CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 1086

November 2011

Mass Education or a Minority Well Educated Elite in the Process of Development: the Case of India

Amparo Castelló-Climent and Abhiroop Mukhopadhyay

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract

This paper analyses whether in developing countries mass education is more growth enhancing than to have a minority well educated elite. Using the Indian census data as a benchmark and enrollment rates at different levels of education we compute annual attainment levels for a panel of 16 Indian states from 1961 to 2001. Results indicate that if the reduction of illiteracy stops at the primary level of education, it is not worthwhile for growth. Instead, the findings reveal a strong and robust significant effect on growth of a greater share of population completing tertiary education. The economic impact is also found to be very large: if one percent of the adult population were to complete tertiary education instead of completing only primary, the annual growth rate could increase by about 4 percentage points. Moreover, we find that a one percentage change in tertiary education has the same effect on growth as a decrease in illiteracy by 13 percentage points. A sensitivity analysis shows the results are unlikely to be driven by omitted variables, structural breaks, reverse causation or atypical observations.

JEL classifications: I28, O11, O50
Keywords: Distribution of education, attainment levels, economic growth, panel data

This paper was produced as part of the Centre's Macro Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements

We would like to thank the participants of the 6th Annual Conference on Economic Growth and Development (New Delhi), 2011 International Growth Conference (New Delhi), 2011 Royal Economic Society Conference (London), 26th Annual Congress of the European Economic Association (Oslo) and seminar participants at the Indian School of Business, for their helpful comments and suggestions. We also thank Anagula Reddy of NUEPA for helping us compile the data. We acknowledge the financial support of the Planning and Policy Research Unit at the Indian Statistical Institute (Delhi). Amparo Castelló also acknowledges financial support from the Spanish Ministry of Education through ECO2008-02675 project and the Ramón y Cajal Programme and she is also grateful to the hospitality of the CEP at the London School of Economics, where part of this research was conducted. Abhiroop Mukhopadhyay wishes to thank Institute of Economic Growth where he was Sir Ratan Tata Senior Fellow when he was working on this paper.

Amparo Castelló-Climent was a visitor from the University of Valencia, Spain with the Macro Programme at the Centre for Economic Performance, London School of Economics Abhiroop Mukhopadhyay is with the Indian Statistical Institute, Delhi.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

According to United Nations, in 2007 about 72 million children in the world did not have access to education.[1] This striking number highlights that in poorer countries a large mass of population work as unskilled labor in low productivity activities. In such economies, governments face the dilemma of whether to focus on policies that extend education to those that are illiterates or on policies that increase the share of well educated workers, who could specialize in high-skill sectors; that have large productivity and fuel economic growth. In order to better understand whether in developing countries mass education is more growth enhancing than to have a minority well educated elite, this paper focuses on a particular developing economy and estimates the effect of different measures of human capital that capture the distribution of education as well as the influence of each level of schooling.

The conventional wisdom about the relevance of each level of schooling is that mass education is the key. A possible explanation for this belief is that the marginal return to schooling is found to be decreasing with the level of schooling (e.g. Psacharopoulos and Patrinos, 2004). Nevertheless, these studies do not take into account the fact that in many countries the majority of university graduates are employed in the public sector, in which the wages do not reflect its market value. Moreover, recent evidence suggests that the pattern of returns is changing and the rate of return to primary education may now be lower than that to post-primary levels of schooling. Colclugh et al. (2010) survey the new evidence that uses data from the 1990s and early 2000s for individual countries and show that in most studies the rate of return to an extra year of education is found to be increasing as the level of education rises. Likewise, recent estimates with cross-country data have also challenged the traditional view by showing an estimated rate of return to an additional year of schooling being higher at the secondary and tertiary levels than at the primary level (e.g. Barro and Lee, 2010). On the other hand, private and social returns of education may differ as well. As noted by Pissarides (2000), educated laborers can be engaged in activities with high private returns but located in sectors that are not growth-enhacing.

[1] The second United Nations Millennium Development Goal is to achieve universal primary education by 2015.

The goal of this paper is to analyze the importance of the composition of human capital in the less developed countries, which are characterized by a large share of population with no education and, therefore, a trade-off between literacy and high skill education may arise. Among the developing countries, India stands out as its governments have, at different points in time, prioritized extending primary schooling among the illiterates as well as in increasing the share of population with tertiary education. High quality engineering and technology-oriented institutions of higher education have been the aim of all the Indian governments since its independence in 1947. The high mass of illiterates along with a non negligible number of a highly educated elite makes India an important case study for how the shape of the distribution of education may affect the economic performance of an economy.

The case of India is also convenient since it is one of the few developing countries with good statistics on relevant variables. Data on real GDP and other determinants of growth for the main Indian states are available on a year basis for the period 1961-2001 (e.g. Besley and Burgess, 2000, 2004; Ghate and Wright, 2011). The advantage of using the cross-sectional and the temporal dimension of the data is that they can be used to estimate a panel data model that controls for state specific effects and, therefore, minimize any omitted variable bias in the analysis. Moreover, data quality varies enormously across countries and it is typically worse in low income countries. The use of sub-national level data also has the advantage that the different levels of education are more comparable across states within a country than across developing economies.

Whereas data on income measures and other relevant variables are available for the states of India on a yearly basis, there are no similar panel data for educational measures. We fill this gap by computing yearly data on educational attainments across the states of India. Specifically, we use the Indian Census as a benchmark, which contains decadal information on the educational levels across the states, and estimate annual observations using data on enrollment rates for different education levels and a variant of the perpetual inventory method that takes into account possible over-reporting of enrollments. We compute the share of population 15 years and above with no schooling, some primary, completed primary, completed middle, completed secondary and completed tertiary for 16 Indian States from 1961 to 2001. We also use these data to compute distributional measures, such as the Gini coefficient and the distribution of education by percentiles.

From a methodological point of view, this paper shows that in developing countries, measures commonly used in the literature such as the average years of education or the human capital Gini coefficient are not sufficient to assess the effect of the level of education on economic growth rates, since they are determined to a large extent by the huge mass of people with no education. For instance, in the case of India, the correlation between the average years of schooling and the share of illiterates is above 0.9. These measures, therefore, mainly pick up the influence of illiterates on the economic performance of the economies. Alternate specifications that include the share of illiterates, the average years of education among the literates and the Gini coefficient among literates also fail in extricating the effect. This is because these distributional statistics are driven by large proportions of people with low levels of education. Hence the average educational attainment is collinear to these popularly used distributional statistics.

We show that when the shares of educational attainment vary a lot in degrees of magnitude, using shares of attainment of each education level does a much better job in bringing out the effect of the distribution of education. Results indicate that the tertiary attainment level is the level of education that had the strongest contribution to the growth rates of the Indian states over the period 1961-2001. We find the result is robust across different specifications, the use of instrumental variables, holds with different frequencies used to compute the growth rates and survives splitting the data into sub-periods that reflect different stages of development of the Indian economy. Moreover, the economic impact is also found to be large: if one percent of the adult population were to complete tertiary education instead of completing only primary, the annual growth rate could increase by about 4 percentage points. We find that a percentage change in tertiary education (a re-allocation from primary and secondary level) has the same effect on the growth rate as a decrease in illiteracy by 13 percentage points.

Overall, we find that if the reduction of illiteracy stops at the primary level of education, the impact on the growth rates is very low. This is in contrast to micro-studies like Psacharapoulus and Patrinos (2004), who report the returns to education are usually higher for low levels of schooling and decline for higher levels of education. However, in line with our macro findings, Bosworth et al. (2007) also find that in India the returns of primary education are relatively lower than the average returns. Likewise, estimating a Mincerian wage equation using data from 2004-2005, Kingdon (2009) shows that, in India, the returns to education are convex, that is, the marginal return to each extra year of schooling raises with the level of education. These findings are consistent with a low quality educational system at the primary level which may lead to higher literacy but does not contribute to skill accumulation. In fact, teachers absenteeism and teacher negligence is common in many Indian schools (e.g. Kremer et al., 2005).

So far the traditional literature that empirically investigates the influence of human capital on economic growth has not emphasized the role of the composition of human capital. Instead, the most common approach has been the use of the average years of schooling of the adult population as a proxy of the stock of human capital (Benhabib and Spiegel, 1994; de la Fuente and Domenech, 2006; Cohen and Soto, 2007), or the Gini coefficient to analyze the effect of the distribution of education (e.g. Castello and Domenech, 2002).[2] However, this paper shows that an aggregate measure of education, such as the average years of schooling or the Gini coefficient, is not sufficient to assess the effect of education on growth in countries characterized by a high number of illiterates. We show that an increase in the average years of education may be the result of an increase in the share of individuals with primary education, secondary schooling or an increase in the share of population with a university degree, each of them with a different effect on the growth rates.

One of the few attempts to analyze the role of the composition of human capital is Vandenbussche, Aghion and Meghir (2006), which focuses on the relevance of tertiary education in innovation activities in a sample of OECD countries that are close to the technological frontier. Our results show that even in the less developed countries, tertiary

[2] The effect of human capital on economic growth has been under debate in the empirical literature. Bils and Klenow (2000) calibrate a model and find that most of the relationship between schooling and growth, found in Barro (1991), can be explained by a channel that goes from expected growth to schooling instead of from schooling to growth. Another challenging finding is that by Pritchett (2001), who in a growth accounting regression shows that the impact of growth in educational capital on growth of per worker GDP is negative. According to Pritchett, explanations for the absence of positive returns in education at the macro level could be that educated individuals work in unproductive sectors, that the supply of educated labor could have been expanded whereas the demand had been stagnated or that the quality of schooling has been so low that it has not increased cognitive skills or productivity.

education may be crucial in shaping the economic performance of a country.

We show some preliminary evidence on the channels through which human capital influences growth. As noted by Kocchar et al. (2006), in India, both manufactures and services are relatively concentrated in skill-intensive output. Furthermore, Arora and Bagde (2010) show that software exports are higher in the states with higher levels of human capital, as measured by the state level engineering baccalaureate capacity. In line with these findings, we explore the impact of different levels of education on sectoral growth rates. Interestingly, results show that different attainment levels affect the growth of each sector. We find that while tertiary education has been key for the impressive growth of the services sector, the industrial growth rate has been mainly affected by the share of population with secondary education.

The organization of the paper is as follows. In the next section we present the data used and the methodology to estimate annual educational attainments across the Indian states. In Section 3 we discuss some specifications and display the econometric models to be estimated. In Section 4 the main results are shown. The robustness of the findings are analyzed in Section 5. Section 6 presents preliminary evidence on the effect of the attainment levels on the sectoral growth rates. Finally, Section 7 discusses the conclusions reached.

2 Data and estimation of educational attainments

2.1 Data

The Indian Census is the most credible source of information on educational attainments across the states of India. It contains decade information on the educational levels of the population classified by age and sex. The educational categories include illiterates, incomplete primary, primary, middle, matriculation, higher secondary, non-technical diploma, technical diploma and graduates and above.[3] We take the five available data points in census ranging from 1961 to 2001 as benchmarks and compute the annual attainment levels with enrollment figures. The total number of students annually enrolled in primary (classes I-V), middle (classes VI-VIII) and secondary/higher secondary (classes IX-XII)

[3] Some of these categories are grouped differently for earlier census years.

are taken from *Growth of Enrolment in School Education 1950-51 to 1993-94* (*Planning, Monitoring & Statistics Division* at the *Ministry of Human Resource Development).* The number of enrollees in classes I-V, VI-VIII, IX-XII for the years 1994-2000 are sourced from annual publications *of Education in India (Government of India, Ministry of Human Resource Development, Department of Education, Planning, Monitoring & Statistics Division).* We use the population by age groups from the census to compute the gross enrollment ratios.[4] Data on enrollments in tertiary education (including, among other things, professional education and diplomas) are sourced from annual publications of *Education in India* (1965-1979) and *Selected Education Statistics* (1980s onwards).

Data on real net state domestic product and standard determinants of the economic growth rates for the main Indian states for the period 1960-2000 are taken from Besley and Burgess (2000, 2004) and updated by Ghate and Wright (2011).

2.2 Estimation methodology of annual educational attainments

We follow a perpetual inventory method to estimate annual attainments. The procedure involves taking data on educational attainments as benchmark stocks and using enrollment data, with appropriate lags, to measure the new entrant flows that add to the stock. Annual observations on school attainment for the population 15 and above are computed as follows. Let t only refer to census years. In our dataset, $t = 1961, 1971, 1981, 1991$ & 2001. Let $H_{j,t}$ denote the population 15 years and above for whom j is the highest level of education attained; $j = 5$ refers to complete tertiary, 4 complete secondary, 3 complete middle, 2 complete primary, 1 incomplete primary and 0 no schooling. *HIGH*, *SEC*, *MDL* and *PRI* are the gross enrollment ratios in tertiary, secondary, middle and primary, respectively.[5] The variable L_m refers to the total population aged m years old. For example, $L_{t,15}$ is the total population aged 15 years old at time t, $L_{t,20-24}$ is the total population ranging between 20 and 24 years old at time t, and so on. The variable δ_t is the mortality rate for the population 15 years and above between year t and $t - 10$ and

[4] The age groups for each educational level are 6-11 years old for primary, 11-14 years old for middle and 14-18 years old for matriculation and higher secondary.

