide additional justification for the parameter values, as they replicate share of expenditures spent on housing for rich countries, and share of output spent on the next generation. The 32% investment rate is the average for the 2000-2014 period as given in the *Penn World Tables* (29%) and *World Development Indicators* (34%). India expenditure figures come from *World Bank*, 2014, the latest available. All of the US values, except for schooling, come from Tamura, Simon and Murphy (2016), The US schooling value comes from Tamura and Witham (2021).

12 Calibration Results

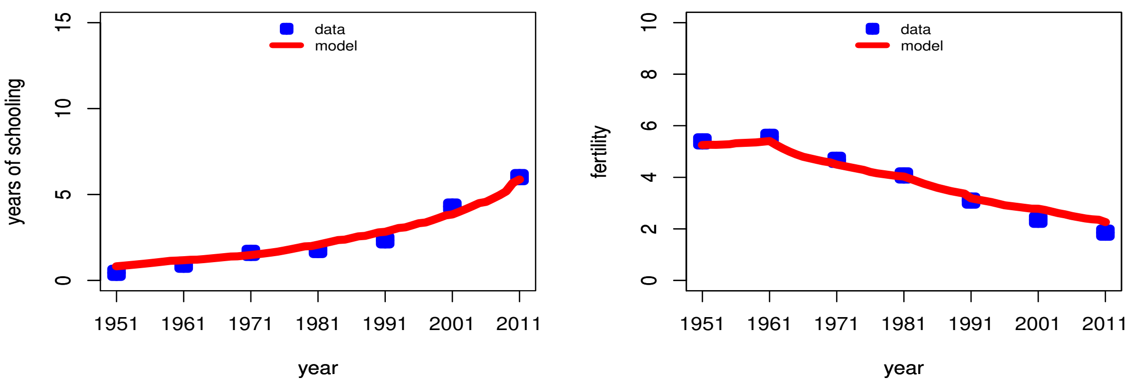


Figure C1: Average years of schooling and TFR for Andhra Pradesh

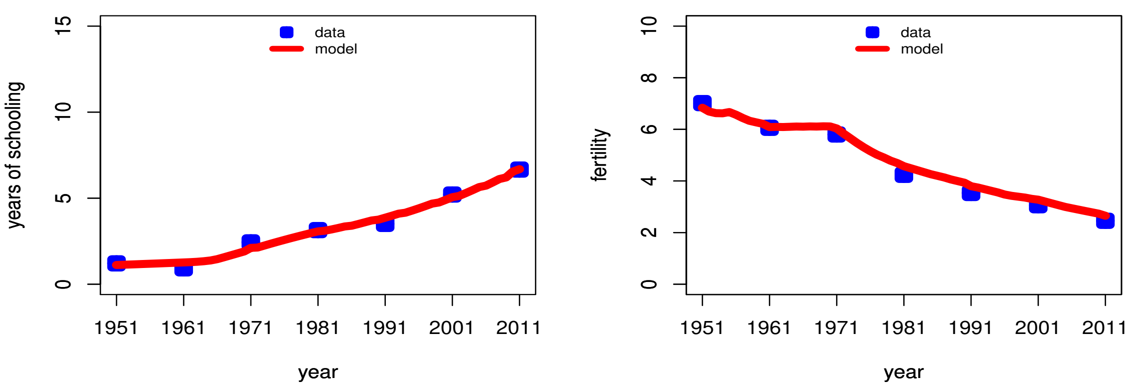


Figure C2: Average years of schooling and TFR for Assam

