EDUCATION IN INDIA: MARKET FAILURES AND POLITICAL CONSIDERATIONS

by

Shira Klien

London School of Economics and Political Science

Thesis submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

This thesis is dedicated to my husband Philip Jonathan Saunders

Abstract

Governments around the world fund schools and are also involved in operating them. There is wide agreement that governments should be involved in provision of education, but the appropriate level and form of their involvement is a subject of debate. The key justifications for government involvement are externalities and credit market imperfections, and this thesis examines these inefficiencies within the context of India's education system.

Chapter 2 assesses human capital externalities in India. It demonstrates that living in a locality with educated individuals has a strong beneficial effect on wages over and above the effect of one's own education. In line with theoretical predictions, the effect is strongest for small geographical areas. In contrast to a general equilibrium interpretation of the results, skilled labour also benefits from a better level of local education. Furthermore, human capital externalities are more pronounced in non-primary industries.

Chapter 3 analyses the effect of credit constraints on education. The principal findings are that credit constraints significantly reduce school attendance and increase wealth inequalities in educational outcomes. Temporary income shocks reduce the probability of attending school, but access to credit mitigates this effect. Finally, the results are not limited to short-term outcomes, but are also seen to be present in long-term outcomes.

Chapter 4 studies how representation of teachers in India's state Upper Houses affects the provision of education. The main results are that teacher representation increases employment of teachers in represented schools and reduces employment in unrepresented schools, with a corresponding effect on educational outcomes. Rather than achieving the intended objectives of teacher representation, teachers seem to have used their political power to shift resources in their favour.

Acknowledgements

I wish to express gratitude to the many people who have made this thesis possible. First of all, I thank Tim Besley and Robin Burgess for their helpful supervision and fine guidance. Their encouragement and advice were essential for bringing this thesis to fruition. I would also like to thank them for providing me with their state-level data of India. Furthermore, I would like to thank Tim for influencing the way I think as an economist, for teaching me the importance of empirical research, and for encouraging me to pursue my research interests. I also wish to thank Robin for his excellent advice on writing methods.

I would like to thank everyone at the STICERD research centre for their friendship, support and discussions, as well as providing me with an excellent work environment. I would particularly like to thank Oriana Badiera, Els Compernolle, Markus Goldstein, Lupin Rahman, Julia Shvets, and Pataporn Sukontamarn for reading earlier drafts and for providing invaluable comments. I extend my warmest thanks to my other friends at LSE: Nazish Afraz, Heski Bar-Isaac, Erlend Berg, Marieke Huysentruyt, Irma Clots-Figueras, Inger Munk, Silvia Pezzini, and many others who helped make working on this thesis much more enjoyable.

This thesis would never been written without the financial support of a number of benefactors. Funding from the AVI fellowship, the C.K. Hobson fellowship, the department of economics at LSE, the ORS studentship, and STICERD, are all gratefully acknowledged.

I am grateful to the National Sample Survey organisation in India for providing me with the NSS data required for the research.

I have presented different parts of this work at various seminars and conferences and I would like to specially thank for their helpful comments the participants at the EOPP seminar, EUDN workshop, IFS conference and the LSE research laboratory workshop. Any residual errors are entirely my own responsibility.

I would also like to thank my family and friends in England and in Israel, whose

love and support kept me sane during difficult moments.

Finally, I would like to thank my dearest husband Phil Saunders, without whom this thesis would not have been completed. I am deeply indebted to his help throughout the years. His love, devotion, and care kept me going. To him I devote this thesis.

"A stable and democratic society is impossible without a minimum degree of literacy and knowledge on the part of most citizens and without widespread acceptance of some common set of values. Education can contribute to both. In consequence, the gain from education of a child accrues not only to the child or to his parents but also to other members of the society. [. . .] Most of us would probably conclude that the gains are sufficiently important to justify some government subsidy."

Milton Friedman, "Capitalism and Freedom"

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Inefficiencies in India's education system are the focus of this thesis. They are examined from three different angles: the benefits to society, household decisions, and the supply side. The first part of the thesis (chapters 2 and 3) examines the existence of market failures, with chapter 2 assessing the existence of human capital externalities in India and chapter 3 analysing the effect of credit market constraints on educational outcomes in India. The second part of the thesis (chapter 4) studies the effect of political power on the supply of education.