[5] Due to classification problems, we cannot treat higher secondary as a separate level. We do not want to include higher secondary as incomplete tertiary as the effect of university/professional degrees may be disproportionately large as compared to higher secondary.

has been estimated by using the formula

$$\delta_t = 1 - \left(\frac{L_{t,15+} - L_{t,15-24}}{L_{t-10,15+}}\right)^{0.1}$$

We estimate annual completion ratios ($CR_{t,j}$) for each educational level using census data (see Appendix for the estimation method used to calculate them). Given the completion rates, the implicit annual stock for tertiary education is given by

$$H_{5,t+i} = H_{5,t} * (1-\delta)^i + \sum_{j=1}^{i} CR_{t,5} * HIGH_{t+j} * L_{t+j,21} * (1-\delta)^{i-j} \tag{1}$$

where the subscript i refers to each year within the decade for which census data is not available, with $i = 1, ..., 9$. Similarly,

$$H_{4,t+i} = H_{4,t} * (1-\delta)^i + \sum_{j=1}^{i} [(CR_{t,4} * SEC_{t+j} * L_{t+j,16}) - (CR_{t,5} * HIGH_{t+j} * L_{t+j,21})] * (1-\delta)^{i-j} \tag{2}$$

$$H_{3,t+i} = H_{3,t} * (1-\delta)^i + \sum_{j=1}^{i} [(CR_{3,t} * MDL_{t+j-1} * L_{t+j,15}) - (CR_{4,t} * SEC_{t+j} * L_{t+j,16})] * (1-\delta)^{i-j} \tag{3}$$

$$H_{2,t+i} = H_{2,t} * (1-\delta)^i + \sum_{j=1}^{i} [(CR_{2,t} * PRI_{t+j-5} * L_{t+j,15}) - (CR_{3,t} * MDL_{t+j-1} * L_{t+j,15})] * (1-\delta)^{i-j} \tag{4}$$

$$H_{0,t+i} = H_{0,t} * (1-\delta)^i + \sum_{j=1}^{i} (1 - PRI_{t+j-5}) * L_{t+j,15} * (1-\delta)^{i-j} \tag{5}$$

$$H_{1,t+i} = 100 - H_{0,t+i} - H_{2,t+i} - H_{3,t+i} - H_{4,t+i} - H_{5,t+i}$$

Our algorithm, by construction, ensures that we match the actual data attainments in the census years. Our method is also relevant for scenarios where enrollment figures are overstated, as is often the case with developing countries. Given that we force the completion ratios to be such that they have to be consistent with the initial attainment and the final attainment of the population every ten years means that they adjust the possible biases in reporting enrollment decade wise for each state. Thus, the algorithm we have followed is less subject to the criticism that confronts the perpetual inventory method used by Barro and Lee (2001).[6] We use census data available for five decades and follow the perpetual inventory method with the modification that it fits perfectly adjacent census years. Since we derive completion rates endogenously such that it fits census years, we are automatically correcting, for any decade, reporting bias in completion rates. The Barro-Lee (2010) is a better method to follow when there is not enough census data, since it does not require decade wise fitting. However, since census data are available for several decades in India, we use our method of fitting decade by decade.

Using the attainment data for each level of schooling we also compute a human capital Gini coefficient:

$$Gini = \frac{1}{2\overline{H}} \sum_{i=0}^{5} \sum_{j=0}^{5} \mid \widehat{x}_i - \widehat{x}_j \mid n_i n_j$$

where $\overline{H}$ are the average years of education of the population 15 years and above, i and j stand for the different levels of education: no schooling (0), incomplete primary (1), complete primary (2), complete middle (3), complete secondary (4) and complete tertiary (5); n_i and n_j are the shares of population with a given level of education, and $\widehat{x}_i$ and $\widehat{x}_j$ are the cumulative duration in years of each education level. We take 0 years for no schooling, 3 years for incomplete primary, 5 years for complete primary, 8 years for complete middle, 10 years for complete secondary and 15 years for complete tertiary.

The path of the share of illiterates and the completed attainment levels for primary,

[6] The perpetual inventory method by Barro-Lee (1993, 2001) has been criticized by Cohen and Soto (2007) and de la Fuente and Domenech (2006) because they show "implausible time series profiles of educational attainment for some countries". This had led to a new algorithm laid out in Barro Lee (2010) that uses critically the education levels by 5 years age intervals in the previous or subsequent five year periods. The problem with the Barro-Lee (1993, 2001) perpetual inventory method is that it depends crucially on enrollment rates. Data from a single census year is taken as a starting point and educational data for other years are then calculated by extrapolating, forward or backwards, using enrollment and completion rates. The new method laid out in Barro Lee (2010) provides better estimates than the perpetual inventory method because enrollment rates are often over stated in developing countries and many countries have data for only one census year.

middle, secondary and tertiary schooling from 1961 to 2001 for the 16 Indian states is displayed in Figure 1. Although all the states have reduced the share of illiterates and increased the population at all attainment levels of schooling, the evolution of the shares across the states is quite different. For example, in 1961 the share of population with no schooling or incomplete primary was more than 80 percent in almost all the states. In 2001, in spite of government focus on education, there were still differences across the states with a share of 24 percent in Kerala and 62 percent in Bihar. Overall, all the states have experienced extraordinary rates of growth in secondary and tertiary education, given the extremely low starting levels. In the case of tertiary education, the figure shows that in 1961 most of the states concentrated around a value of 0.4 percent. Among these states, Gujarat and Andhra Pradesh display values of tertiary education above 6 percent in 2001. The highest tertiary attainment levels in 2001 are found in Maharashtra and Karnataka, with shares of 8.16 percent and 7.54 percent, respectively. At the bottom of the distribution are Bihar, Rajasthan and Assam, with a share of tertiary schooling around 4.5 percent. In all figures Kerala stands out as an outlier state with much lower levels of illiteracy and much higher share of population with secondary and tertiary schooling than the rest of the states. Tables A1, A2 and A3 in the Appendix provide summary statistics and correlation among the main educational variables.

3 Empirical Model

In this section, we wish to specify models that allow us to test whether a state with a low average educational attainment but with the educated completing tertiary education, can grow faster than a state where education is more widespread but where the average education level among the educated is not very high. In other words, we would like to see the effect of the distribution of educational achievement in the population on economic growth after controlling for other covariates. In particular, we would like to investigate, as a thought experiment, whether a state can afford to have a higher proportion of people who haven't completed primary education, that is illiterate and people with incomplete primary education and yet grow faster because a larger share of its literate population has completed tertiary education. From here on, for lack of a better term, we will refer to those who are illiterate or haven't completed primary schooling as *illiterate*; the complement set of people we will term $literate$[7].

A specification that has been used before (e.g. Castello and Domenech, 2002) includes the average years of education and the Gini of education as additional regressors in an otherwise standard econometric model of growth. This standard model, often used in the context of cross country growth regressions, would regress growth rates on the usual covariates like capital stock, initial GDP and other variables, the choice of which differ depending on the focus of the paper. The inclusion of the Gini, controlling for average years of schooling, measures the distributional impact of education on growth.

To fix notation, let g_{it} denote the growth rate of per capita GDP,[8] y, of the ith state, between years t and $t+1$; denote the average years of schooling by Edu_{it} and the Gini coefficient of education by $Gini_{it}$. Let μ_i capture all state i specific time invariant heterogeneity and Z_{it} be the other observables which determine growth (which we discuss later). Then:

$$g_{it} = \alpha + \mu_i + \beta_1 y_{it} + \beta_2 Edu_{it} + \beta_3 Gini_{it} + \Pi Z_{it} + \xi_{it} \tag{6}$$

In equation (6), β_3 represents the effect of the distribution of education for a given average level of education. We also consider other variants of equation (6), where the Gini coefficient is replaced by, in one specification, the share of total education held by the top 1 percent of the educated people of the economy, and in another specification, by the top 10 percent. These alternate specifications are motivated by Voitchovsky (2005), who, in the context of income inequality, states that aggregate indicators of inequality, as measured by the Gini coefficient, could mask the different effect that the lower and upper part of the income distribution have on growth. We include only indicators of the upper part of the distribution since the large number of illiterates gives values equal to zero to the bottom percentiles. We refer to this group of regression equations as Model 1.

However, in economies with high proportion of illiterates, the average years of education and the Gini capture the same idea. Since a very small fraction of the population are

[7] This is not to ignore the importance of some primary education. However, it is tedious to refer to the composite groups by alternate names like *barely educated*, *etc.*For ease of presentation, we refer to those with very little or no education as *illiterates* and the share of population with primary, secondary and tertiary refer to complete attainment levels.

[8] We use the per capita Net Real State Domestic Product (NRSDP) from Ghate and Wright (2010).

educated, the total stock of education is concentrated. Thus, in countries with low educational attainment, all we may be able to capture with this specification is that illiteracy is bad. But this is not the focus of the paper.

To extricate better the effect of higher education among the educated, we consider the effect of the distribution of education among those who have at least completed primary schooling (*literates*), after controlling for, among other things, the proportion of *illiterate* people (S^{ILL}) and the average years of schooling among the *literates* (Edu^{LIT}). The inclusion of these controls keep the size of the pie constant. Thus, in Model 2, we estimate the following empirical model:

$$g_{it} = \alpha + \mu_i + \beta_1' y_{it} + \beta_2' S_{it}^{ILL} + \beta_3' Edu_{it}^{LIT} + \beta_4' Gini_{it}^{LIT} + \Pi' Z_{it} + \xi_{it}' \qquad (7)$$

Note, however, that in developing countries, where the education structure of society is very concentrated, even the use of the Gini among the *literate* ($Gini^{LIT}$) may not be very informative. This is because the share of population who complete and stop at lower levels of education will be enormous as compared to those who complete higher studies. Nevertheless, we estimate this specification to illustrate a methodological issue in model specification while dealing with countries/states where the distribution of education is very uneven and where one is interested in the effect of the education level held by very few.

As an alternative, we suggest a specification that may do a better job in extricating the impact of tertiary education: one that is held by very few. Consistent with the notation before, let us denote the share of the labor force with completed education level j by S^j, where $j = PRI, SEC$ and $TERT$. PRI denotes complete primary and middle schooling, SEC represents complete and incomplete secondary (includes higher secondary), and $TERT$ stands for complete tertiary education. Thus, in model 3 we estimate:

$$g_{it} = \alpha + \mu_i + \delta_1 y_{it} + \delta_2 S_{it}^{PRI} + \delta_3 S_{it}^{SEC} + \delta_4 S_{it}^{TERT} + \Pi'' Z_{it} + \xi_{it}'' \qquad (8)$$

We omit the share of *illiterates* in the labor force. Hence δ_i measures what would be the effect on growth if a unit share of the *illiterates* in the labor force were to acquire i *th* level of education. If education is useful in promoting growth, δ_i would always be greater than 0.

The parameters in (8) show the merit of having populations with different levels of education. However, they show the trade-offs with respect to the omitted category, i.e., *Illiterates*. The results of the estimation of this specification would essentially establish that more education is good (if all parameters are positive) and which education level is most productive for growth. For example, we would expect $\delta_4 > \delta_3 > \delta_2$. The main motivation behind this specification is to demonstrate that describing the distribution in shares is better when education is so concentrated among the top. It is of course important to point out here that equations (7) and (8) are not equivalent. However in so far, as model 3 describes more about the distribution, it conveys more information.

Building on these regressions, our objective is to compare two alternatives. First, we want to look at the impact of a higher share of tertiary education on growth, the share of illiterates kept constant. This is in spirit similar to a policy that would concentrate on those who have completed primary education and ensure that a larger proportion of them complete tertiary education.

Thus, in Model 4, we estimate the following equation:

$$g_{it} = \gamma_0 + \mu_i + \gamma_1 y_{it} + \gamma_2 S_{it}^{ILL} + \gamma_3 S_{it}^{TERT} + \Psi Z_{it} + \omega_{it} \tag{9}$$

As an alternative, we compare it with a scenario where more people are given education. Thus there is reduction in illiteracy. However, such a reduction imposes no assumption on which level of education the literates complete. Thus we estimate

$$g_{it} = \gamma'_0 + \mu'_i + \gamma'_1 y_{it} + \gamma'_2 S_{it}^{ILL} + \Psi' Z_{it} + \omega'_{it} \tag{10}$$

Using parameters γ_3 and γ'_2 in equations (9) and (10), we want to compare the effects on growth of lowering S_{it}^{ILL} in one scenario and raising S_{it}^{TERT} in another situation.

It may be contended though, that the effects of tertiary education on growth depend on what is the stock of illiterates in the economy. For example, there may be a threshold maximum share of illiterates above which the economy may not be responsive to increases in tertiary education. We allow for this flexibility by estimating a specification that includes an interaction term $S_{it}^{TERT} * S_{it}^{ILL}$ in addition to the variables in equation (9).

There are two major econometric issues that arise in the estimation of our specifications. The first one relates to the methods used to estimate dynamic panel data models.

The usual problems are dealt with using Arellano and Bond (1991) and Blundell and Bond (1998) estimators in the growth literature. These estimators involve using all available lagged values as instruments to take care of endogeneity that spring out of the dynamic panel structure. However, in our case, these methods are not useful as these estimators have been devised for problems where i is large and t is small. Indeed, the endogeneity problem is substantial mostly in the case of large i and small t. Since our data set has 40 years of data on 16 states, we have long t and small i. In this scenario, the inconsistency, if any, for using the fixed effect estimator in a dynamic model are likely to be small (e.g. Nickell, 1981). Thus we can run fixed effects methods to estimate all of our specifications.

What is potentially more problematic for our exercise is endogeneity even after we take into account fixed effects. What could be critical for us, especially since we use state level data, is endogenous choice of where to locate. While migration rates within India have been found to be low especially during the period of our sample (e.g. Munshi and Rosenzweig, 2009), it could be contended that migrants with tertiary education are more mobile and locate themselves in urban centres that are rich in the first place. Therefore, what we pick up as the effect of growth of tertiary education in our within estimator is the effect of income growth. Notice though, that controlling for initial income already factors out this confounding factor and the effect of tertiary education is over and above that. Thus, to validate our results, we investigate if growth of S_t^{TERT} depends on past growth rates of states. For example, it can be contended that people with tertiary education do not just settle down in states with higher income but in states that have grown faster in the previous periods. To investigate this, we regress S_t^{TERT} on $S_{t-\tau}^{TERT}$ and the lagged growth rate. If the coefficient of lagged growth rate is insignificant, this would further substantiate our result that growth of share of tertiary education causes and is not led by past growth.

We explore this possibility in Table 1, which displays the effect of lagged growth and lagged per capita income on the current share of population with tertiary education. Results show that the share of tertiary education is not determined by the previous growth rates, the coefficient of lagged growth is not statistically significant in any specification. Moreover, the lagged per capita income does not have a significant effect on tertiary education either, as displayed in columns (2) and (4).

Nevertheless, to purge our estimators of any additional endogeneity, in section 5 we test the robustness of the results by using instrumental variables. We use the lagged values of tertiary education as instruments. Furthermore, we also use the lagged value of the wage rate of individuals with tertiary education in urban areas. The identifying assumption is that larger wages for the population with tertiary education will incentivize individuals to acquire tertiary education but have no other direct effect on growth. We provide the Sargan p-value and the Cragg-Donald F-statistic to test for the validity of the instruments.