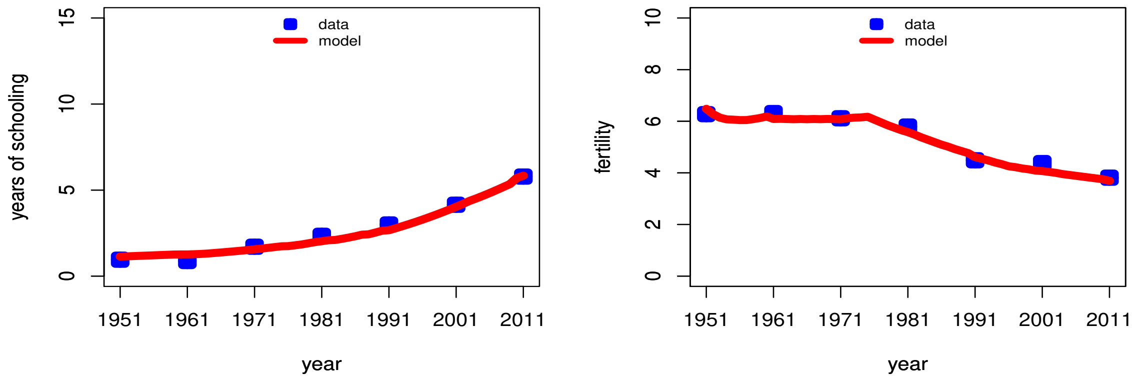


Figure C3: Average years of schooling and TFR for Bihar

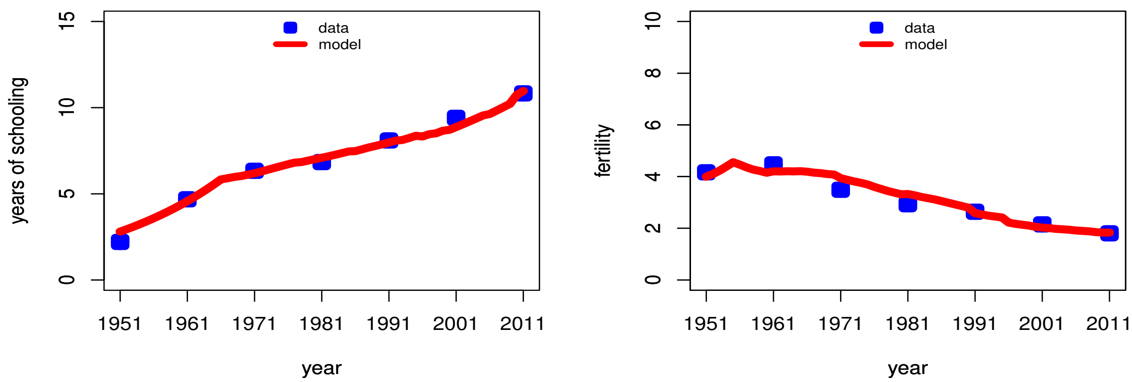


Figure C4: Average years of schooling and TFR for Delhi

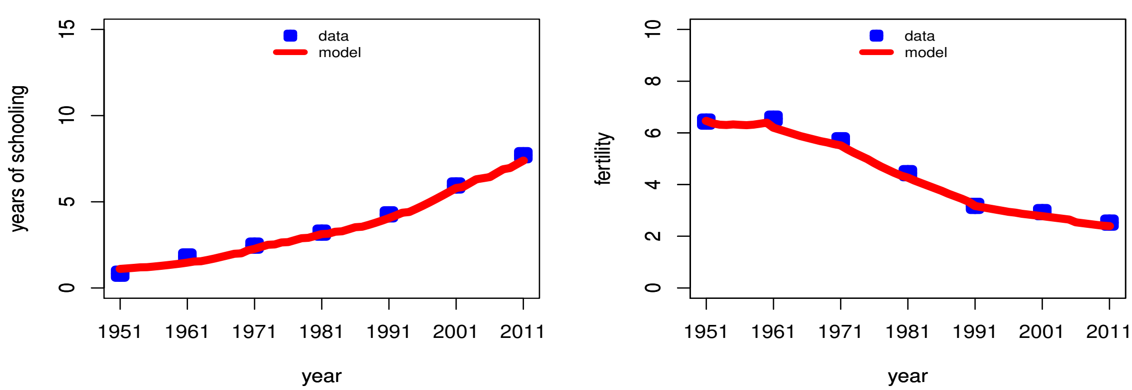


Figure C5: Average years of schooling and TFR for Gujarat

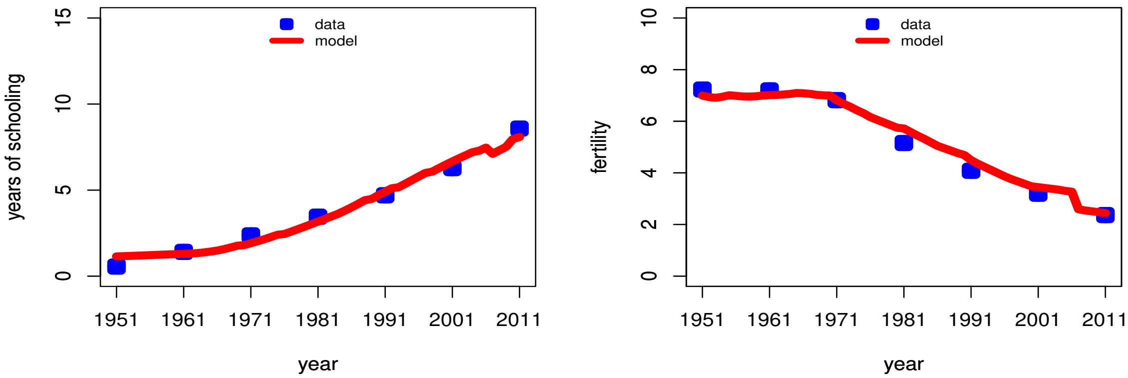


Figure C6: Average years of schooling and TFR for Haryana

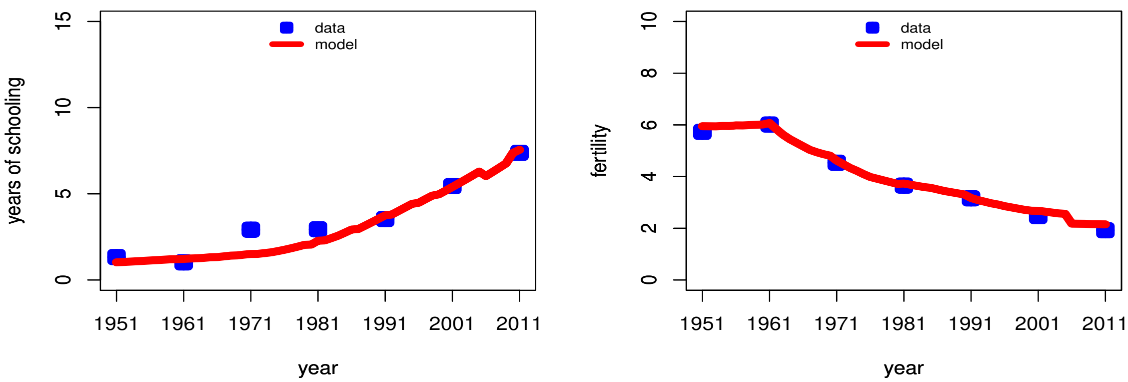


Figure C7: Average years of schooling and TFR for Karnataka

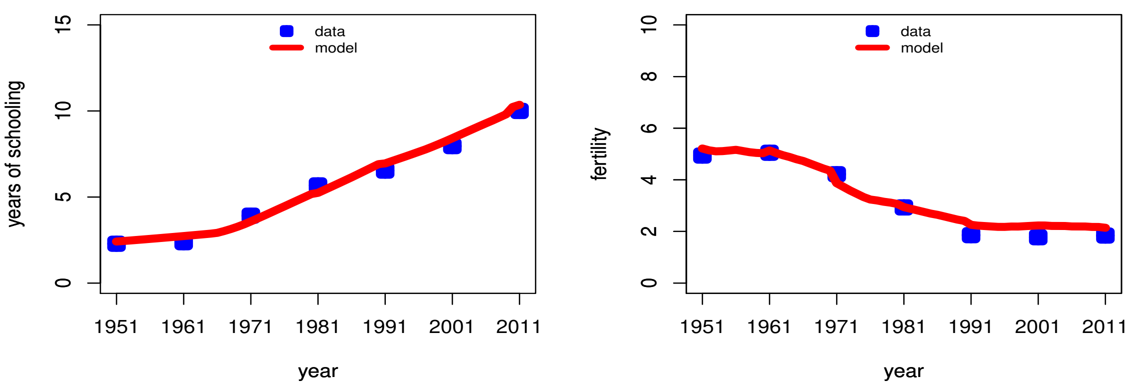


Figure C8: Average years of schooling and TFR for Kerala

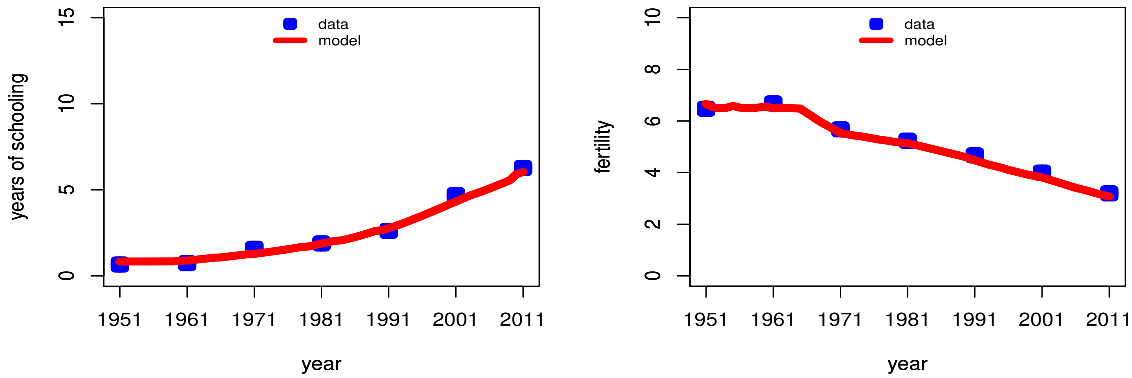


Figure C9: Average years of schooling and TFR for Madhya Pradesh

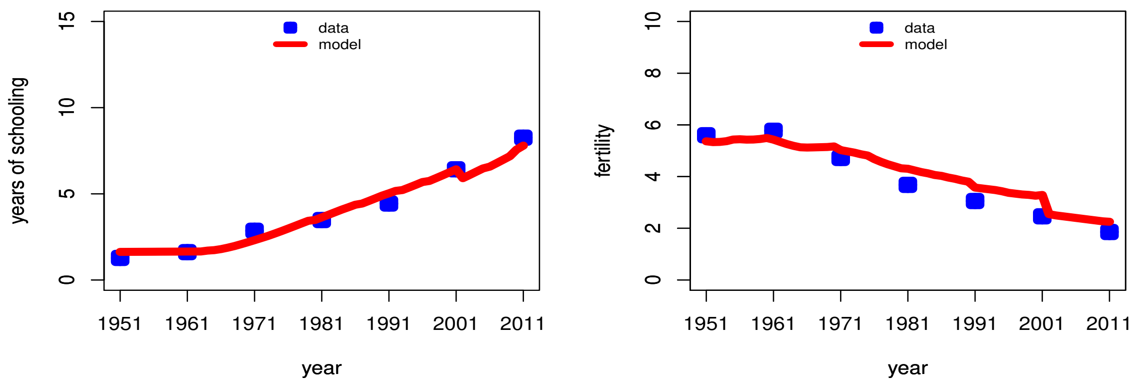