1.1 Motivation

The provision of education is a major role of governments around the world. Governments not only provide most of the funding for schools but generally also operate the education sector. Although it is widely agreed that governments should be involved in the provision of education, the level and form of involvement is a subject of debate.

Economic theory views education as an investment decision (Becker (1964)). Individuals bear the direct cost of tuition and other educational inputs, and also the indirect costs of forgoing income during their studies. In return their future income is significantly higher than the income of uneducated individuals. Micro-economic research finds returns of around 6-12% per additional schooling year, albeit varying with time, with state, and with level of technology.[1] Indeed, during the twentieth century education has been a good investment and the world has witnessed a large increase in the share of the educated population across the world.[2]

[1] See Card (1999) for a comprehensive survey of the literature.

[2] For example, in the year 2000 OECD countries had an average secondary school completion rate

However, even if we find clear evidence of large benefits to individuals from schooling, this does not in itself imply the need for large-scale government intervention. Large private returns mean that individuals have strong incentives to acquire schooling without need for government intervention. The standard justifications given for the level of governmental involvement in education are externalities, economies of scale, market failures mainly in the credit market, and redistributive motives. Because of these issues, uncoordinated private decisions are unlikely to lead to optimal social allocation.

Economic theory does not make clear whether externalities are larger in developing countries or developed countries. At low levels of education a single additional year of average schooling can have a dramatic effect on the political awareness of the population or the ability to adopt new technology. In theory however, externalities may only be present in countries above a certain threshold of development, or only at very high levels of education.

Credit constraints are likely to be present in developed countries as well as in developing countries, because human capital is embodied in individuals and cannot act as collateral, so lenders are reluctant to provide loans to finance education. However, given the existing subsidised loans and aids in place in the USA and the UK it seems that credit constraints are unlikely to be a significant barrier to education in those countries. In contrast, households in developing countries are more exposed to risk and official mechanisms to mitigate their risk are limited.

Redistribution motives view education as an important tool in acquiring future income. It is argued that by providing quality education to all citizens regardless of their wealth, the government grants them equal opportunities and enables wealth inequalities to shrink. However, in many countries especially in the developing world, rich students are over-represented in the education system, so education provision can actually be a regressive policy. Under those circumstances a better mechanism for redistribution might be to provide funds directly to poor citizens rather than to fund primarily rich citizens via the education system.

Regardless of the question about the level of government involvement, it is important to understand the governments' incentives for providing education. Political economy literature makes it clear that governments are not always pursuing the best interests of society, with political motivation and institutional limitations sometimes leading to inefficiencies in the government provision of education.

---

of only 45% for 55-64 year olds, compared with an average of 72% for 25-34 year olds. (Hanushek (2002))

This thesis examines the effect of externalities, credit constraints, and political representation on education in a developing country, India. It takes advantage of the unique characteristics of this developing country to study market imperfections. Specifically it examines the existence of human capital externalities where educational outcomes are particularly low and therefore externalities might be of a different magnitude to those in developed countries. In addition, it takes advantage of well-documented credit constraints to study the effect of credit availability on education. It also makes use of the significant state variation in educational outcomes in India to study the effect of teachers' political representation on the provision of education.

This introductory chapter begins by discussing the unique characteristics and distinctive problems of the education sector in developing countries, and then covers the specific features of education in India. Three sub-sections next provide a literature review: the first examines theory and evidence on human capital externalities, the second describes the literature on credit market failures in developing countries, and the third discusses the effect of political institutions on educational outcomes in developing countries. The chapter concludes by describing the layout of the thesis, its main results, and its contribution.

1.2 Background

Education in Developing Countries

Policymakers place immense importance on improving educational outcomes in the developing world. For example two out of the eight millennium goals adopted by the United Nations focus on education. One goal is for all children to attend and complete primary school by 2015, and the second goal is to achieve gender equality at all levels of education by 2005.

In fact, educational achievements have improved dramatically in almost all developing countries since 1960, but despite significant progress toward universal primary education and sharp increases in middle school and secondary school enrolment, there is still substantial room for improvement. In 2000, around 850 million adults (aged 15 or older) in developing countries, amounting to roughly a quarter of the adult population in those countries, were illiterate (UNESCO (2002)).