4 Main Results

In all regressions, in addition to the variables that capture the impact of education, we use rainfall during the year $t+1$, $Rainfall_{i,t+1}$, to proxy for agricultural shocks; its square, to allow for the non linear impact of rainfall on growth since too much rain can turn into floods that may be harmful for growth; the initial per capita income, y_t, to control for convergence in income across states; the adult (15+) population of the state POP_t^{15+}, to control for the labour force in the state and the rural population of the state, POP_t^R, to proxy for the size of the agriculture sector. Following previous work by Besley and Burguess (2000, 2004) and others, we use total expenditures, $TEXP_t$, and development expenditures, $DEXP_t$, to measure fiscal policy at the state level. As noted in the previous section, we control for state level heterogeneity by running fixed effects regressions.

We start by analyzing the influence that inequality in the distribution of education may have on the growth rates. As noted above, a standard formulation usually includes the average years of education and the Gini coefficient of education to capture the distributional impact of education on growth. The results, displayed in Table 2, show that the average years of education do have an impact on growth. The positive marginal effect of an unit increment of average education suggests that there are gains to education. However, the Gini coefficient of education is insignificant. As discussed before, this could be either because there are no distributional impacts on growth or because it is not possible to extricate the impact of a very small proportion of highly educated people using this specification. Recall that in most developing countries the large number of illiterates among the adult population suggests that typical distributional measures, such as the Gini coefficient, may not convey more information than what is captured by the low average years of education. Alternative specifications that include the share of education attained by the top end of the distribution- the share of the top 1 percent or the top 10 percent (columns (2) and (3))- are not statistically significant either.

As an alternative, we compute the Gini coefficient and the average years of education among the literate and introduce the share of the labor force that are illiterates as a regressor. The aim is to look at whether the Gini coefficient among the literates is significant after we separate out the effect of illiterates and control for the average years of education among those literate. We find that controlling for the aggregate education in society (given by the share of the illiterates and the average years of schooling among the literates), the Gini coefficient among the literates is still not statistically significant (column (4)). However, the effect of the share of illiterates is negative and statistically significant at 1 percent level. The estimated coefficient of the share of illiterates is -0.33 and an increase in one standard deviation in the share of illiterates (0.147) is associated with a 4.8 percentage points reduction in the growth rates. Likewise, results also show that the coefficient of the average years of education among the literates is positive and significant. The estimated coefficient of 0.026 implies that the effect of an increase in one standard deviation in Edu^{Lit} (0.59) is associated with an increment in growth by 1.5 percentage points. The significance of Edu^{Lit} and share of illiterates hints that the distribution of education may indeed matter but this specification is still opaque on whether distribution among literates matter.

The effect of the additional controls on growth are, reassuringly, as expected. We find that greater rainfall has a positive impact on economic growth. An increase in the total revenue expenditure discourages the growth rates whereas greater development expenditure has a beneficial influence. Results also show that a greater share of population living in rural areas is negatively related to economic growth whereas a larger labour force has a positive effect on growth. Finally, in line with Nagaraj et al. (2000), we also find evidence of conditional convergence across the states; the coefficient of initial per capita income is negative and significant at 1 percent level. Most of these results are robust across other specifications in the paper.

In order to better extricate the distributional impact of education, we look at the specification which includes the share of the labor force with different levels of education, as stated in Model 3. In Table 3 the share of all schooling levels are included and the share of illiterates is the omitted category. Therefore, the coefficient of each share is its trade-off with S^{ILL}. Results in column (1) show that the coefficient of complete primary education is insignificant. This is not surprising given the low quality of primary education in India (e.g., Kremer et al., 2005). Just completing primary education is no different in economic terms than being illiterate or having some incomplete primary education.

Whereas complete primary education seems to have an insignificant impact on the growth rates, we find a strong impact of higher levels of education. The coefficient of tertiary education is positive, significant and quite large in absolute value. However, the significance of secondary education alternates with the significance of the coefficient of S^{TERT}, which points out to a collinearity problem. Indeed, the correlation between S^{SEC} and S^{TERT} is 0.940. We confirm the collinearity problem by dropping both variables, one at a time. When we drop S^{TERT}, S^{SEC} becomes positive and significant and vice versa for S^{TERT} in column (3). That the share of tertiary education has a larger marginal impact can be gauged by a much larger magnitude of its coefficient across all specifications (the coefficient for tertiary education ranges from 4.001 to 4.383). For example, using coefficients in column (1), a one standard deviation increase in S^{TERT} (0.019) is associated with a 7.6 percentage point increase in the growth rates. Nevertheless, one must keep in mind that this large marginal coefficient reflects a re-allocation of the share from illiterates to tertiary. The marginal effect of the share of secondary education is lower (column (2)) at 0.962 and a standard deviation increase in the share of complete secondary education (0.055) increases growth rates by 5.3 percentage points.

These estimation results point out to the usefulness of using the shares of each level of education. In contrast to results that posit that increasing average years of education matters, we find a more nuanced result that only complete secondary and higher education is productive. The use of shares seems to be an important part of finding distributional impacts in context of developing countries where conventional specifications that include measures of distribution (like Gini, top 10 percent) seem to give insignificant results.

As pointed out before, our aim in this paper is slightly different. We want to compare two scenarios: one where illiteracy is reduced but there are no restrictions made to which level of education the literates reach, and the second, where the share of literates is held constant but a greater proportion of people with complete post primary education also complete tertiary education. Our next empirical exercises are motivated by this question.

Columns (4-6) show the estimation results of Model 4. In column (4), we can calculate the returns to a lower share of illiterates with no restrictions on whether the literates complete at most primary level, a tmost secondary level or tertiary education. An economy with 10 percentage lower illiteracy in this scenario grows faster by 3.3 percentage points.[9]

Alternatively, consider the results in column (5). As described before, with this estimation result, we can study the returns to an increase in tertiary education with the share of illiterates held constant. An increment of one standard deviation of the share the tertiary education (0.019) raises growth rates by as much as 8.23 percentage points. Recall that this derivative reflects an increase in the share of post primary educated people who complete tertiary education. We also find that this result does not depend on the level of literacy in society. The interaction term in the estimation of equation (10) turns out to be insignificant. In addition, we find that the share of illiterates turns out to be insignificant in columns (5) and (6). This is consistent with earlier results that show that there is an insignificant impact if the *illiterate* complete only primary schooling.

Given these results, what can one say about the trade-off between a large proportion of lowly educated adult population and a small proportion of literate but highly educated population. Comparing columns (4) and (5), we can compute the changes in tertiary education and illiteracy levels which yield the same change in growth rates. Using our results, we find that a percentage change in tertiary education (a re-allocation from primary and secondary level) has the same effect on the growth rate as a decrease in illiteracy by 13 percentage points[10].

[9] This return is a weighted average of the returns to completed levels of primary, secondary and tertiary education, with the weights implicitly being driven by what, empirically, is the proportion of these incremental literate in each education category.

[10] While we do not seek to do a fiscal cost-benefit analysis in this paper, it may be useful to have an idea of the cost of these trade-offs at an all India level. A one percent increase in tertiary education completion would cost at least 47 billion rupees (Rs 7117 per student at 1993-94 prices), while a 12 percent points increase would cost at least 106 billion (1229 rupees per student). Using per student real costs for elementary and tertiary education given in "Report on Working Group on Higher Education - 11th Five Year Plan" (http://www.aicte-india.org/downloads/higher_education_XIplan.pdf)

5 Robustness of the results

Instrumental Variables

To minimize any biases due to reverse causation, we test the robustness of the results by using instrumental variables. Our identification strategy consists in using the fourth and fifth lags as instrument for tertiary education after checking the exclusion restriction holds. Further, we also use the lagged value of the wages of university-level degree employees in urban areas, as higher past wages will encourage individuals to acquire tertiary education.

Table 4 displays the estimates of Models 3 and 4 using instrumental variables. In line with the fixed-effect estimator, the findings reveal a strong effect of tertiary education on economic growth. Likewise, the estimated coefficient of primary and secondary schooling are not statistically significant in any specification. It is worth noting that the use of instrumental variables increases the estimated coefficient of the attainment levels in tertiary education, which leads to the suggestion that previous findings were not driven by reverse causation.

The strong effect of the share of population with tertiary education on the economic growth rates is also remarkable when we estimate Model 4. Columns (5-7) show that the only variable that is statistically significant in almost all specifications is the share of population with tertiary education. The economic significance of an increase in the share of population with tertiary education is also relevant. According to columns (5) and (7), controlling for the share of illiterates and holding other things constant, results imply that an increase in one standard deviation in the share of population with tertiary education (0.019) increases growth rates between 11.46 and 15.94 percentage points. In other words, if one percent of the people with primary or secondary education in the labor force were to complete tertiary education, the annual growth rate would increase between 6.03 and 8.39 percent. Although the effect seems to be huge, the share of population with tertiary education is still very low in India and even increasing one percent level has taken several years. In our sample, the average attainment level in tertiary education for the 16 states in 2001 was only 6.1 percent.

The validity of the results depends on the accuracy of the instruments. The bottom part of the table shows the number of lags used as instruments and the tests for the validity of the instruments. Results suggest that the instruments are not weak, the F-statistic is very high and above 10 in all cases. When we have multiple endogenous variables, to test for weak identification we use the Cragg-Donald F-statistic. The null hypothesis is that a given group of instrument is weak against the alternative that it is strong. The Cragg-Donald F-test is very high in all cases and higher than the critical value in Stock and Yogo (2005), suggesting that in all the specifications we can reject the null hypothesis that instruments are weak.[11] A greater concern is whether the instruments are orthogonal to the error term in the second stage since the lagged values of the educational variables may not be a truly source of exogenous variation. However, the Sargan p-value of the null hypothesis that the instruments are uncorrelated with the error term cannot be rejected at the 1 percent level in most of the cases, indicating that the excluded instruments are statistically valid.[12]

Overall, the estimates with the lagged values of wages are larger than those when the lagged values of the tertiary attainment levels are used as instruments. In addition to the differences regarding the source of variation and the degree of correlation with the error term in each case, the sample period is also different. The wages of the tertiary educated population are only available from 1980 onwards. Therefore, when this variable is used as an instrument, the period of analysis is 1981-2001 instead of 1961-2001. Next we analyze the robustness of the results to different sample periods.

Period Break-down

It has been contended that India's growth process underwent a structural change sometime in early 1980s.[13] To take into account possible structural breaks, we estimate our

[11] The critical value is a function of the number of included endogenous regressors, the number of instrumental variables, and the desired maximal bias of the IV estimator relative to the OLS. For example, with 2 endogenous regressors and 4 instruments, the critical value based on 2SLS maximum bias of the IV estimator relative to OLS of 0.05 at significance level of 5 percent is 11.04. The null of weak instruments is rejected in the case that the Cragg-Donald F-statistic on the excluded instruments exceeds this critical value.

[12] When we use wages as instruments the model is just identified and the Sargan p-value is not provided. In this case we test for the exclusion restriction by checking that the wages of the population with tertiary education is not statistically significant related to growth once the share of population with tertiary education is controlled for.

[13] The exact time of the turn around in the growth rates of the Indian economy and the possible explanations that caused such change have been the issue of several academic papers. See for example, Rodrik and Subramanian (2005), Virmani (2006), Basu (2008) and Chetan and Wright (2011)

equations separately for the period 1961-1981 and 1981-2001. The first period represents a planning era in Indian history, one marked by a low annual rate of growth. The second period has been one of rapid growth with the services sector showing substantial growth. In our sample, the average growth rate of per capita NRSDP during the period 1981-2001 (sample average of 2.93%) is almost twice the average growth rate during the years 1961-1981 (sample average of 1.49%).[14] With the changes the economy experienced during these years, it is possible the effect of the composition of education on economic growth has also changed over time.

The estimation results in Table 5 yield results similar to those discussed before. The average years of education, the share of illiterates and the average years of education among the literate are significant in both sub-periods (Models 1 and 2). While the returns to a unit increase in the average years of education is 0.12 (column 1) between 1961-1981, the marginal effect of a similar increase is larger at 0.15 in the latter period. Similarly, controlling for the average education level among literates, lowering the share of illiterates has a larger impact in the second sub-period. However, as before, most standard variables that represent inequality of education (Gini, top 10 percent, Gini^{LIT}) are insignificant in both periods.

To further unravel what may have changed, we look at the estimation results of the returns to various levels of education (Model 3). Notice that for the period 1961-1981, the share of lower levels of education (complete primary and secondary education) are significant. A unit change in the share of complete primary education increases growth rates by 0.18 percentage points. However, tertiary and secondary education are still important. When we omit secondary education due to collinearity problems, we find that the marginal impact of tertiary education is significant and has a magnitude of 5.976.

The results of the latter period 1981-2001 resonate completely with the overall result. We find that the effect of tertiary education is robustly significant across all specifications. Moreover, the marginal effect of complete primary education is again insignificant. This suggests that there has indeed been a structural break in the returns to education for growth. While in the earlier period, the whole distribution of education mattered for

[14] This is different from the average growth rate of India as a whole. First we have omitted the smaller states. Second, the sample average treats each state equally, whereas a national average would weigh each growth rate by the share of the state in the national GDP.

growth, what matters for the latter period is the share of tertiary education.

These observations become clearer when we look at results for Model 4. We notice that for the period 1961-1981, keeping constant the share of tertiary education, a decrease in the share of illiterates (an increase in the share with complete primary and secondary schooling) increases growth rates by 0.2 percentage points. However, even for this period, the returns to an increment in the share of tertiary education (a movement from primary/secondary to tertiary) are high. Indeed they are higher than for the latter period 1981-2001 when the share of tertiary education is the only education level that matters[15]. For the latter period, echoing results for the whole period, the share of illiteracy is insignificant reflecting that a unit decrease in *illiterates* towards higher schooling levels did not yield any return for growth rates.

Another result that our period wise break-down yields is the significance of the interactive term for the period 1961-1981. While the coefficient of the term linear in S^{TERT} is negative, the partial effect at the mean value of S^{ILL} is 7.11. Thus, states with a large S^{ILL} showed a larger marginal effect with respect to share of tertiary education. In the latter period, the interaction term is insignificant and again the results are similar to what we get for the whole period.