Figure C10: Average years of schooling and TFR for Maharashtra

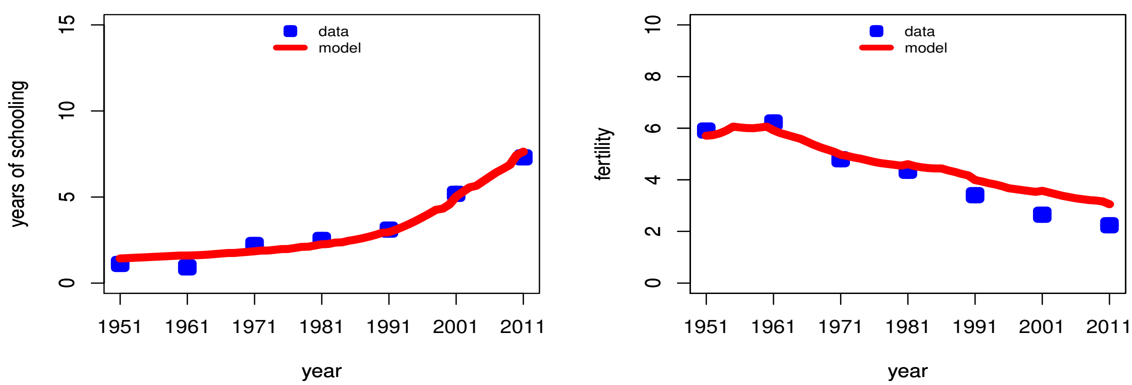


Figure C11: Average years of schooling and TFR for Odisha

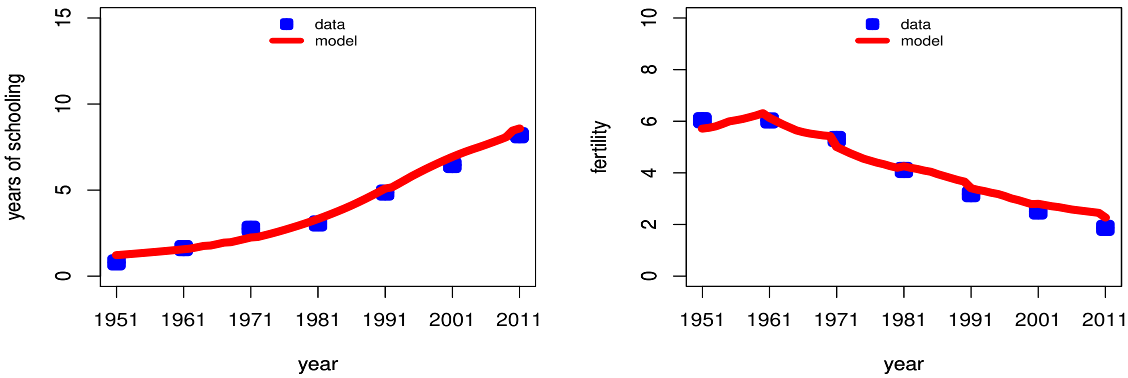


Figure C12: Average years of schooling and TFR for Punjab

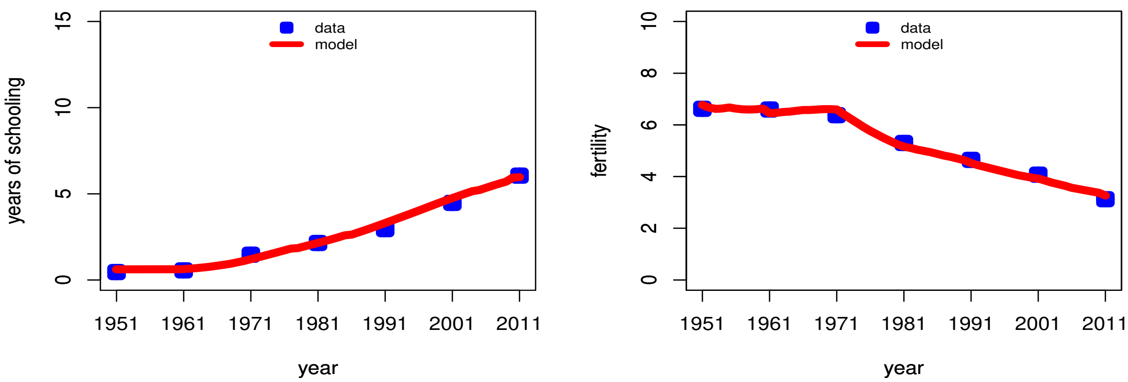


Figure C13: Average years of schooling and TFR for Rajasthan

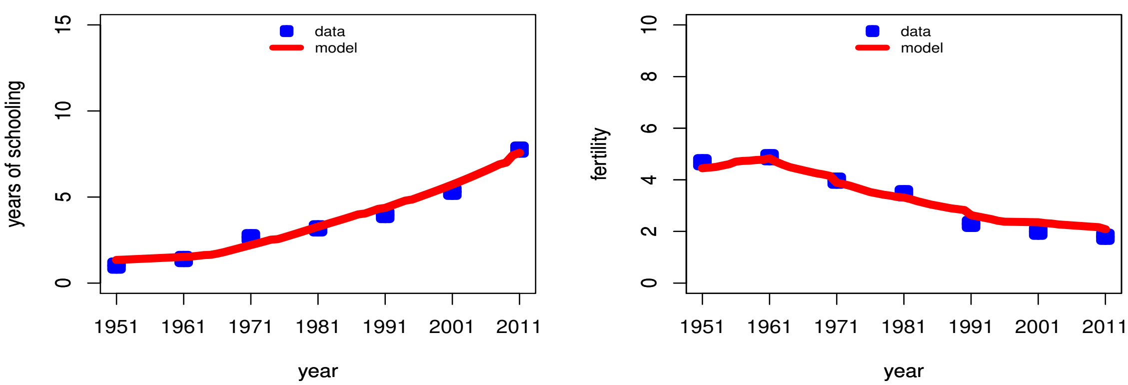


Figure C14: Average years of schooling and TFR for Tamil Nadu

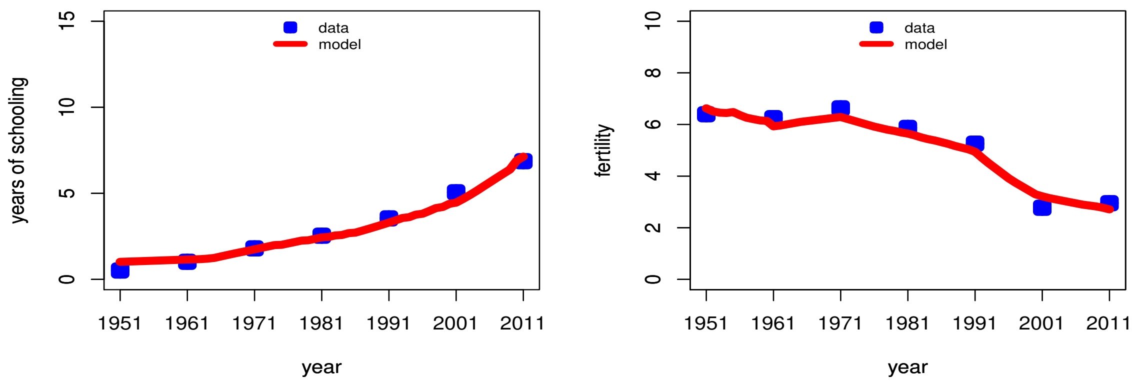


Figure C15: Average years of schooling and TFR for Uttar Pradesh

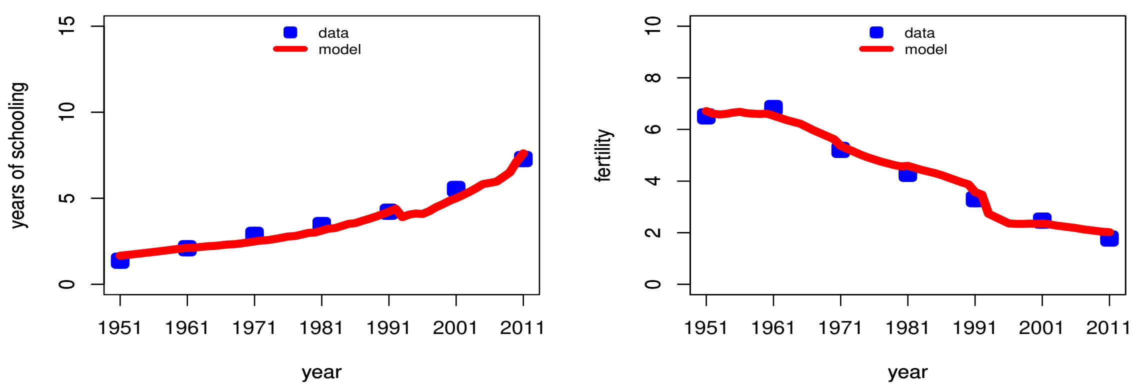


Figure C16: Average years of schooling and TFR for West Bengal

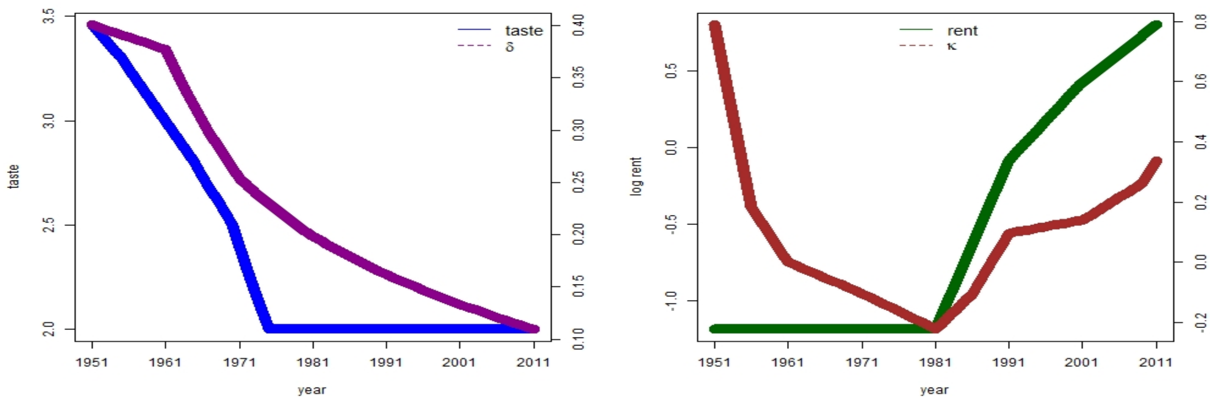


Figure C17: a) Taste and δ, b) κ and rent for Andhra Pradesh