Five-year and ten-year average growth rates

To increase the number of observations in the sample and, therefore, the accuracy of the estimates, we have computed annual growth rates. However, yearly growth rates may incorporate short-run disturbances and increase the noise in the estimates. We test for the robustness of the results by averaging the growth rates over five-year and ten-year period to reduce yearly serial correlation from business cycles.

In Table 6 the growth rates are computed as annual averages over the next five or ten years, $\Delta \ln y_{t,t+5}$ or $\Delta \ln y_{t,t+10}$, and the explanatory variables are measure in period t. This specification also reduces endogeneity caused by simultaneous determination of variables since explanatory variables are measured at the beginning of the 5-year or 10-year period. The specifications include all the explanatory variables in previous tables, however, to save space we only show the estimated coefficients of the educational variables

[15] The higher marginal effect of tertiary education in 1961-1981 is to a large extent a lower base effect. Indeed, returns to a standard deviation increase in tertiary education is lower during 1961-1981 than 1981-2001.

in each model.

The upper part of Table 6 shows the results of estimating Models 1 and 2. The results are qualitatively similar to those with annual data. On the one hand, the average years of schooling have a positive and significant effect on growth, whereas a greater share of illiterates is associated with lower growth rates. On the other hand, the measures of inequality in the distribution of education suggest that education inequality is not significantly related to growth in the Indian states.

Likewise, the estimated results of Model 3, displayed in the middle part of the table, also confirm the previous findings regarding the predominant effect on growth of a higher share of population with tertiary schooling, as compared with the attainment levels in primary and secondary education. However, results show that when the annual growth rate is computed as an average over a five-year or a ten-year period, the estimated coefficients are smaller. For example, using the results in columns (1) and (4), a one standard deviation increase in the share of population with tertiary education increases the growth rates between 3.5 (column (4)) and 4.4 (column (1)) percentage points, as opposed to an increase by 7.6 percent if estimates are computed with annual observations.

Finally, the bottom part of the table shows the estimates of Model 4. Results show that whereas the positive and highly significant effect of the attainment levels in tertiary education is quite robust and holds across all the specifications, the negative effect of the share of illiterates is not that robust. In contrast to the estimates of Model 2, the findings reveal a positive coefficient for the share of illiterates in the estimation of equation (9). The main difference is that Model 2 accounts for the average years of education among the educated population and, therefore, controls for all the levels of schooling, whereas Model 4 only accounts for tertiary education. Thus, it might be the case that the share of illiterates is also picking up the negative effect of the lower levels of schooling (see results in column (4) in Model 3).

Variation across-states

When variables are highly persistent, first differences and fixed effects techniques, while controlling for unobservable heterogeneity, may increase the measurement error bias by increasing the variance of the measurement error relative to the variance of the true signal (Griliches and Hausman, 1986). In addition, the use of higher-frequency data may exacerbate these problems (e.g. Pritchett, 2000). In Table 7 we exploit the variation across the states to test for the robustness of the results. We estimate equations (6-9) with pooled OLS.[16] Overall, the results are qualitatively similar to those found with the fixed effect estimator, which suggests that results not only hold within a state across time but also across states. Thus, other things being equal, the states with higher share of population with tertiary education have experienced, on average, higher growth rates.

Presence of outliers

Finally, we test the robustness of the results to the presence of outliers. In Tables 2 and 3, we identify the observations whose residuals are higher than twice the standard errors of the estimated residuals. We include dummies for the states and years identified as atypical observations and the results scarcely change.[17] Overall, we find a strong and highly significant effect on growth of a greater share of population with tertiary education.

6 Composition of Human Capital and Sectoral Growth Rates

The channels through which human capital influences growth are not analyzed in the paper. Nevertheless, a possible explanation for the relevant influence of tertiary education is that highly educated workers have been employed in sectors with high productivity. As preliminary evidence, we analyze the effect of the levels of schooling on the growth rates of per capita income in different sectors.[18] In particular, in Table 8 we analyze the effect of primary, secondary and tertiary education on the growth rates of per capita income in the industry (columns 1-4) and service (columns 5-8) sectors. The results suggest a differential effect of each attainment level on the growth rates of the sectors. Interestingly, columns

[16] Estimation with Random Effects shows similar results.

[17] We identify two types of outliers, those whose residuals are higher than the absolute value of two times the standard errors of the estimated residuals and those whose residuals are lower than that value. The states with a higher value are Andhra Pradesh (1972), Bihar (1995), Gujarat (1987 and 1991), J&K (1970), Madhya Pradesh (1966 and 1979), Orissa (1974, 1976 and 1987) and Rajasthan (1969 and 1987). The states with a lower value include Andhra Pradesh (1975), Bihar (1994), Gujarat (1971 and 1973), Haryana (1967), J&K (1964 and 1986), Madhya Pradesh (1978), Maharashtra (1999), Orisa (1989) and Rajasthan (1967, 1970, 1973 and 1978).

[18] We use Ghate and Wright (2011) sector shares in the Net State Domestic Product (NSDP) and, assuming the same price deflator for all sectors, compute each sector per capita Net Real State Domestic Product (NRSDP).

(1-4) show that an increment in the share of population with secondary education is the attainment level with the greatest effect in the industry sector; the coefficient of secondary education is positive and statistically significant even when tertiary education is controlled for. However, a greater share of population with tertiary education is what matters for the growth rates in the service sector. Other things being equal, an increment in one standard deviation in the share of population with tertiary education (0.019) is found to have increased the growth rate in the service sector by about 10 percentage points.

As pointed out in the literature, the growth rate in the service sector has been the primary source of the increase in India´s growth (e.g. Bosworth, et al., 2007). This is in contrast, for example with China, in which the growth of the industrial sector has also played a fundamental role in the growth upsurge (e.g. Bosworth and Collins, 2008). In India, in year 2000, the share in total income of each sector was 30 percent of agriculture, 23 percent of industry, and 47 percent of services. Thus, the large share of the service sector in the total income, relative to that of the industry, and the large influence of the university-level graduates in the growth of this sector could explain why the population with tertiary education have played a fundamental role in explaining the growth rates of total income in the Indian states.

7 Conclusions

The link between human capital and growth is a well established one. However, there has been less emphasis on the distribution of education among the literates and its effect on growth. Most of the work has been done with data from OECD countries and there has been less emphasis on developing countries. Among developing countries, India has, at various points in its history, emphasized both setting up of tertiary institutions as well as reducing illiteracy by attempting universal access to primary education. Given variations between states and over time in illiteracy rates and the share of adult population with tertiary education, we seek to investigate the distributional impact of education on growth in this context. Data for developing countries, especially time series, are difficult to get for most countries and are many a times non comparable. For example, a particular level of education across countries may reflect very different quality of knowledge that is accumulated. Thus, working with data on Indian states is an attempt to look within a comparable data set. We investigate the link between the distributional impacts of education and growth using data from 16 major states of India for the period 1961-2000.

Many developing countries are characterized by very skewed distributions with large proportion of the labor force who are illiterate and a relatively tiny proportion of people who complete tertiary education. In this paper, we show that the usual measures of inequality like the Gini or the education level of the top 1% or top 10%, are not able to extricate the important effect of tertiary education. The reason is a statistical one: given the large proportion of illiterate people, the average education that controls for the size of the pie is collinear with most commonly used distribution descriptives like the Gini or the education level of the top 1%. We find that in such scenarios, using shares of the adult population with different levels of education yield better results. Our result lays a case for its use for other similar contexts in developing countries.

While data from India is better than from most other developing countries, annual data on educational attainments are not available. We adopt a perpetual inventory method, which uses census data on educational attainment every ten years and enrolment figures reported annually, to construct the annual series for educational attainment. Our method deals with often reported biases in enrollments by using an algorithm such that it fits consecutive census data on attainments. We use this data to estimate a fixed effects model of growth. In the process we also take into account some endogeniety issues. In particular, we are aware of the possibility of endogenous location of tertiary educated people. We take care of this using appropriate instruments.

Our results show that a reduction of the share of illiterates, even if the new literates complete primary education, has scarcely an impact on growth. On the contrary, we find a strong and robust effect of increasing the population with tertiary education. Whereas the result may indicate a low quality primary education system, it should be viewed in light of the strategy of many developing countries that stress on getting children to school to make them literate. A relevant example is that of India which has recently enforced the "Right to Education" that lays the ground for universal access to primary education. While this is important, our results point out that there are equally important and higher economic gains from focussing on school and tertiary education completion. This is not to understate the role of illiteracy which we find has a negative impact in the economy. However, empowering the existing share of literates with skills and education that a tertiary education brings is another growth strategy that cannot be ignored.

In this paper, we do not explore, in a substantive way, the pathways of how education affects growth. Preliminary results suggest that the effect depends on the sectoral composition of growth. The interesting result that different levels of education are important for growth of different sector suggests several research questions. Do different sectors grow because of existing education policies? Is inequality of education another feature of kuznetsian sector dynamics? Or are our results specific to countries that have grown through the services sector, in contrast to others, like China, where the industrial sector has a larger role in determining growth? Further investigation into these links is a natural extension for future research.

References

[1] Arellano, M. and S. R. Bond (1991). "Some Test of Specification for Panel Data: Monte Carlo Evidence and and Application to Employment Equations". *Review of Economic Studies*, 58 (2), 277-97.

[2] Arora, A. and S. K. Bagde (2010). "Human Capital and the Indian Software Industry," NBER Working Paper No. 16167.

[3] Barro, R. (1991). "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106(2), pp. 407-43.

[4] Barro, R. and J-W Lee (1993). "International Comparisons of Educational Attainment," *Journal of Monetary Economics*, 32, 363-94.

[5] Barro, R. and J-W Lee (2001). "International Data on Educational Attainment Updates and Implications," *Oxford Economic Papers*, III, 541-63.

[6] Barro, R. and J-W Lee (2010). "A New Dataset of Educational Attainment in the World, 1950-2010," NBER Working Paper N^o 15902.

[7] Basu, K. (2008). "The Enigma of India´s Arrival: A Review of Arvind Virmani´s *Propelling India: From Socialist Stagnation to Global Power*," *Journal of Economic Literature*, 46 (2), pp. 396-406.

[8] Benhabib, J. and M. Spiegel (1994). "The Role of Human Capital in Economic Development." Journal of Monetary Economics, 34 (2), pp. 143-174.

[9] Besley, T. and R. Burgess (2000). "Land Reform, poverty Reduction, and Growth: evidence from India," *Quarterly Journal of Economics*, 115(2), pp. 389-430.

[10] Besley, T. and T. Burgess (2004). "Can Labor Regulation Hinder Economic Performance? Evidence From India," *Quarterly Journal of Economics*, 19(1), pp. 91-134.

[11] Bils, M and P. Klenow (2000). "Does Schooling Cause Growth." American Economic Review, 90(5), pp.1160-1183.

[12] Blundell, R. and S. Bond (1998). "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models". *Journal of Econometrics*, 87, pp. 115-43.

[13] Bosworth, Barry, and Susan M. Collins (2008). "Accounting for Growth: Comparing China and India." *Journal of Economic Perspectives*, 22(1), pp. 45–66.

[14] Bosworth, B., S. Collins and A. Virmani (2007). "Sources of Growth in the Indian Economy," NBER working Paper N^o 12901.

[15] Castelló, A. and R. Doménech (2002). "Human Capital Inequality and Economic Growth: Some New Evidence". *Economic Journal*, 112 (478), pp.187-200.

[16] Cohen, D. and M. Soto (2007). "Growth and Human Capital: Good Data, Good Results," *Journal of Economic Growth*, 12(1), pp.51-76.

[17] Colclough, C.; G. Kingdon and H. Patrinos (2010). "The Changing Pattern of Wage Returns to Education and its Implications," *Development Policy Review*, 28(6), pp. 733-47.

[18] de la Fuente, A. and R. Doménech (2006). "Human Capital in Growth Regressions: How Much Difference Does Data Quality Make," *Journal of the European Economic Association*, 4(1), pp.1-36.

[19] Ghate, Ch. and S. Wright (2011). "The "V-Factor": Distribution, Timing and Correlates of the Great Indian Growth Turnaround," *Journal of Development Economics,* forthcoming.

[20] Griliches, Z. and J. Hausman (1986). "Errors in Variables in Panel Data". *Journal of Econometrics,* 31, 93-118.

[21] Kingdon, G. G. (2009). "The Economic Benefits of Education in India," mimeo.

[22] Kochhar, K., U. Kumar, R. Rajan, A. Subramanian and I. Tokatlidis (2006). "India's Pattern of Development: What Happend, What Follows?" *Journal of Monetary Economics*, 53, pp. 981-1019.

[23] Kremer, M, K. Muralidharan, N. Chaudhury, J. Hammer and F. H. Rogers (2005). "Teacher Absence in India," *Journal of the European Economic Association*, 3, pp. 658-67.

[24] Miguel, E., S. Satyanath and E. Sargenti (2004). "Economic Shocks and Civil Conflict: An Instrumental Variable Approach," *Journal of Political Economy*, 112(4), pp.25-53.

[25] Munshi, K. & M. Rosenzweig (2009) "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth," *NBER Working Papers* 14850, National Bureau of Economic Research

[26] Nagaraj, A., A. Varoudakis and M. A. Veganzones (2000). "Long-Run Growth Trends and Convergence Across Indian States," *Journal of International Development*, vol. 12, pp. 45-70.

[27] Nickell, S. (1981). "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, pp. 1417-26.

[28] Pissarides, Ch. (2000). "Human Capital and Growth: A Synthesis Report," OECD Development Centre Working Paper N^o 168.

[29] Pritchett, L. (2000). "Understanding Patterns of Economic Growth: Searching for Hills among Plateaus, Mountains and Plains," *The World Bank Economic Review*, Vol. 14 (2), pp. 221-250.

[30] Pritchett, L. (2001). "Where Has All the Education Gone?" *The World Bank Economic Review*, 15(3), pp. 367-91.

[31] Psacharopoulos and H. Patrinos (2004). "Returns to Investment in Education: A Further Update," Education Economics, Vol. 12 (2), pp. 111-34.

[32] Rodrik, D. and A. Subramanian (2005). "Fom "Hindu Growth" to Productivity Surge: The Mystery of the Indian Growth Transition," IMF Staff Papers, Vol. 52, Number 2.

[33] Stock, J. H., and M. Yogo (2005). "Testing for Weak Instruments in Linear IV Regression" In Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, ed. D. W. Andrews and J. H. Stock, pp. 80–108. Cambridge University Press.

[34] Vandenbussche, J., P. Aghion and C. Meghir (2006). "Growth, Distance to Frontier and Composition of Human Capital." *Journal of Economic Growth*,11, pp. 97-127.

[35] Virmani, A. (2006). "Propelling India: From Socialist Stagnation to Global Power," Academic Foundation.

[36] Voitchovsky, S. (2005) "Does the Profile of Income Inequality Matter for Economic Growth?: Distinguishing between the effect of inequality in different parts of the income distribution," *Journal of Economic Growth*, 10, 273-96.

8 Appendix

8.1 Estimation of Completion Ratios

We use the census data and assume that completion ratios hold constant over the decade. As an illustration, let us explain how the completion ratio for tertiary education is calculated. If we want to compute the completion ratio between census years t and $t-10$ ($CR_{5,t}$), we start with the stock of people with tertiary education in the year $t-10$ ($H_{5,t-10}$). Given mortality rates δ, $(1-\delta)^{10}$ proportion of them will survive by the year t. Next, we look at how many new entrants are added to this stock. We assume that the average student completes tertiary education by the age of 21. Thus we need to account for how many people in the cohort $21-30$ in the year t have completed tertiary education[19]. Let $\overline{HIGH}_{t,t-10}$ be the average enrollment ratios over the decade. Thus, the new stock of tertiary educated people over the decade would be $CR_{5,t} * \overline{HIGH}_{t,t-10} * L_{t,21-30}$. Given everything else, the completion ratio for tertiary educational level for the decade between t and $t-10$ is then given by the following formula:

$$CR_{5,t} = \frac{H_{5,t} - (1-\delta)^{10} * H_{5,t-10}}{L_{t,21-30} * \overline{HIGH}_{t,t-10}} \tag{11}$$

For calculating the other completion ratios, we use the following assumptions. Student finish primary schooling by the age of 11, middle schooling by the age of 14, secondary school by 16[20]. While $\overline{SEC}_{t,t-10}$ is the average enrollment rate for secondary schooling, $\overline{MDL}_{t-1,t-11}$ is the average middle school enrollment between the years $t-1$ and $t-11$. Similarly $\overline{PRI}_{t-5,t-15}$ is the average enrollment rates in primary education for each decade. The last two enrollment rates are constructed based on lags with different start and end years because the minimum age to be counted in our attainment figures is 15. For example, it takes people who pass out of middle schooling one year to be counted. Thus for the cohort of age 15, what matters for the completion ratio is what was the enrollment rate among them one year back. Arguments analogous to the one made for tertiary education yield the following:

$$CR_{4,t} = \frac{H_{4,t} - (1-\delta)^{10} * H_{4,t-10} + L_{t,21-30} * \overline{HIGH}_{t,t-10} * CR_{5,t}}{L_{t,16-27} * \overline{SEC}_{t,t-10}} \tag{12}$$

$$CR_{3,t} = \frac{H_{3,t} - (1-\delta)^{10} * H_{3,t-10} + L_{t,21-30} * \overline{SEC}_{t,t-10} * CR_{4,t}}{L_{t,15-24} * \overline{MDL}_{t-1,t-11}} \tag{13}$$

$$CR_{2,t} = \frac{H_{2,t} - (1-\delta)^{10} * H_{2,t-10} + L_{t,15-24} * MDL_{t,t-10} * CR_{3,t}}{L_{t,15-24} * \overline{PRI}_{t-4,t-14}} \tag{14}$$

[19] One advantage of looking at the 21-30 cohort in the year t is that it already takes into account mortality rates over the period.

[20] Due to variations over time and across states in higher secondary, we do not look at that education category separately

Appendix

Table A1
Summary Statistics

	1961-2001					1961-1981					1981-2001				
	Obs.	Mean	Std. Dv	Min	Max	Obs.	Mean	Std. Dv.	Min	Max	Obs.	Mean	St.Dv	Min	Max
$\Delta \ln y$	641	0.023	0.073	-0.224	0.383	321	0.015	0.078	-0.224	0.314	336	0.029	0.067	-0.186	0.383
lny	642	8.650	0.385	7.691	9.636	322	8.447	0.260	7.691	9.113	336	8.842	0.383	7.911	9.636
S^{years}	656	2.986	1.273	0.697	7.111	336	2.159	0.833	0.697	5.191	336	3.811	1.065	1.953	7.111
Gini_h	656	0.679	0.118	0.305	0.910	336	0.742	0.093	0.406	0.910	336	0.617	0.106	0.305	0.792
$\mathrm{S}_{Years}^{Literates}$	656	7.964	0.590	5.946	9.135	336	7.559	0.486	5.946	9.122	336	8.369	0.348	7.577	9.135
$\mathrm{Gini}_h^{Literates}$	656	0.184	0.013	0.132	0.207	336	0.182	0.013	0.132	0.204	336	0.187	0.013	0.154	0.207
$\mathrm{S}^{Illiterates}$	656	0.662	0.147	0.218	0.969	336	0.751	0.111	0.385	0.970	336	0.573	0.120	0.218	0.780
$\mathrm{S}^{\Pr imary}$	656	0.214	0.089	0.013	0.534	336	0.178	0.085	0.013	0.465	336	0.249	0.077	0.130	0.534
$\mathrm{S}^{Secondary}$	656	0.096	0.055	0.009	0.264	336	0.056	0.027	0.009	0.126	336	0.135	0.045	0.047	0.264
$\mathrm{S}^{Tertiary}$	656	0.029	0.019	0.002	0.088	336	0.014	0.008	0.002	0.036	336	0.043	0.014	0.016	0.088
Rainfall	650	1.297	0.676	0.292	4.003	330	1.283	0.687	0.317	4.003	336	1.310	0.667	0.292	3.431
Total Expenditure	603	0.263	0.420	0.002	2.954	331	0.035	0.035	0.002	0.224	288	0.515	0.496	0.030	2.954
Development Expenditure	603	0.168	0.257	0.001	1.644	331	0.024	0.025	0.001	0.143	288	0.329	0.298	0.022	1.644
Population15+	656	25.938	17.815	2.108	102.922	336	19.840	12.669	2.108	64.639	336	31.991	19.919	3.532	102.922
Rural population	652	32.720	23.228	2.965	137.363	332	27.038	17.770	2.965	90.690	336	38.297	26.336	4.709	137.363

.

Table $A2$
State-wise Means

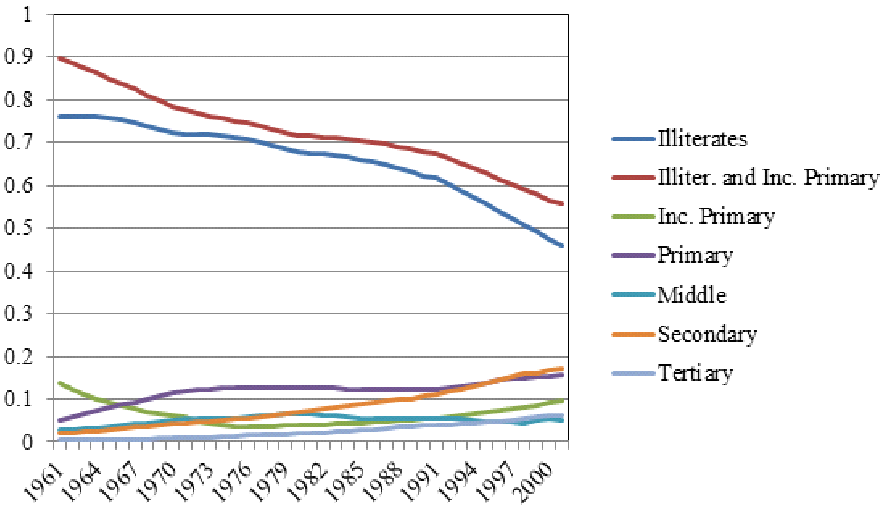
State	$\mathbf{S}^{Illiterates}$	$\mathbf{S}^{Primary}$	$\mathbf{S}^{Secondary}$	$\mathbf{S}^{Tertiary}$	$\mathbf{S}^{years}$	$\mathbf{Gini}_h$
ANDHRA PRADESH	0.723	0.170	0.081	0.026	2.397	0.744
ASSAM	0.650	0.246	0.085	0.020	2.998	0.663
BIHAR	0.768	0.132	0.079	0.021	2.187	0.774
GUJARAT	0.631	0.230	0.108	0.031	3.247	0.639
HARYANA	0.647	0.209	0.113	0.032	3.067	0.691
JAMMU KASHMIR	0.723	0.153	0.098	0.027	2.507	0.754
KARNATAKA	0.649	0.212	0.107	0.032	3.121	0.672
KERALA	0.395	0.427	0.138	0.040	5.207	0.410
MADHYA PRADESH	0.761	0.152	0.062	0.024	2.232	0.753
MAHARASHTRA	0.568	0.272	0.121	0.039	3.814	0.602
ORISSA	0.726	0.198	0.055	0.021	2.454	0.709
PUNJAB	0.603	0.223	0.138	0.035	3.436	0.656
RAJASTHAN	0.780	0.137	0.062	0.021	2.039	0.779
TAMIL NADU	0.613	0.249	0.111	0.027	3.365	0.626
UTTAR PRADESH	0.740	0.153	0.082	0.025	2.357	0.758
WEST BENGAL	0.620	0.256	0.089	0.036	3.351	0.638
Total	0.662	0.214	0.096	0.029	2.986	0.679

Table A3
Correlations among the main variables

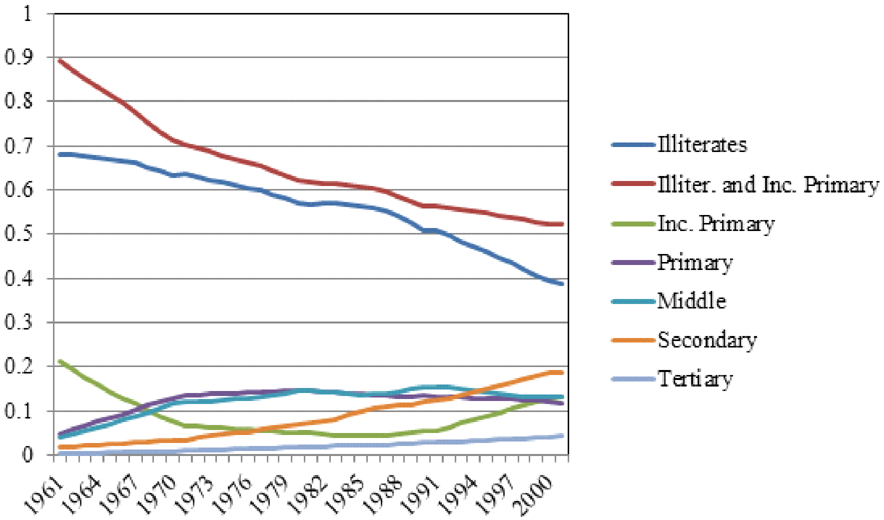
	$\ln y$	$\Delta \ln y$	S^{years}	Gini	S^{LIT}	$Gini^{LIT}$	S^{TERT}	S^{SEC}	S^{PRI}	S^{ILL}
$\ln y$	1									
$\Delta \ln y$	0.011	1								
S^{years}	0.731	0.090	1							
Gini	-0.643	-0.075	-0.963	1						
S^{LIT}	0.456	0.062	0.491	-0.283	1					
$Gini^{LIT}$	0.308	0.079	0.123	-0.101	0.067	1				
S^{TERT}	0.781	0.110	0.895	-0.764	0.728	0.256	1			
S^{SEC}	0.787	0.098	0.916	-0.797	0.707	0.110	0.940	1		
S^{PRI}	0.533	0.060	0.878	-0.919	0.109	0.047	0.624	0.650	1	
S^{ILL}	-0.714	-0.087	-0.986	0.951	-0.420	-0.102	-0.854	-0.884	-0.928	1

Note: 656 Observations.