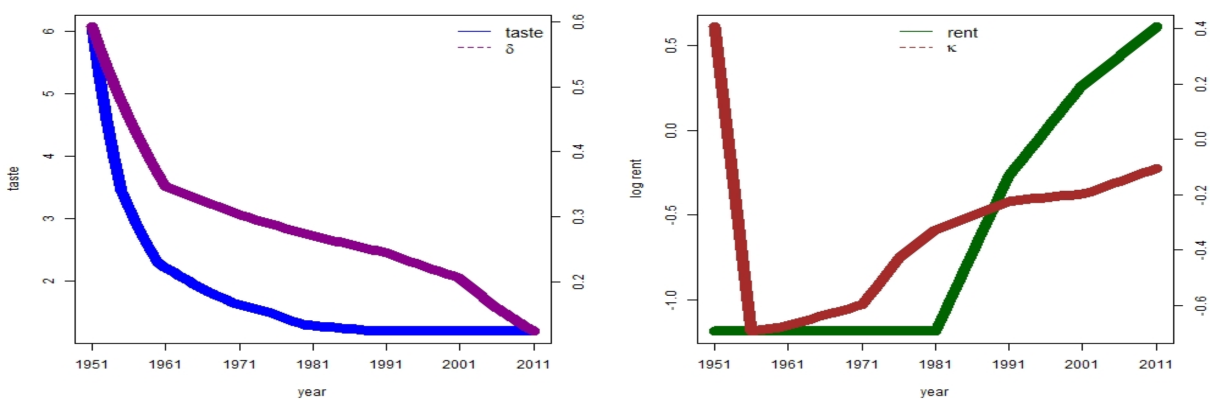


Figure C18: a) Taste and δ, b) κ and rent for Assam

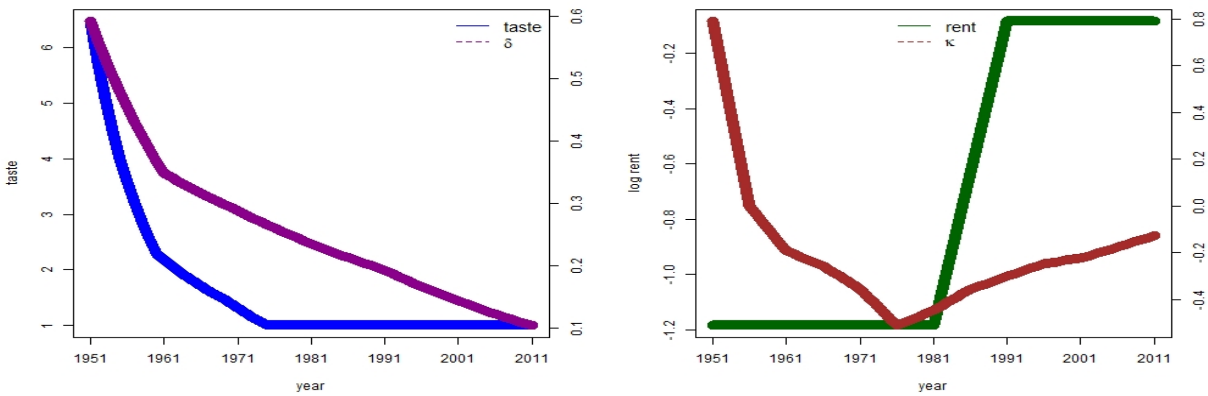


Figure C19: a) Taste and δ, b) κ and rent for Bihar

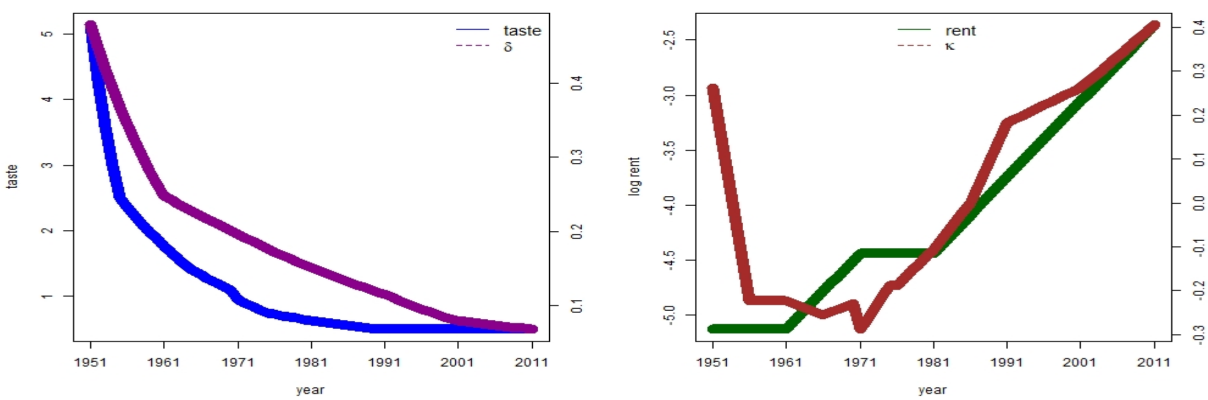


Figure C20: a) Taste and δ, b) κ and rent for Delhi

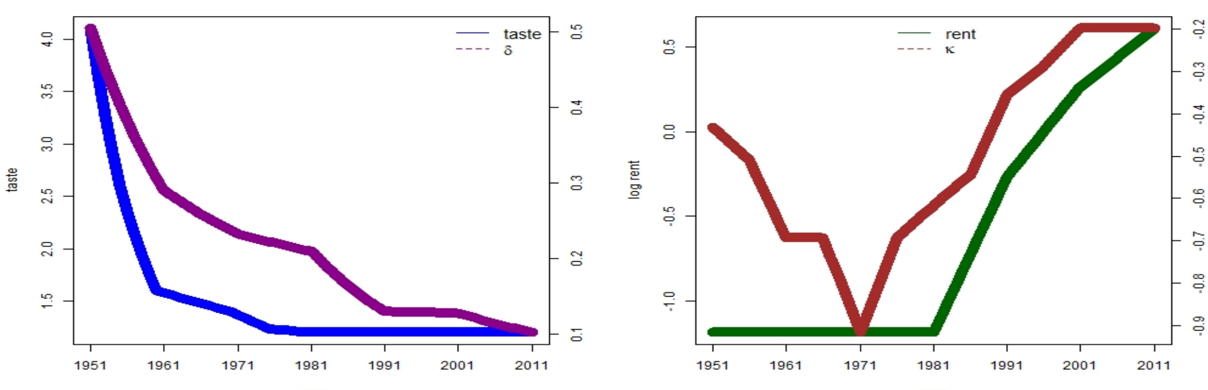


Figure C21: a) Taste and δ, b) κ and rent for Gujarat

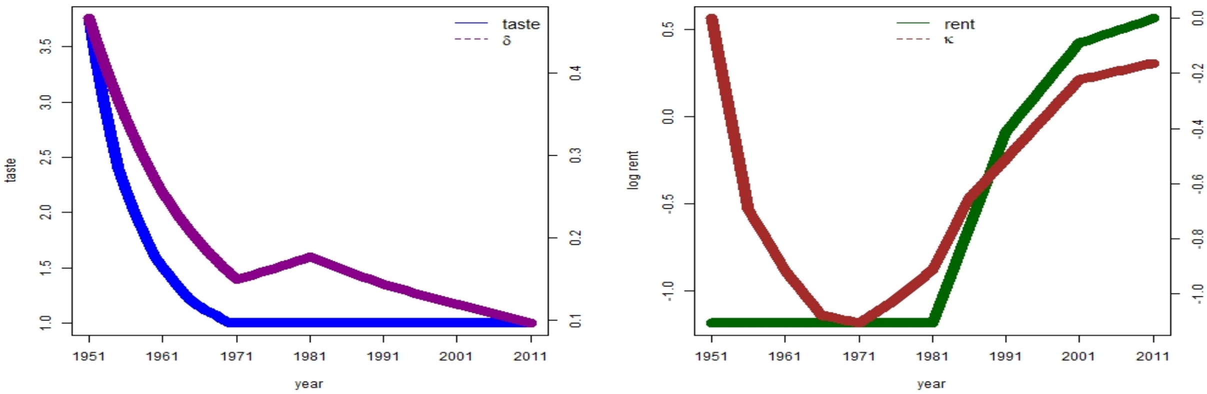


Figure C22: a) Taste and δ, b) κ and rent for Haryana

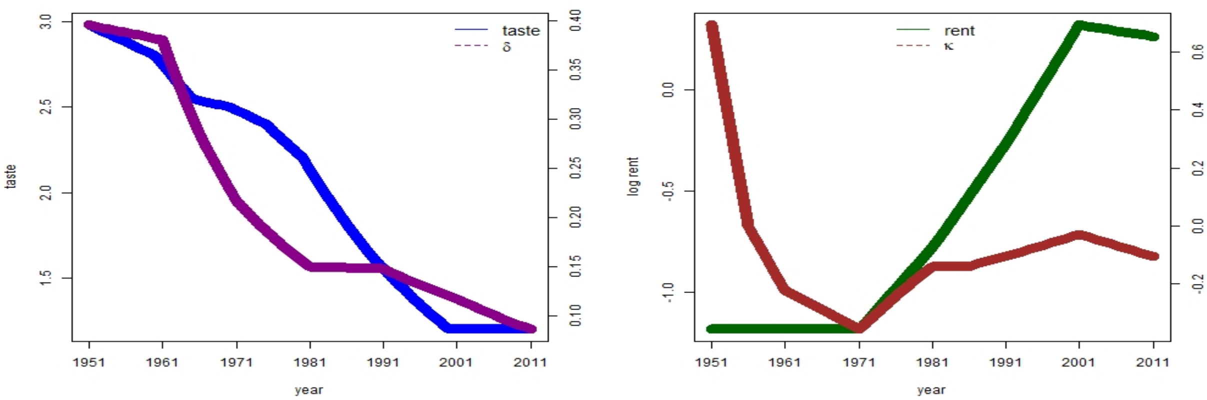


Figure C23: a) Taste and δ, b) κ and rent for Karnataka

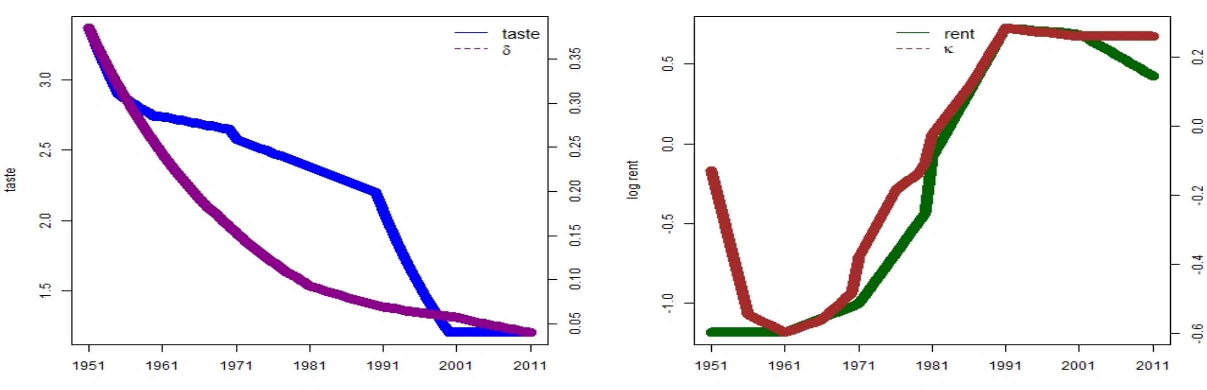


Figure C24: a) Taste and δ, b) κ and rent for Kerala

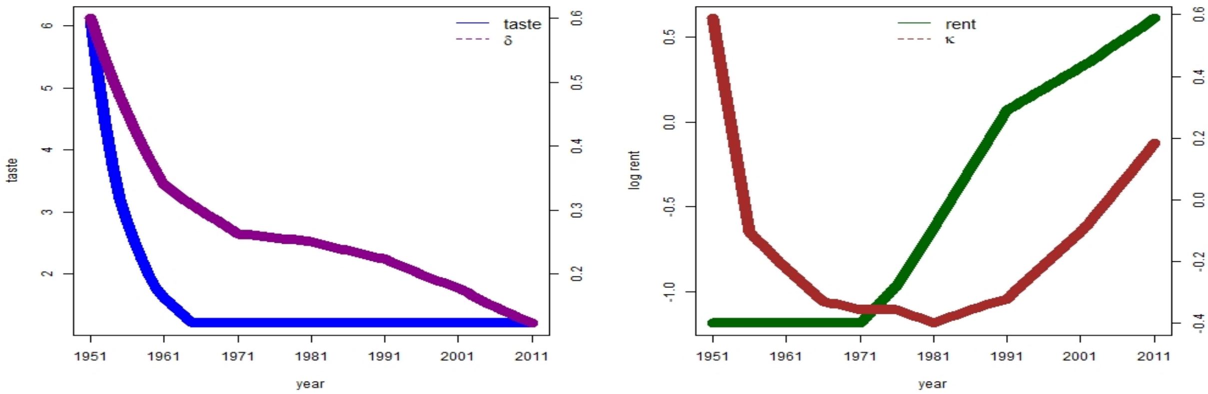


Figure C25: a) Taste and δ, b) κ and rent for Madhya Pradesh