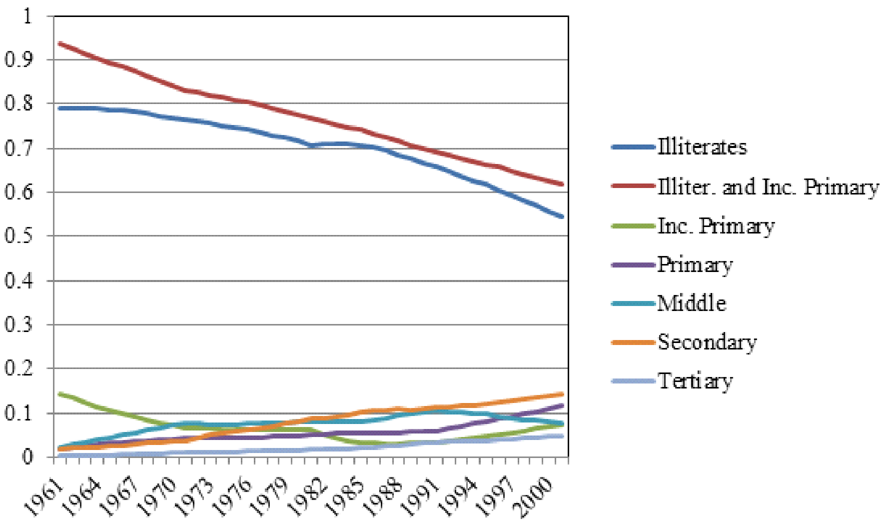
ANDHRA PRADESH



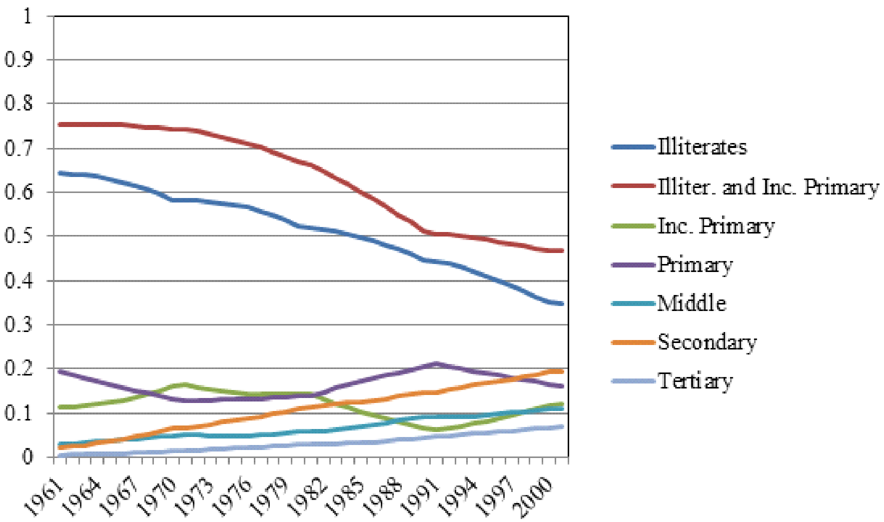
ASSAM



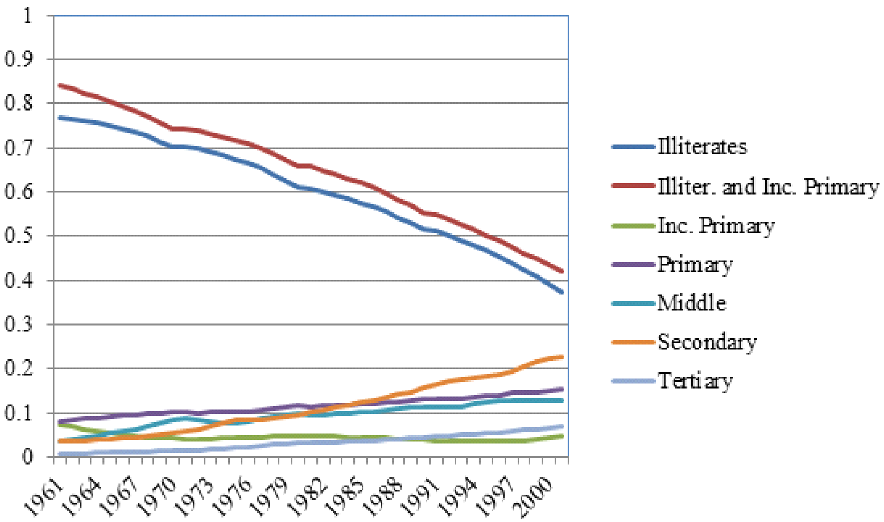
BIHAR



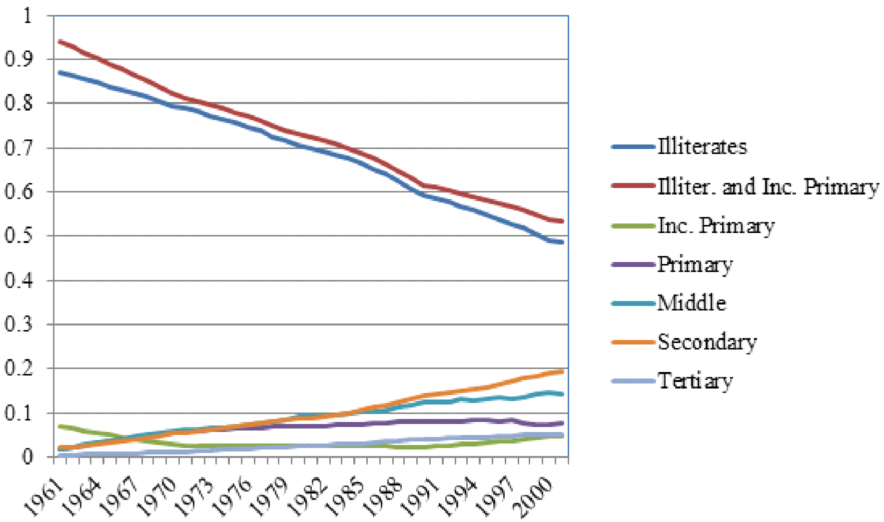
GUJARAT



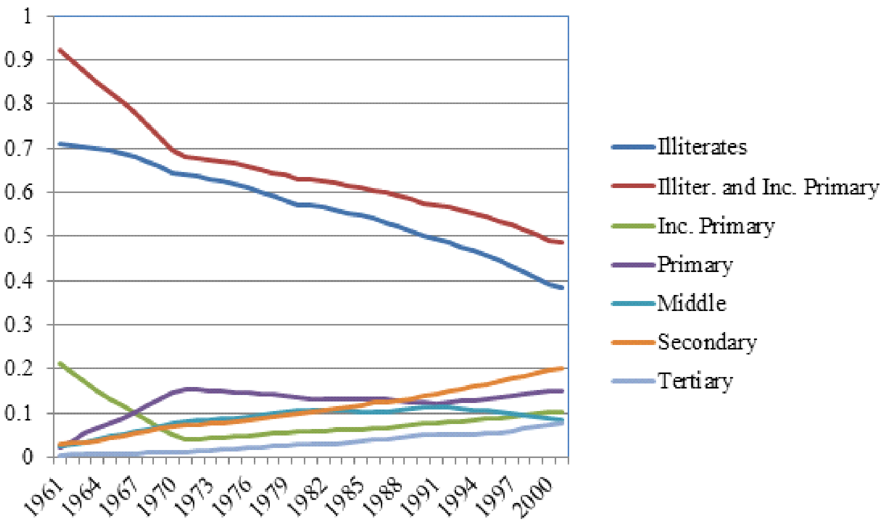
HARYANA



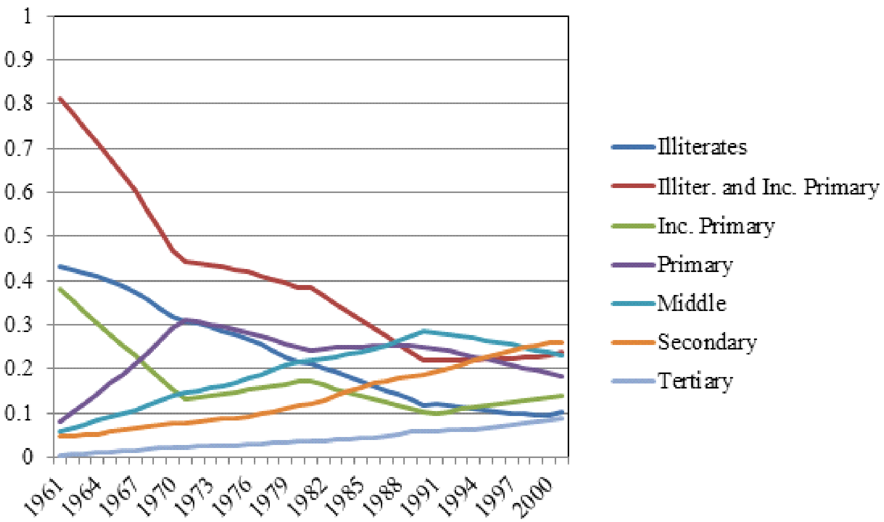
JAMMU AND KASHMIR

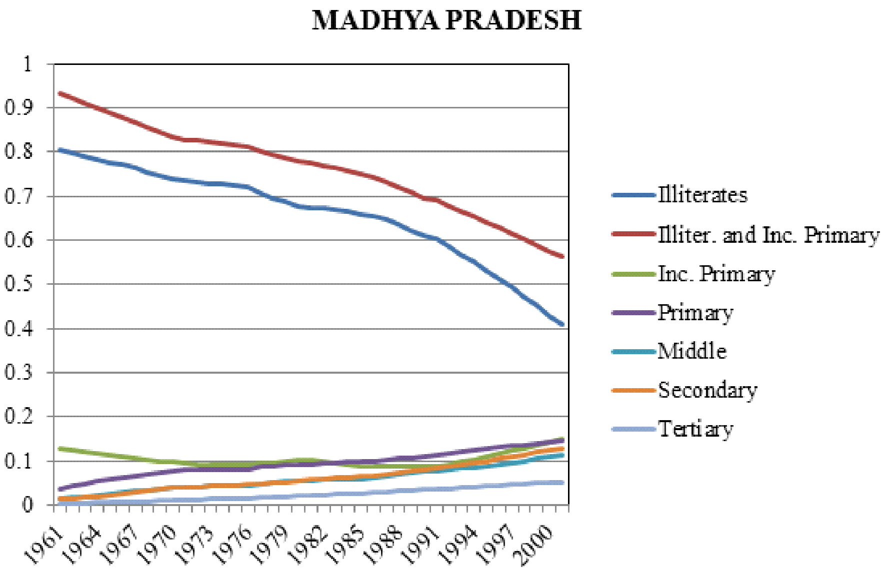


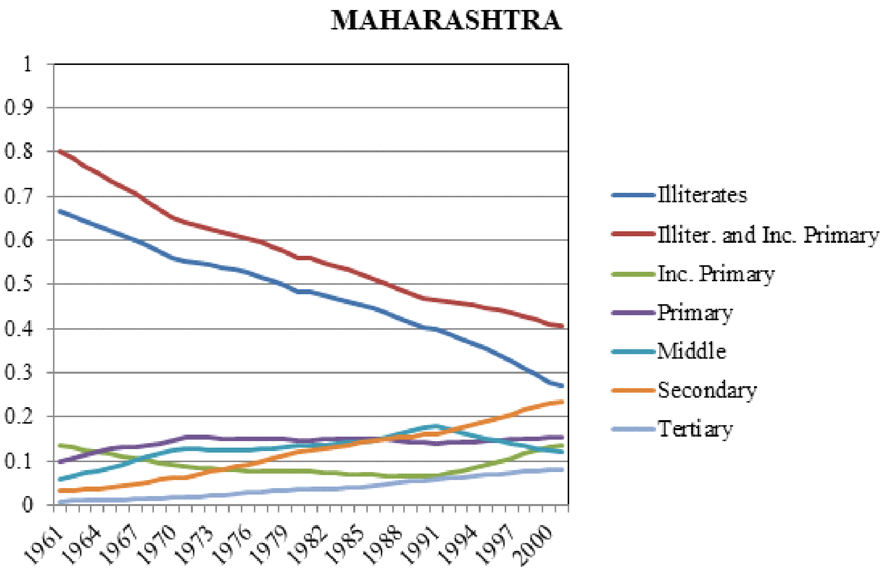
KARNATAKA

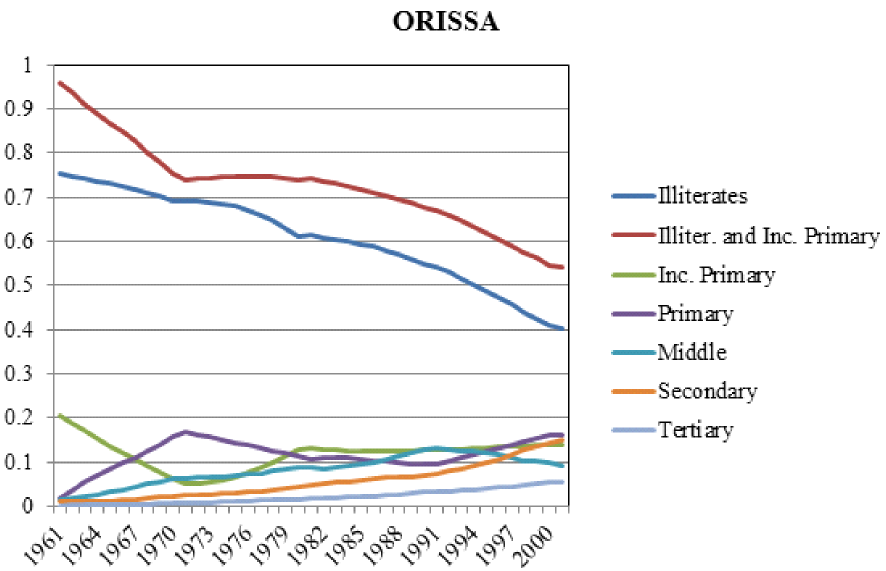


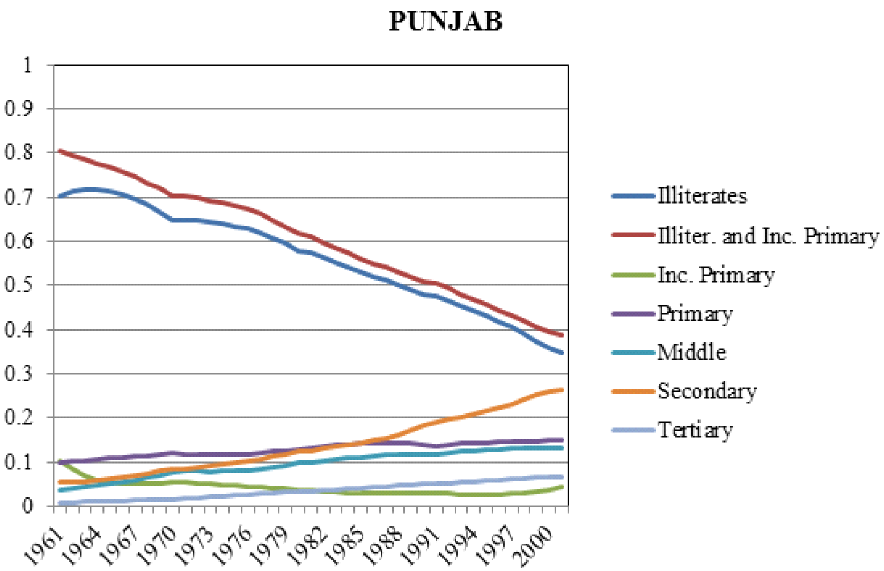
KERALA




MADHYA PRADESH
Illiterates
Illiter. and Inc. Primary
Inc. Primary
Primary
Middle
Secondary
Tertiary


MAHARASHTRA
Illiterates
Illiter. and Inc. Primary
Inc. Primary
Primary
Middle
Secondary
Tertiary


ORISSA
Illiterates
Illiter. and Inc. Primary
Inc. Primary
Primary
Middle
Secondary
Tertiary


PUNJAB
Illiterates
Illiter. and Inc. Primary
Inc. Primary
Primary
Middle
Secondary
Tertiary

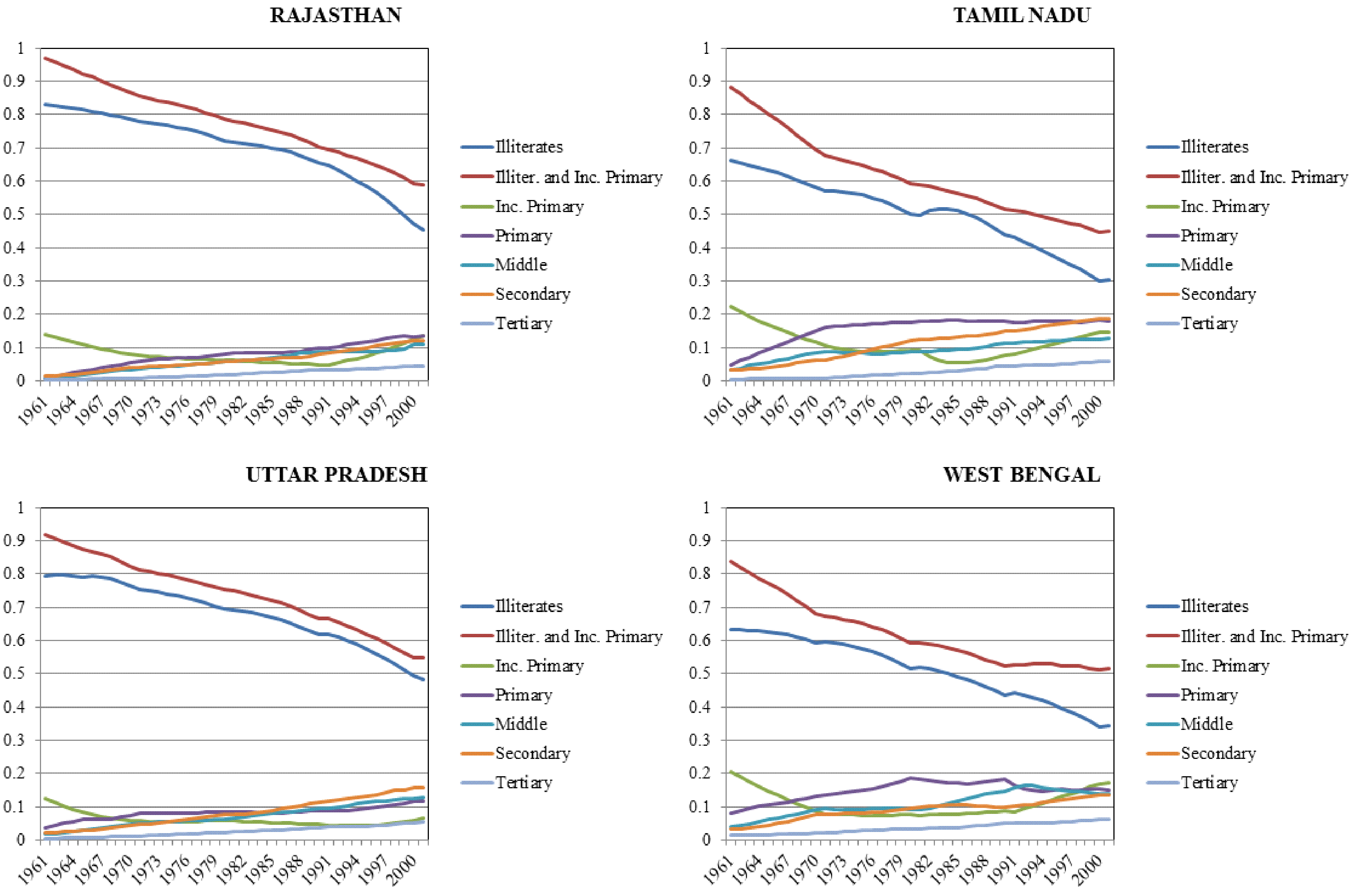


Figure 1- Evolution of attainment levels over time.

Table 1

Dependent variable: Population with tertiary education, $S_{i,t}^{TERT}$

	FE (1)	FE (2)	FE (3)	FE (4)
$\Delta \ln y_{i,t-2}$	0.00037	0.00033		
	(0.00036)	(0.00038)		
$S_{i,t-1}^{TERT}$	1.02446^{a}	1.02317^{a}		
	(0.00211)	(0.00325)		
$lny_{i,t-1}$		0.00008		
		(0.00012)		
$\Delta \ln y_{i,t-10}$			-0.00053	-0.00001
			(0.00147)	(0.00180)
$S_{i,t-10}^{TERT}$			1.22747^{a}	1.21533^{a}
			(0.01295)	(0.02890)
$lny_{i,t-10}$				0.00099
				(0.00208)
Constant	0.00072^{a}	0.00007	0.01041^{a}	0.00229
	(0.00005)	(0.00097)	(0.00065)	(0.01711)
R^2	0.998	0.998	0.972	0.972
Observ.	609	609	481	481

Note: Robust standard errors in parenthesis. a, b, c stand for significance level at 1, 5 and 10 per cent respectively.

Table 2

Dependent variable: Growth rate of per capita income, $\Delta \ln y_{i,t,t+1}$

	Model 1			Model 2
	(1)	(2)	(3)	(4)
$Edu_{i,t}$	0.052^b	0.047^a	0.047^a	
	(0.023)	(0.009)	(0.009)	
$Gini^h_{i,t}$	-0.023			
	(0.301)			
$S^{TOP1}_{i,t}$		-0.372		
		(0.351)		
$S^{TOP10}_{i,t}$			-0.076	
			(0.089)	
$\mathrm{S}^{ILL}_{i,t}$				-0.332^a
				(0.059)
$Edu^{Lit}_{i,t}$				0.026^b
				(0.0110
$Gini^{h^{Lit}}_{i,t}$				0.310
				(0.347)
$\ln y_{i,t}$	-0.379^a	-0.385^a	-0.384^a	-0.369^a
	(0.036)	(0.037)	(0.037)	(0.036)
$Rainfall_{i,t+1}$	0.245^a	0.243^a	0.244^a	0.246^a
	(0.030)	(0.007)	(0.030)	(0.007)
$Ra\inf all^2_{i,t+1}$	-0.045^a	-0.045^a	-0.045^a	-0.045^a
	(0.007)	(0.030)	(0.007)	(0.030)
$TEXP_{i,t}$	-0.125^c	-0.106	-0.113	-0.086
	(0.069)	(0.074)	(0.071)	(0.074)
$DEXP_{i,t}$	0.310^b	0.285^a	0.293^a	0.272^b
	(0.125)	(0.132)	(0.129)	(0.133)
$POP^{RURAL}_{i,t}$	-0.007^a	-0.008^a	-0.008^a	-0.007^a
	(0.002)	(0.003)	(0.003)	(0.002)
$POP^{15+}_{i,t}$	0.008^a	0.009^a	0.009^a	0.007^a
	(0.003)	(0.003)	(0.003)	(0.003)
$Constant$	2.964^a	3.044^a	3.041^a	2.972^a
	(0.400)	(0.323)	(0.328)	(0.326)
R^2	0.340	0.340	0.340	0.330
Observ.	592	592	592	592

Note: Fixed Effects estimation. Robust standard errors in parenthesis. a, b, c stand for significance level at 1, 5 and 10 per cent respectively.

Table 3

Dependent variable: Growth rate of per capita income, $\Delta \ln y_{i,t,t+1}$

	Model 3			Model 4		
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathrm{S}_{i,t}^{TERT}$	4.001^a		4.383^a		4.329^a	3.487^a
	(0.991)		(0.604)		(0.775)	(0.825)
$\mathrm{S}_{i,t}^{SEC}$	0.122	0.962^a				
	(0.229)	(0.145)				
$\mathrm{S}_{i,t}^{PRI}$	-0.010	0.107	-0.012			
	(0.075)	(0.071)	(0.075)			
$\mathrm{S}_{i,t}^{ILL}$				-0.332^a	-0.002	-0.027
				(0.059)	(0.074)	(0.073)
$\mathrm{S}_{i,t}^{ILL} * \mathrm{S}_{i,t}^{TERT}$						1.826
						(1.281)
$\ln y_{i,t}$	-0.415^a	-0.375^a	-0.414^a	-0.369^a	-0.414^a	-0.423^a
	(0.038)	(0.036)	(0.038)	(0.036)	(0.038)	(0.039)
$Rainfall_{i,t+1}$	0.239^a	0.247^a	0.238^a	0.246^a	0.238^a	0.236^a
	(0.030)	(0.029)	(0.007)	(0.030)	(0.030)	(0.007)
$Ra\inf all_{i,t+1}^2$	-0.045^a	-0.046^a	-0.044^a	-0.045^a	-0.044^a	-0.044^a
	(0.007)	(0.007)	(0.030)	(0.007)	(0.007)	(0.030)
$TEXP_{i,t}$	-0.082	-0.135^c	-0.075	-0.086	-0.076	-0.045
	(0.069)	(0.070)	(0.068)	(0.074)	(0.069)	(0.078)
$DEXP_{i,t}$	0.219^c	0.324^b	0.206^c	0.272^b	0.209^c	0.167
	(0.125)	(0.129)	(0.122)	(0.133)	(0.124)	(0.136)
$POP_{i,t}^{RURAL}$	-0.006^b	-0.007^a	-0.006^b	-0.007^a	-0.006^b	-0.007^a
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$POP_{i,t}^{15+}$	0.006^b	0.008^a	0.006^b	0.007^a	0.006^b	0.007^b
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$Constant$	3.291^a	2.925^a	3.290^a	2.972^a	3.292^a	3.392^a
	(0.320)	(0.301)	(0.320)	(0.326)	(0.330)	(0.349)
R^2	0.360	0.340	0.360	0.330	0.360	0.360
$Observ.$	592	592	592	592	592	592

Note: Fixed Effects estimation. Robust standard errors in parenthesis. a, b, c stand for significance level at 1, 5 and 10 per cent respectively.

Table 4

Dependent variable: Growth rate of per capita income, $\Delta \ln y_{i,t,t+1}$

	Model 3				Model 4			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$S^{TERT}_{i,t}$	6.483^a	5.008^a	9.348^b	6.417^a	6.029^a	5.525^a	8.391^b	4.799
	(1.423)	(0.717)	(4.638)	(1.558)	(1.282)	(1.477)	(3.723)	(4.365)
$S^{SEC}_{i,t}$	-0.445		-0.747					
	(0.305)		(1.046)					
$S^{PRI}_{i,t}$	-0.096	-0.069	-0.003	0.021				
	(0.165)	(0.160)	(0.338)	(0.301)				
$S^{ILL}_{i,t}$					0.183	0.132	0.249	-0.014
					(0.167)	(0.172)	(0.485)	(0.391)
$S^{ILL}_{i,t} * S^{TERT}_{i,t}$						0.135		4.802
						(1.532)		(7.993)
$\ln y_{i,t}$	-0.444^a	-0.446^a	-0.501^a	-0.522^a	-0.444^a	-0.440	-0.500^a	-0.484^a
	(0.044)	(0.044)	(0.092)	(0.095)	(0.044)	(0.045)	(0.093)	(0.093)
$Rainfall_{i,t+1}$	0.223^a	0.225^a	0.180^a	0.180^a	0.224^a	0.224^a	0.188^a	0.176^b
	(0.031)	(0.031)	(0.011)	(0.043)	(0.007)	(0.007)	(0.011)	(0.013)
$Ra\inf all^2_{i,t+1}$	-0.041^a	-0.042^a	-0.036^a	-0.035^a	-0.042^a	-0.042^a	-0.038^a	-0.033^a
	(0.008)	(0.008)	(0.041)	(0.011)	(0.007)	(0.031)	(0.041)	(0.044)
$TEXP_{i,t}$	-0.036	-0.061	0.016	-0.031	-0.050	-0.051	-0.018	0.012
	(0.073)	(0.071)	(0.107)	(0.080)	(0.074)	(0.080)	(0.089)	(0.117)
$DEXP_{i,t}$	0.141	0.188	-0.048	0.027	0.161	0.167	-0.010	-0.046
	(0.130)	(0.127)	(0.181)	(0.149)	(0.131)	(0.139)	(0.161)	(0.193)
$POP^{RURAL}_{i,t}$	-0.005^c	-0.005^b	-0.014^b	-0.017^a	-0.005^c	-0.005^c	-0.014^b	-0.017^b
	(0.003)	(0.003)	(0.007)	(0.007)	(0.003)	(0.003)	(0.007)	(0.007)
$POP^{15+}_{i,t}$	0.005	0.006^c	0.016	0.022^b	0.005	0.006	0.018^b	0.020^b
	(0.003)	(0.003)	(0.010)	(0.009)	(0.003)	(0.003)	(0.009)	(0.009)
$Constant$	3.582^a	3.580^a	4.048^a	4.153^a	3.395^a	3.404^a	3.747^a	3.849^a
	(0.372)	(0.372)	(0.806)	(0.806)	(0.404)	(0.427)	(0.829)	(0.820)
R^2	0.335	0.339	0.288	0.326	0.336	0.338	0.296	0.298
$Observ.$	526	526	288	272	526	526	288	288
First Satage								
Instrument	L4 L5 $S^{TERT}_{i,t}$	L4 L5 $S^{TERT}_{i,t}$	L1 $Wage^{TERT}_{URB}$	L2 $Wage^{TERT}_{URB}$	L4 L5 $S^{TERT}_{i,t}$	L4 L5 $S^{TERT}_{i,t}$	L1 $Wage^{TERT}_{URB}$	L1 $Wage^{TERT}_{URB}$
						L4 L5 $S^{TERT}_{i,t} * S^{ILL}_{i,t}$		L1 $Wage^{TERT}_{URB} * S^{ILL}_{i,t}$
Cragg-Donald F-Stat.	308.89	1169.47	32.14	165.01	419.88	236.67	47.21	37.32
Sargan p-value	0.262	0.485			0.344	0.113		

Note: Fixed Effects (FE) and Instrumental Variable (IV) estimation. Robust standard errors in parenthesis. a, b, c stand for significance level at 1, 5 and 10 per cent respectively.