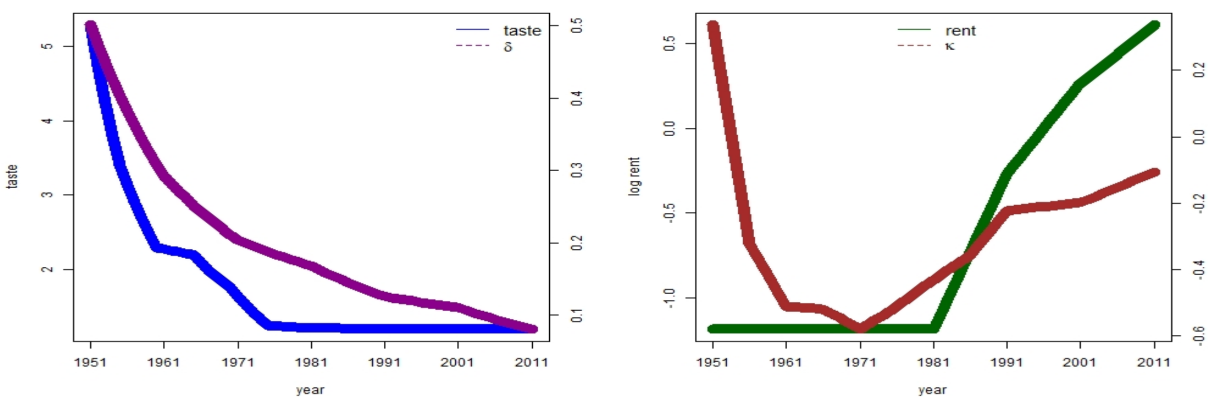


Figure C26: a) Taste and δ, b) κ and rent for Maharashtra

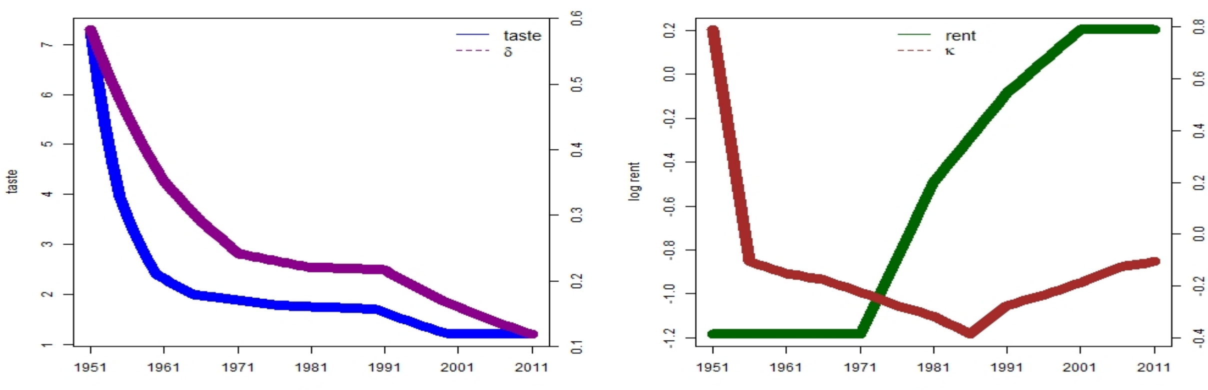


Figure C27: a) Taste and δ, b) κ and rent for Odisha

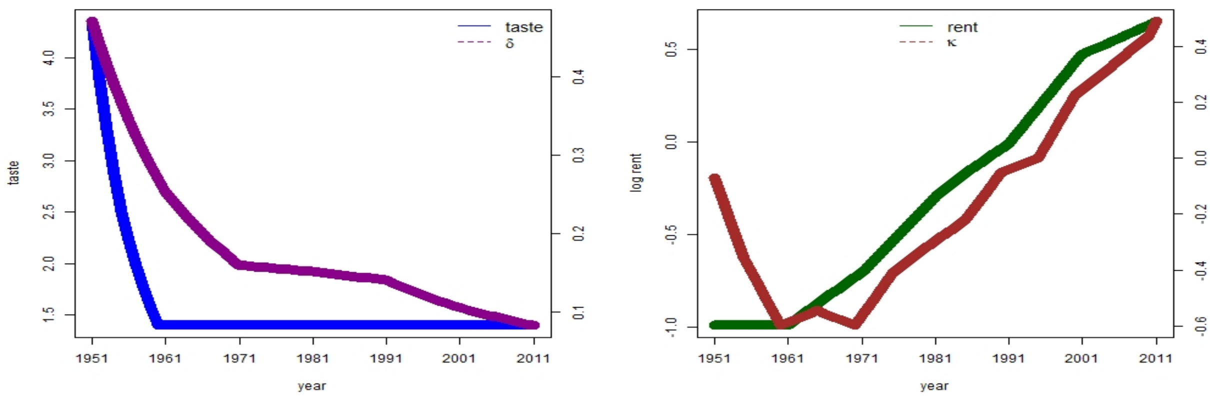


Figure C28: a) Taste and δ, b) κ and rent for Punjab

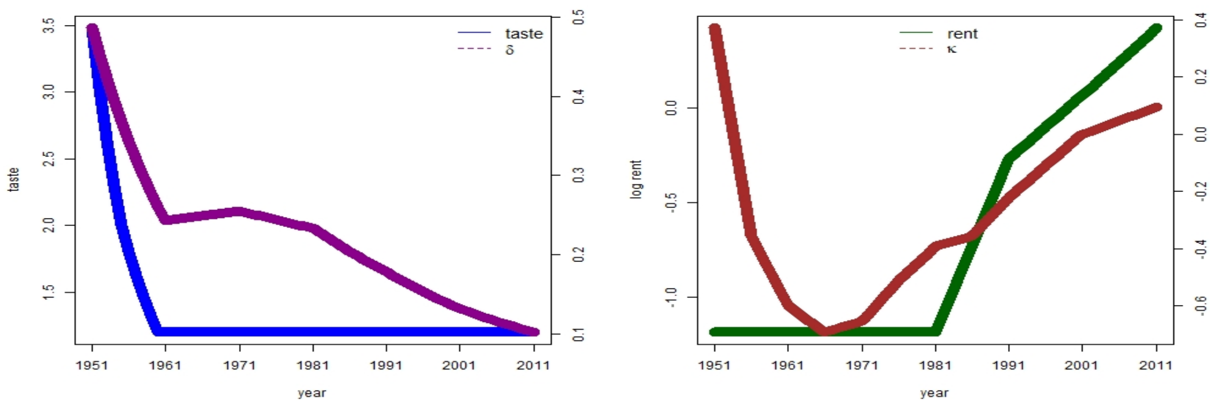


Figure C29: a) Taste and δ, b) κ and rent for Rajasthan

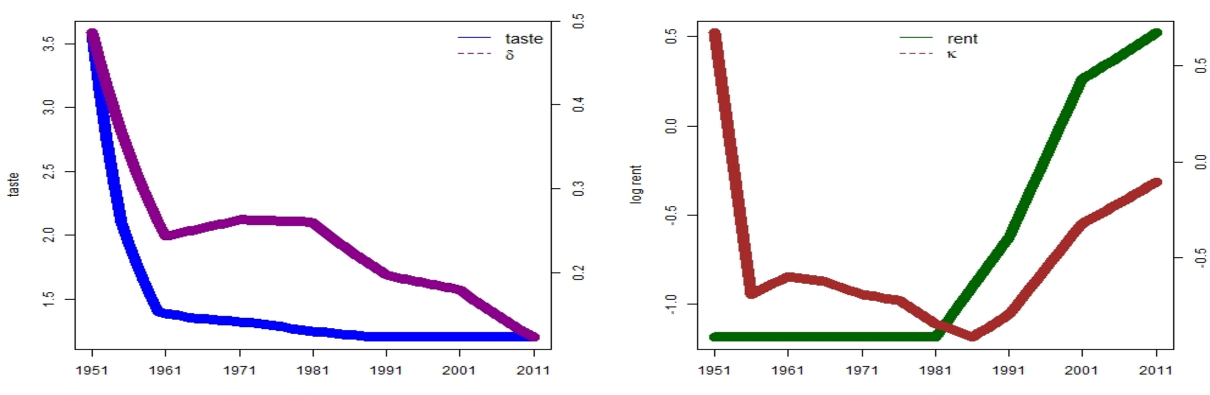


Figure C30: a) Taste and δ, b) κ and rent for Uttar Pradesh

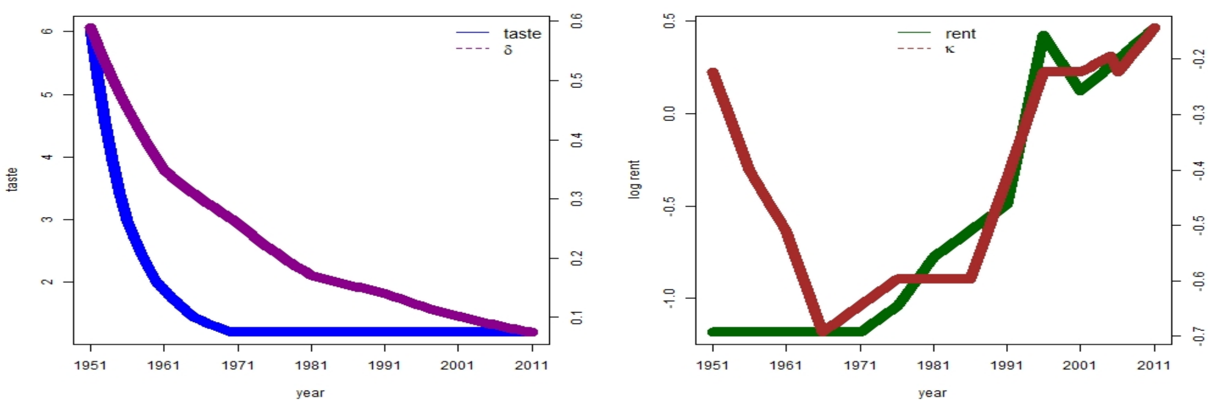


Figure C31: a) Taste and δ, b) κ and rent for West Bengal

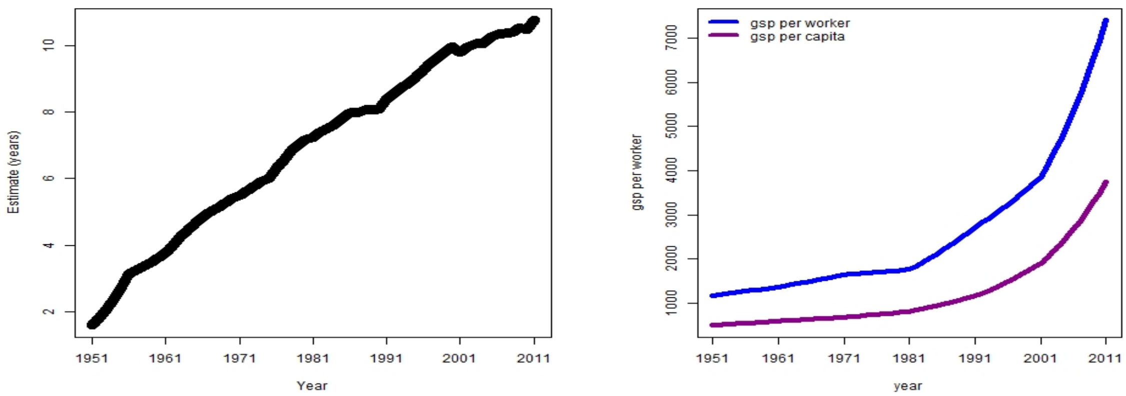


Figure C32: Young schooling and GSP per worker for Andhra Pradesh

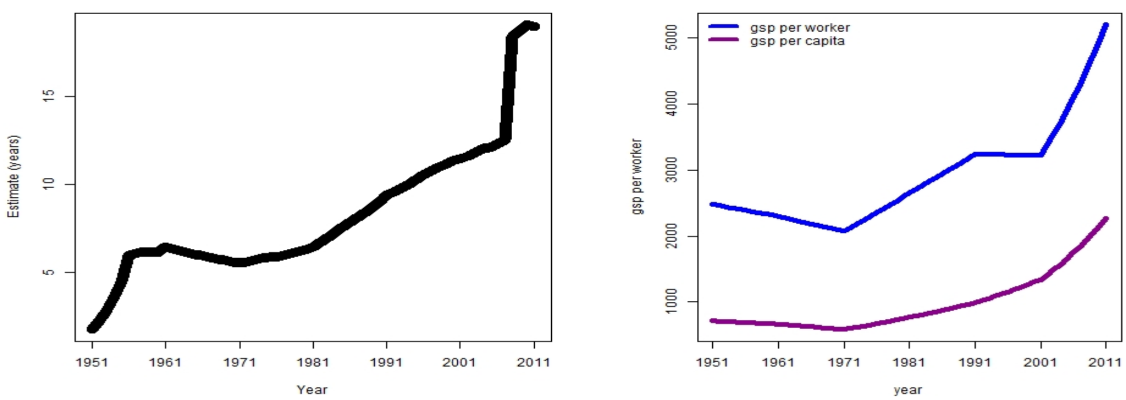


Figure C33: Young schooling and GSP per worker for Assam

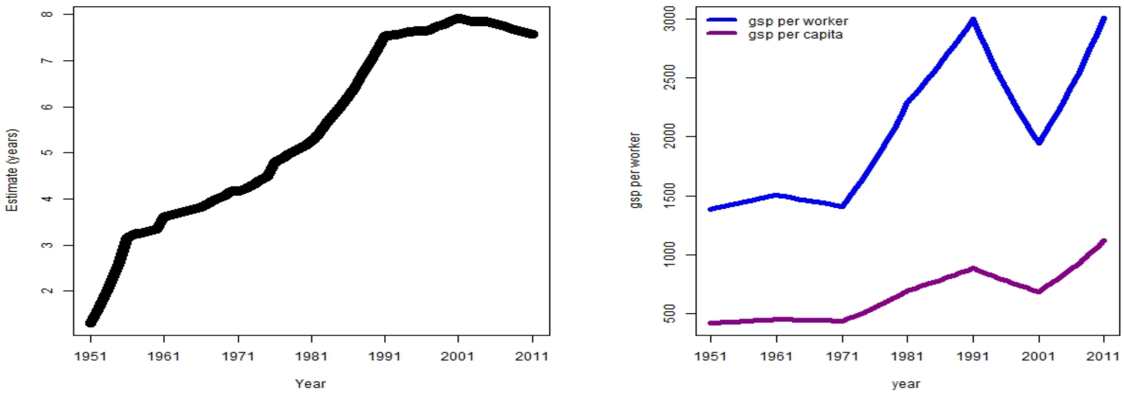


Figure C34: Young schooling and GSP per worker for Bihar

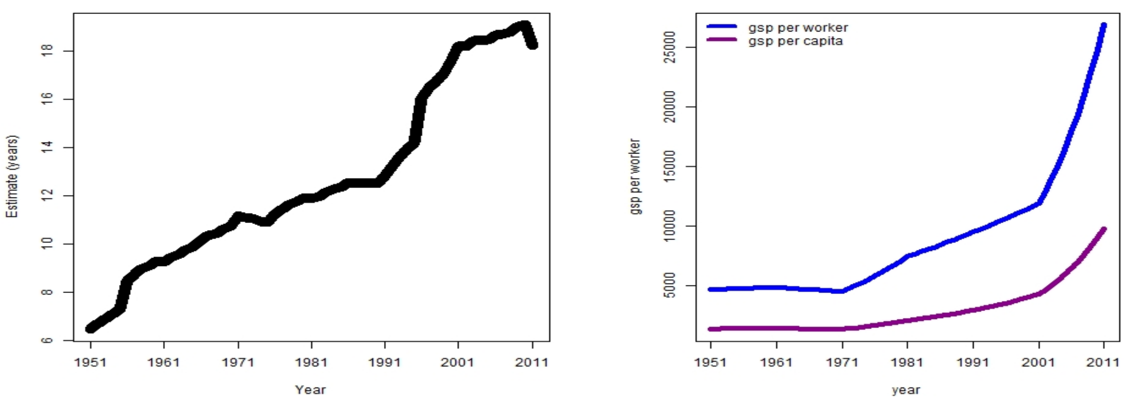


Figure C35: Young schooling and GSP per worker for Delhi

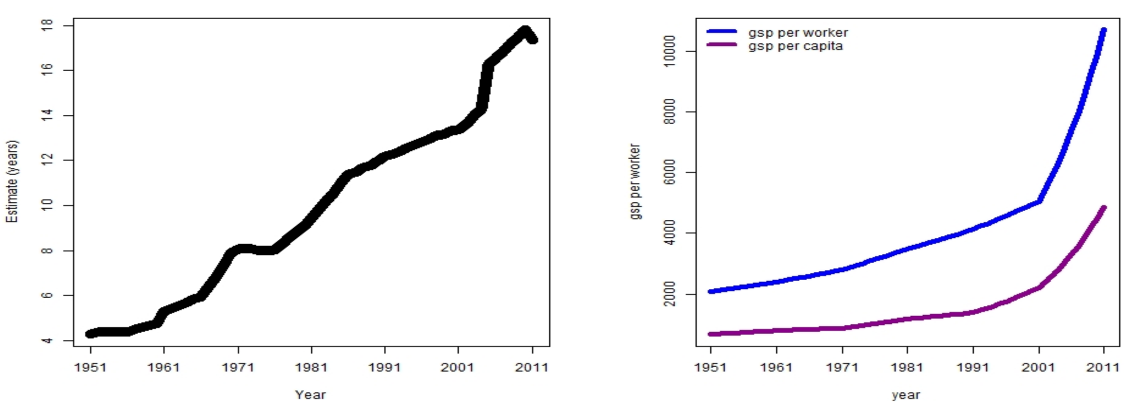


Figure C36: Young schooling and GSP per worker for Gujarat

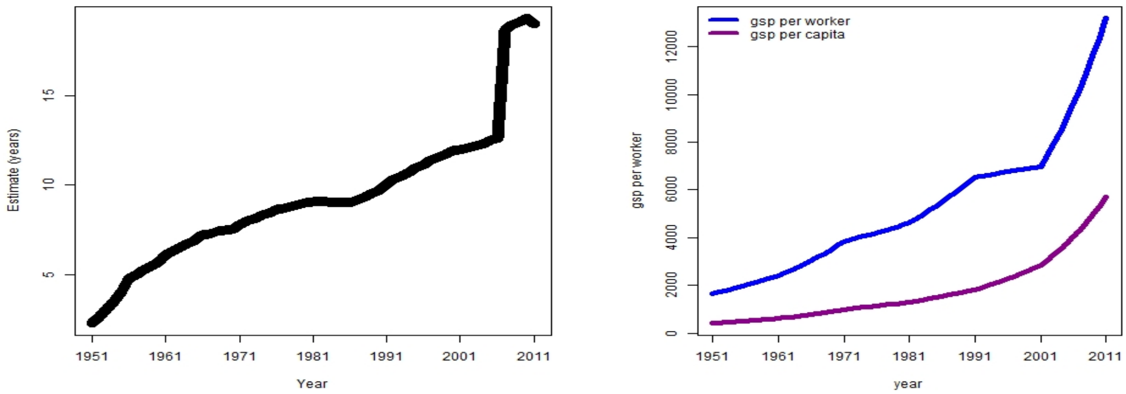


Figure C37: Young schooling and GSP per worker for Haryana

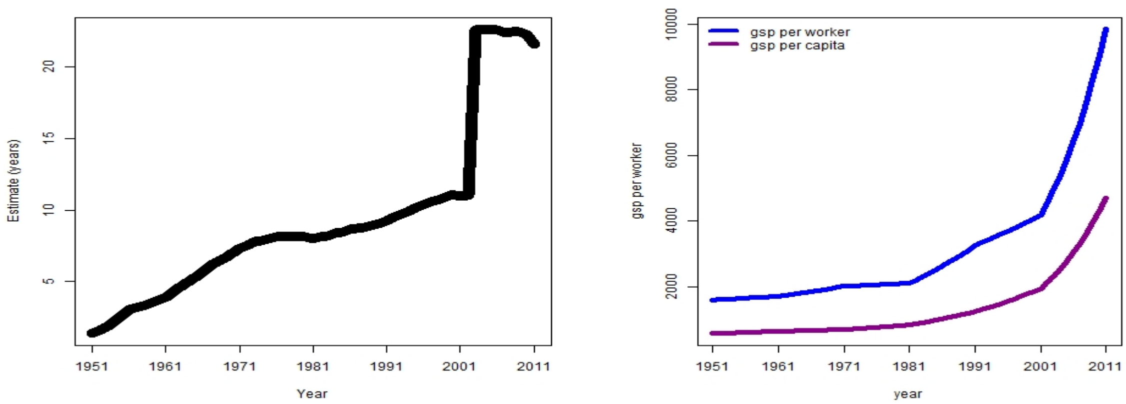


Figure C38: Young schooling and GSP per worker for Karnataka

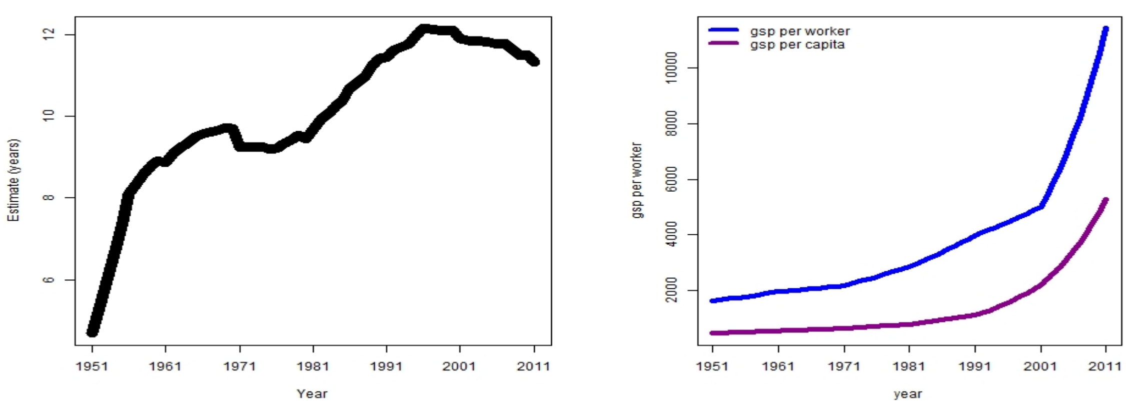


Figure C39: Young schooling and GSP per worker for Kerala

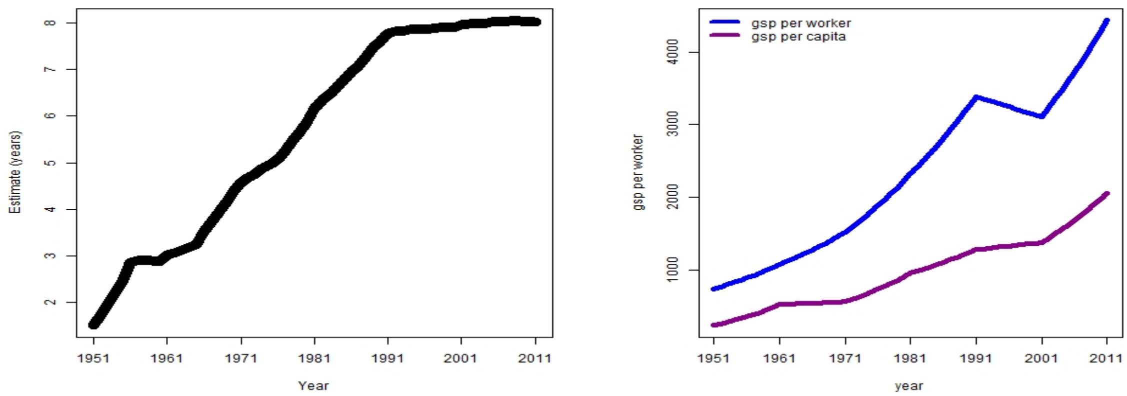


Figure C40: Young schooling and GSP per worker for Madhya Pradesh