Table 5

Dependent variable: Growth rate of per capita income, $\Delta \ln y_{i,t,t+1}$

	1961-1981			1981-2001		
	(1)	(2)	(3)	(4)	(5)	(6)
			Model 1			
$Edu_{i,t}$	0.126^a	0.060^a	0.064^a	0.152^b	0.092^a	0.098^a
	(0.047)	(0.015)	(0.018)	(0.068)	(0.032)	(0.032)
$Gini^h_{i,t}$	0.558			0.746		
	(0.606)			(0.772)		
$S^{TOP1}_{i,t}$		-1.028^b			0.158	
		(0.458)			(2.094)	
$S^{TOP10}_{i,t}$			-0.205			0.100
			(0.144)			(0.357)
R^2	0.530	0.540	0.530	0.320	0.320	0.320
			Model 2			
$S^{ILL}_{i,t}$	-0.496^a			-0.611^a		
	(0.087)			(0.191)		
$Edu^{Lit}_{i,t}$	0.032^b			0.066^c		
	(0.015)			(0.037)		
$Gini^{h^{Lit}}_{i,t}$	0.351			0.257		
	(0.350)			(1.180)		
R2	0.530			0.320		
			Model 3			
$S^{TERT}_{i,t}$	2.870		5.976^a	3.253^c		4.969^a
	(1.906)		(1.427)	(1.836)		(1.270)
$S^{SEC}_{i,t}$	1.267^b	1.835^a		0.654	1.358^a	
	(0.550)	(0.410)		(0.500)	(0.365)	
$S^{PRI}_{i,t}$	0.183^c	0.247^a	0.172	0.214	0.330	0.203
	(0.105)	(0.083)	(0.107)	(0.317)	(0.328)	(0.307)
R^2	0.540	0.530	0.530	0.340	0.330	0.330
			Model 4			
$S^{ILL}_{i,t}$	-0.442^a	-0.217^b	-0.436^a	-0.720^a	-0.321	-0.260
	(0.084)	(0.109)	(0.121)	(0.173)	(0.289)	(0.276)
$S^{TERT}_{i,t}$		4.956^a	-4.989^c		3.675^b	5.744^b
		(1.621)	(2.536)		(1.802)	(2.240)
$S^{ILL}_{i,t} * S^{TERT}_{i,t}$			16.138^a			-5.364
			(3.680)			(3.897)
R^2	0.517	0.530	0.550	0.307	0.330	0.340
Observ.	320	320	320	288	288	288

Additional controls:

$\ln y_{i,t}$, $Rainfall_{i,t+1}$, $Ra\inf all^2_{i,t+1}$, $TEXP_{i,t}$, $DEXP_{i,t}$, $POP^{RURAL}_{i,t}$, $POP^{15+}_{i,t}$

Note: Fixed Effects (FE) estimation. Robust standard errors in parenthesis. a, b, c stand for significance level at 1, 5 and 10 per cent respectively.

Table 6

Dependent variable: Growth rate of per capita income, $\Delta \ln y_{i,t,t+1}$

	5-year average			**10-year average**		
	(1)	(2)	(3)	(4)	(5)	(6)
				Model 1		
$Edu_{i,t}$	0.029^c	0.017^a	0.021^a	0.003	0.008^b	0.007^c
	(0.015)	(0.006)	(0.007)	(0.012)	(0.003)	(0.004)
$Gini^h_{i,t}$	0.117			-0.069		
	(0.188)			(0.154)		
$S^{TOP1}_{i,t}$		-0.188			0.001	
		(0.166)			(0.093)	
$S^{TOP10}_{i,t}$			0.002			-0.005
			(0.041)			(0.028)
R^2	0.530	0.530	0.530	0.770	0.770	0.770
			Model 2			
$S^{ILL}_{i,t}$	-0.115^a			-0.042^b		
	(0.034)			(0.018)		
$Edu^{Lit}_{i,t}$	0.008			0.002		
	(0.007)			(0.004)		
$Gini^{h^{Lit}}_{i,t}$	0.417^b			0.176		
	(0.196)			(0.131)		
R^2	0.530			0.760		
			Model 3			
$S^{TERT}_{i,t}$	2.334^a		2.004^a	1.833^a		1.008^a
	(0.554)		(0.323)	(0.423)		(0.209)
$S^{SEC}_{i,t}$	-0.107	0.379^a		-0.258^b	0.121	
	(0.145)	(0.094)		(0.119)	(0.088)	
$S^{PRI}_{i,t}$	-0.038	0.023	-0.036	-0.032	0.010	-0.027
	(0.042)	(0.044)	(0.041)	(0.026)	(0.032)	(0.025)
R^2	0.590	0.520	0.580	0.840	0.760	0.810
			Model 4			
$S^{ILL}_{i,t}$		0.043	0.033		0.043^c	0.049^c
		(0.040)	(0.039)		(0.023)	(0.025)
$S^{TERT}_{i,t}$		2.201^a	1.845^a		1.252^a	1.414^a
		(0.391)	(0.491)		(0.259)	(0.351)
$S^{ILL}_{i,t} * S^{TERT}_{i,t}$			0.760			-0.317
			(0.750)			(0.552)
R^2	0.590	0.590	0.590		0.820	0.820
Observ.	124	124	124	61	61	61

Additional controls:

$\ln y_{i,t}$, $Rainfall_{i,t+1}$, $Ra\inf all^2_{i,t+1}$, $TEXP_{i,t}$, $DEXP_{i,t}$, $POP^{RURAL}_{i,t}$, $POP^{15+}_{i,t}$

Note: Fixed effects estimation. Robust standard errors in parenthesis. a, b, c stand for significance level at 1, 5 and 10 per cent respectively.

Table 7

Dependent variable: Growth rate of per capita income, $\Delta \ln y_{i,t,t+1}$

	(1)	(2)	(3)
	Model 1		
$Edu_{i,t}$	0.046^{a}	0.012^{b}	0.016^{a}
	(0.011)	(0.005)	(0.006)
$Gini^{h}_{i,t}$	0.351^{a}		
	(0.117)		
$S^{TOP1}_{i,t}$		-0.120	
		(0.283)	
$S^{TOP10}_{i,t}$			0.020
			(0.066)
	Model 2		
$\mathrm{S}^{ILL}_{i,t}$	-0.124^{a}		
	(0.045)		
$Edu^{Lit}_{i,t}$	0.015^{b}		
	(0.007)		
$Gini^{h^{Lit}}_{i,t}$	0.895^{a}		
	(0.290)		
R^2	0.070	0.060	0.060
	Model 3		
$\mathrm{S}^{TERT}_{i,t}$	1.471^{b}		1.925^{a}
	(0.635)		(0.397)
$\mathrm{S}^{SEC}_{i,t}$	0.170	0.474^{a}	
	(0.173)	(0.109)	
$\mathrm{S}^{PRI}_{i,t}$	-0.063	-0.040	-0.054
	(0.056)	(0.055)	(0.056)
R^2	0.080	0.080	0.080
	Model 4		
$\mathrm{S}^{ILL}_{i,t}$	-0.106^{a}	0.031	0.015
	(0.040)	(0.053)	(0.060)
$\mathrm{S}^{TERT}_{i,t}$		1.971^{a}	1.531^{b}
		(0.500)	(0.667)
$\mathrm{S}^{ILL}_{i,t} {*} \mathrm{S}^{TERT}_{i,t}$			0.797
			(1.011)
R^2	0.062	0.080	0.080
Observ.	592	592	592
Additional controls: $\ln y_{i,t}$, $Rainfall_{i,t+1}$, $Ra\inf all^{2}_{i,t+1}$, $TEXP_{i,t}$, $DEXP_{i,t}$, $POP^{RURAL}_{i,t}$, $POP^{15+}_{i,t}$			

Note: Pooled-OLS estimation. Robust standard errors in parenthesis. a, b, c stand for significance level at 1, 5 and 10 per cent respectively.

Table 8

Dependent variable: Sectoral growth rate, $\Delta \ln y_{i,t}^{SECTOR}$

	Industry				Services			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$S_{i,t}^{TERT}$	-0.419		1.554		5.877^a		5.095^a	
	(1.712)		(1.178)		(1.379)		(0.838)	
$S_{i,t}^{SEC}$	0.654^c	0.558^a			-0.236	0.781^a		
	(0.392)	(0.289)			(0.286)	(0.172)		
$S_{i,t}^{PRI}$	0.125	0.112	0.114		-0.049	0.117	-0.045	
	(0.120)	(0.130)	(0.120)		(0.087)	(0.091)	(0.087)	
$S_{i,t}^{ILL}$				-0.236^c				-0.299^a
				(0.123)				(0.073)
$\ln y_{i,t}^{SECTOR}$	-0.138^a	-0.138^a	-0.131^a	-0.129^a	-0.195^a	-0.132^a	-0.194^a	-0.112^a
	(0.046)	(0.044)	(0.046)	(0.043)	(0.031)	(0.023)	(0.030)	(0.023)
$Rainfall_{i,t+1}$	0.080^c	0.079^c	0.075	0.078	0.107^a	0.117^a	0.108^a	0.116^a
	(0.041)	(0.042)	(0.011)	(0.011)	(0.026)	(0.026)	(0.026)	(0.008)
$Ra\inf all_{i,t+1}^2$	-0.009	-0.009	-0.008^c	-0.008^c	-0.020^b	-0.022^a	-0.020^b	-0.021^a
	(0.011)	(0.011)	(0.042)	(0.042)	(0.008)	(0.008)	(0.008)	(0.027)
$TEXP_{i,t}$	-0.032	-0.026	0.004	-0.014	-0.078	-0.128^b	-0.091^c	-0.105^c
	(0.072)	(0.075)	(0.070)	(0.072)	(0.057)	(0.057)	(0.054)	(0.055)
$DEXP_{i,t}$	-0.086	-0.099	-0.155	-0.106	0.174	0.271^b	0.198^c	0.240^b
	(0.169)	(0.181)	(0.166)	(0.178)	(0.107)	(0.110)	(0.103)	(0.108)
$POP_{i,t}^{RURAL}$	-0.009	-0.009	-0.009	-0.009	0.003	0.000	0.003	0.000
	(0.007)	(0.006)	(0.007)	(0.006)	(0.003)	(0.003)	(0.003)	(0.003)
$POP_{i,t}^{15+}$	0.015^c	0.015^c	0.015^c	0.015^c	-0.004	0.001	-0.003	0.001
	(0.008)	(0.008)	(0.008)	(0.008)	(0.003)	(0.003)	(0.003)	(0.003)
$Constant$	0.740^a	0.748^a	0.715^a	0.929^a	1.302^a	0.801^a	1.289^a	0.958^a
	(0.267)	(0.248)	(0.268)	(0.347)	(0.223)	(0.160)	(0.218)	(0.222)
R^2	0.100	0.100	0.090	0.090	0.190	0.160	0.190	0.150
$Observ.$	583	583	583	583	583	583	583	583

Note: Fixed Effects estimation. Robust standard errors in parenthesis. a, b, c stand for significance level at 1, 5 and 10 per cent respectively.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1085	Holger Breinlich	Heterogeneous Firm-Level Responses to Trade Liberalisation: A Test Using Stock Price Reactions
1084	Andrew B. Bernard J. Bradford Jensen Stephen J. Redding Peter K. Schott	The Empirics of Firm Heterogeneity and International Trade
1083	Elisa Faraglia Albert Marcet Andrew Scott	In Search of a Theory of Debt Management
1082	Holger Breinlich Alejandro Cuñat	A Many-Country Model of Industrialization
1081	Francesca Cornaglia Naomi E. Feldman	Productivity, Wages and Marriage: The Case of Major League Baseball
1080	Nicholas Oulton	The Wealth and Poverty of Nations: True PPPs for 141 Countries
1079	Gary S. Becker Yona Rubinstein	Fear and the Response to Terrorism: An Economic Analysis
1078	Camille Landais Pascal Michaillat Emmanuel Saez	Optimal Unemployment Insurance over the Business Cycle
1077	Klaus Adam Albert Marcet Juan Pablo Nicolini	Stock Market Volatility and Learning
1076	Zsófia L. Bárány	The Minimum Wage and Inequality - The Effects of Education and Technology
1075	Joanne Lindley Stephen Machin	Rising Wage Inequality and Postgraduate Education
1074	Stephen Hansen Michael McMahon	First Impressions Matter: Signalling as a Source of Policy Dynamics
1073	Ferdinand Rauch	Advertising Expenditure and Consumer Prices
1072	Alberto Galasso Mark Schankerman Carlos J. Serrano	Trading and Enforcing Patent Rights

1071	Daniel Hale John Coleman Richard Layard	A Model for the Delivery of Evidence-Based PSHE (Personal Wellbeing) in Secondary Schools
1070	David Marsden	Individual Voice in Employment Relationships: A Comparison Under Different Forms of Workplace Representation
1069	Ross Levine Alexey Levkov Yona Rubinstein	Racial Discrimination and Competition
1068	Klaus Adam Albert Marcet	Internal Rationality, Imperfect Market Knowledge and Asset Prices
1067	Yanhui Wu	A Simple Theory of Managerial Talent, Pay Contracts and Wage Distribution
1066	Yanhui Wu	Managerial Incentives and Compensation in a Global Market
1065	Nicholas Bloom Helena Schweiger John Van Reenen	The Land that Lean Manufacturing Forgot? Management Practices in Transition Countries
1064	Klaus Adam Pei Kuang Albert Marcet	House Price Booms and the Current Account
1063	Stephen Hansen Michael McMahon	How Experts Decide: Identifying Preferences versus Signals from Policy Decisions
1062	Paul Dolan Daniel Fujiwara Robert Metcalfe	A Step Towards Valuing Utility the Marginal and Cardinal Way
1061	Marek Jarocinski Albert Marcet	Autoregressions in Small Samples, Priors about Observables and Initial Conditions
1060	Christos Genakos Kai Uwe Kühn John Van Reenen	Leveraging Monopoly Power by Degrading Interoperability: Theory and Evidence from Computer Markets
1059	Klaus Adam Albert Marcet	Booms and Busts in Asset Prices
1058	Michael W. L. Elsby Jennifer C. Smith Jonathan Wadsworth	The Role of Worker Flows in the Dynamics and Distribution of UK Unemployment

The Centre for Economic Performance Publications Unit
Tel 020 7955 7673 Fax 020 7955 7595
Email info@cep.lse.ac.uk Web site http://cep.lse.ac.uk