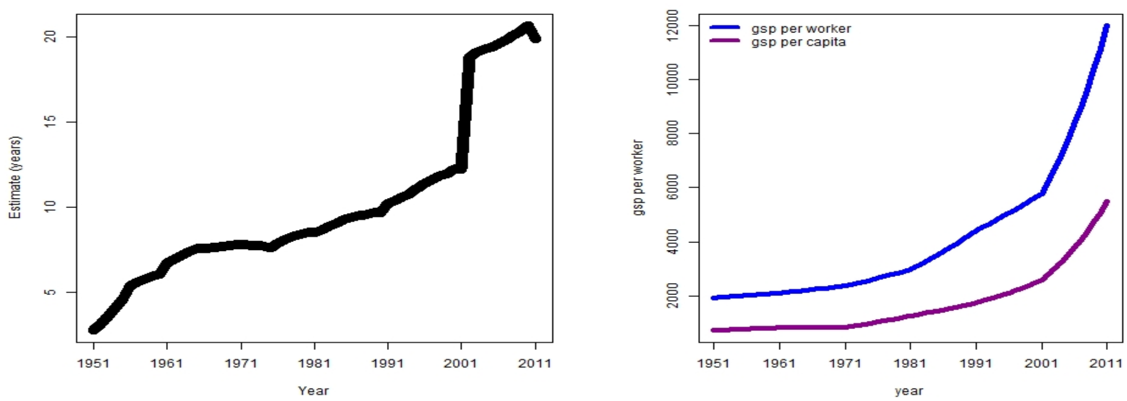


Figure C41: Young schooling and GSP per worker for Maharashtra

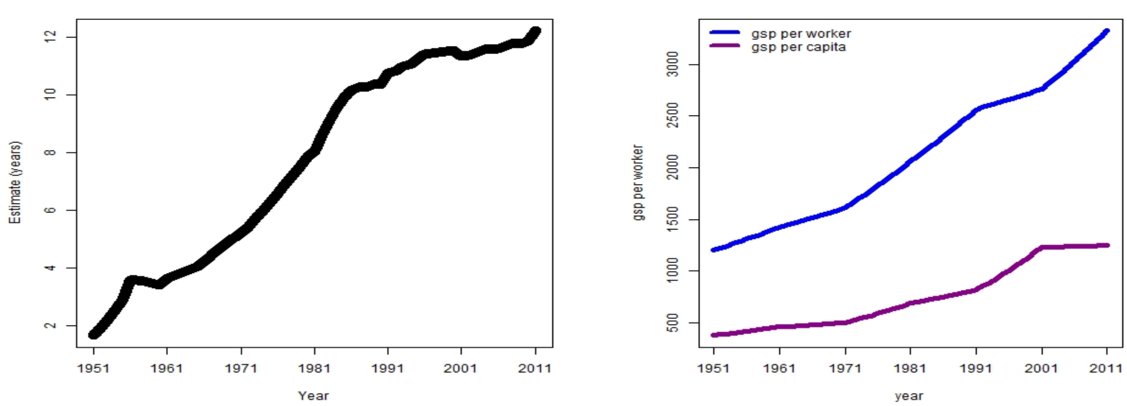


Figure C42: Young schooling and GSP per worker for Odisha

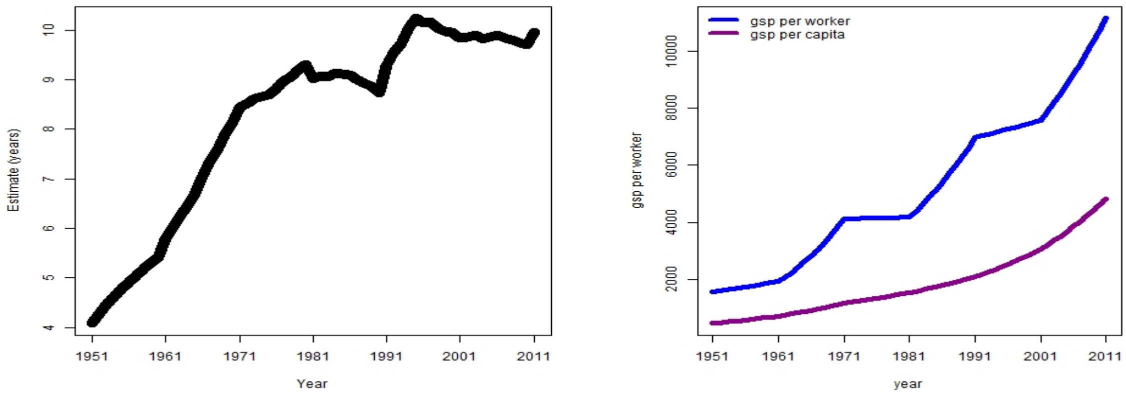


Figure C43: Young schooling and GSP per worker for Punjab

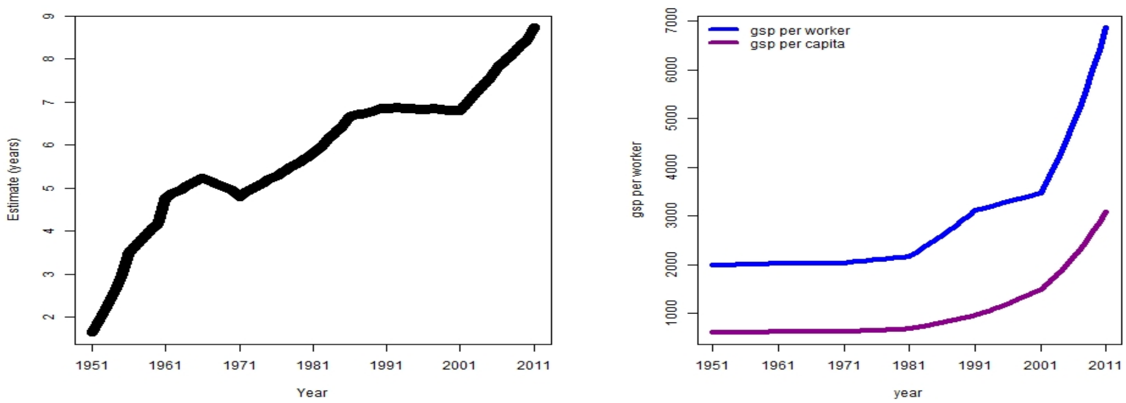


Figure C44: Young schooling and GSP per worker for Rajasthan

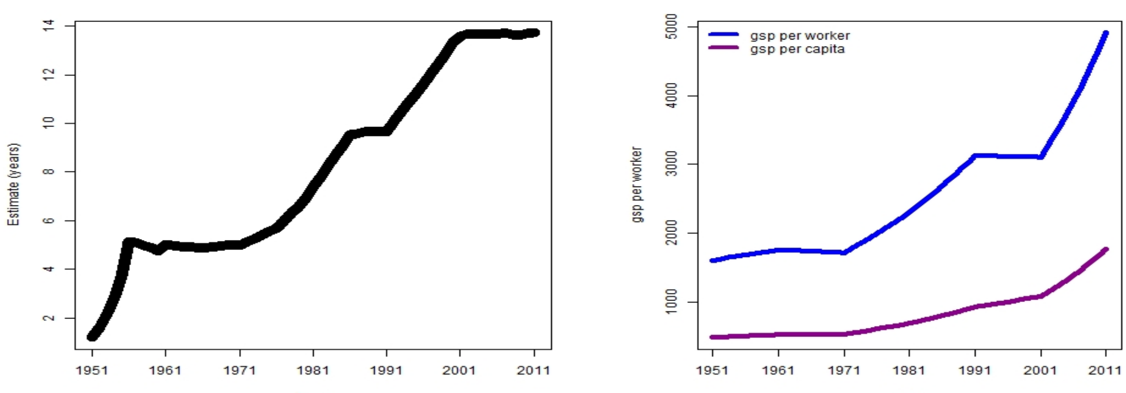


Figure C45: Young schooling and GSP per worker for Uttar Pradesh

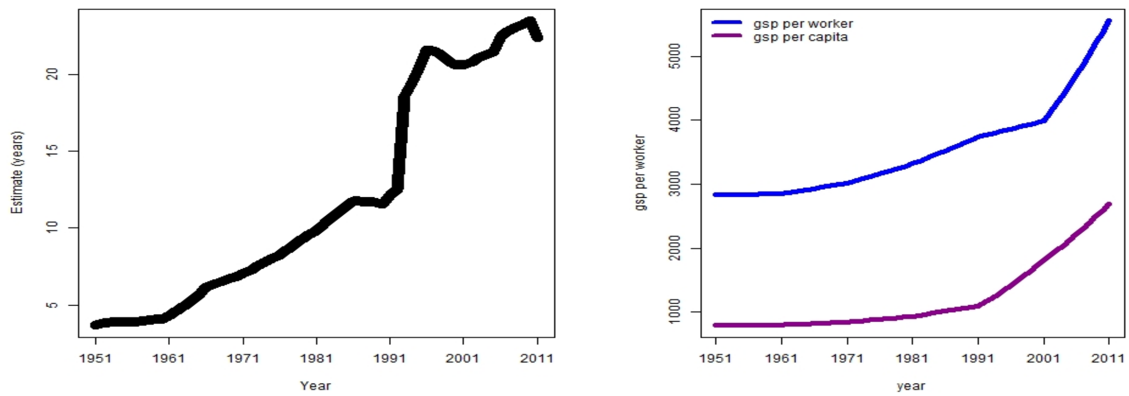


Figure C46: Young schooling and GSP per worker for West Bengal

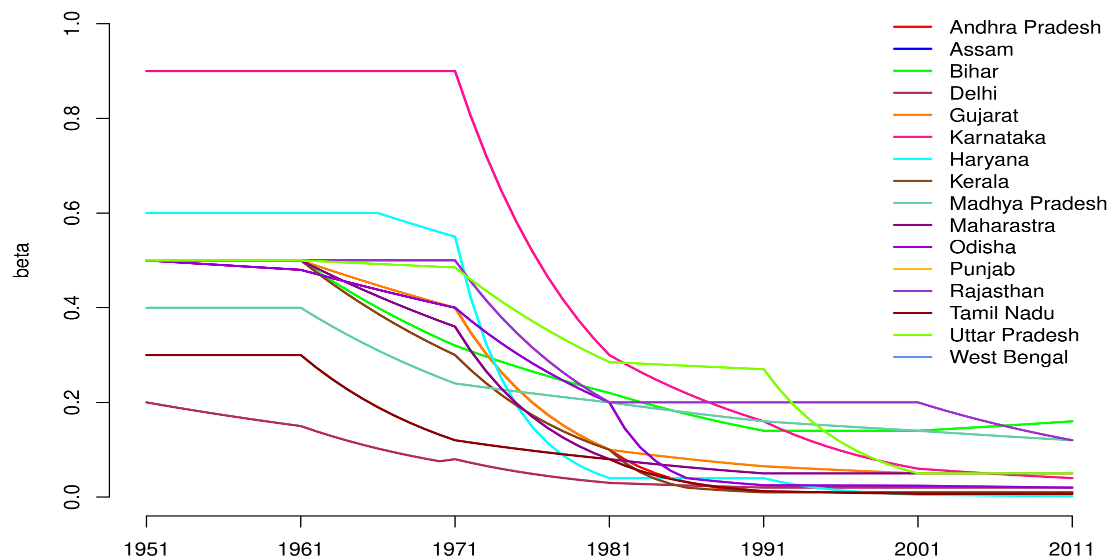
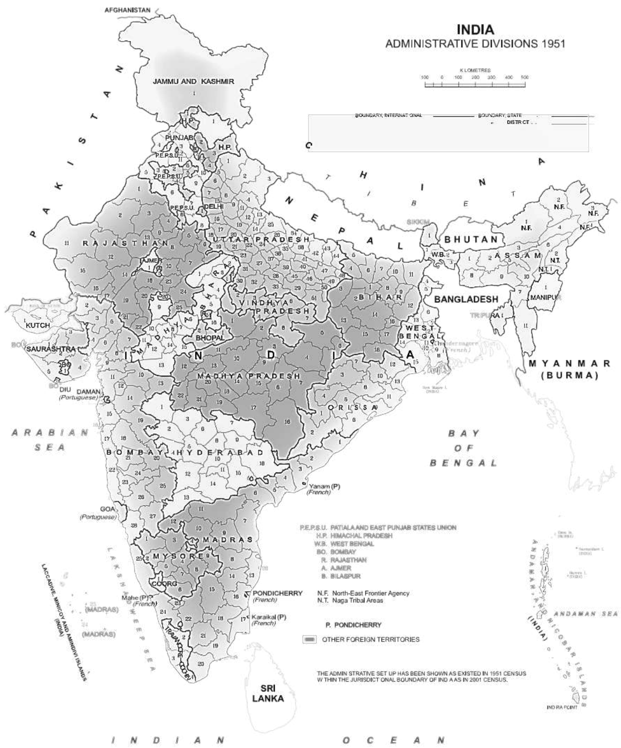


Figure C47: β by state

13 Miscellaneous figures

Figure M1: India administrative divisions- 1951 ©mapmyindia.com



Foreign Territories	French	Portuguese
	5	3

Administrative Divisions	INDIA	Ajmer	Assam	Bihar	Bilaspur	Bhopa	Bombay	Coorg	Delhi	Himachal Pradesh	Hyderabad	Jammu and Kashmir	Kutch	Madhya Pradesh	Madras
No. of Districts	310	1	17	18	1	2	28	1	1	4	16	1	1	22	26

Administrative Divisions	Madhya Bharat	Manipur	Mysore	Orissa	Patiala and East Punjab States Union	Punjab	Rajasthan	Saurashtra	Sikkim	Travancore-Cochin	Tripura	Uttar Pradesh	Vindhya Pradesh	West Bengal	Andaman and Nicobar Islands
No. of Districts	16	1	9	13	8	13	25	5	1	4	1	51	8	15	1

Figure M2: India in 2014 ©mapmyindia.com

india-political-map 2014.jpg

Figure M3: Labor force weighted average years of education by region

Figure M4: Gross State Product per Worker by Region (1951-2011)