Educational achievements are highly uneven in developing countries, especially in South Asia. For example, although in 2000 South Asia had a relatively high 4.6 average years of schooling compared to just 3.4 years in Sub-Saharan Africa, it had a lower

literacy rate than Sub-Saharan Africa (Glewwe and Kremer (2005)). This is caused by a small percentage of the population achieving high levels of education while others stay illiterate. Gender gap in educational achievements is also significant in many developing countries with 57% of children out of school being girls (UNESCO (2003)).

Government expenditure per primary school student in developing countries is about 7 percent of per capita GDP, compared with 18.8 percent in developed countries. In contrast, the gap is significantly narrower for spending per secondary school student: 15 percent of per capita GDP for developing countries compared with 21 percent for developed countries. Governments in developing countries spend on average 34 times more on a student in university education than they spend on a student in primary education and 14 times more than on a student in secondary education. The equivalent figures for developed countries are 1.8 and 1.4 (UNESCO (2003)). These figures indicate a very regressive spending structure in developing countries, as the richer children are those who typically attend universities.

Teachers' salaries are by far the largest part of government spending on education, and in some countries these expenditures account for more than 90% of the total cost. The main reason for this is that teachers' salaries in developing countries are very high relative to average salary and GDP per capita. This is due to limited availability of skilled labour, and also political considerations. The low investment in educational inputs mean that sometimes students learn outside in the open air, in many cases students share textbooks, and in some places parents must provide books and other facilities.

Apart from discrepancies in the financing of education, the education systems in developing countries differ in other ways from those in developed countries. In many countries teachers have weak incentives to perform their duties, due to a lack of appropriate supervision and accountability. In some countries teachers come from a higher social status than their students and parents cannot complain about their performance. As a result there can be a high incidence of teacher absence in developing countries.[3]In many countries, teachers coerce parents to pay for "extra lessons" after school or on weekends to prepare students for important examinations. In such situations, increased teaching effort at school could reduce the demand for such extra lessons, and thus reduce teachers' income.

Following recommendations by the World Bank many countries decentralised the

[3]For example Chaudhury etc. (2005) find that 25% of teachers in India were absent from school during an unannounced visit.

delivery of public goods including education. The idea was to strengthen the relationship between beneficiaries and providers of local public goods. It was believed that this arrangement would increase accountability of providers as the community could monitor them more easily, and would also allow each community to adjust public goods to the specific needs of the community. On the other hand, decentralisation can damage weak minorities lacking political representation.

Education in India

The decentralisation and liberalisation processes in India of the last decade brought educational policy in India into the spotlight of public debate. Over the years there has been a substantial mismatch between goals and resource allocation, and although article 45 of the constitution urged the Government of India to provide free and compulsory education up to the age of fourteen by 1960, universal education has not been achieved anywhere in India to date. However, India's success with higher education has masked the enormity of the problem. Many opinions, including those of Drèze and Sen, make a case for a fundamental change in educational policy, and for massive action by the Government of India to achieve economic growth and eliminate poverty. To support his argument Drèze (1995) highlights India's poor educational achievements compared with those of China, who had a similar problem of mass illiteracy and endemic poverty in the late 1940's.

As further support for a change in educational policy, economists emphasise the poor educational achievements from an international perspective. India in the last three decades appeared to be in a difficult situation even compared with regions that are often considered as less developed. For instance female literacy rates were much lower in India than in Sub-Saharan Africa. According to the Indian Census of 1991, the overall literacy rate in India was 50%.[4]Although the Census of 2001 indicates a large improvement in literacy rates, with an overall increase to 68%,[5]India is still ranked as low as 82nd out of 116 countries for which relevant data is available, in terms of proportion of public expenditure on education to GNP.[6]

Although responsibility for school education is shared between the states and the central government, in practice states bear much more responsibility for education than the central government. Therefore, educational achievements in India are highly uneven

[4]Male literacy rate 64%, and female literacy rate just 39%.

[5]For males to 75%, and for females to 54%.

[6]In 1990 the expenditure on education was around 3% of GNP. The percentages were supposed to increase after decentralisation to 6%, however, according to the World Development Report 2000 educational expenditure was still only 3.4% of GNP in 1996.

and there are striking contrasts in literacy levels between different states, different districts, different sectors (urban and rural), those with different social status levels (castes), and genders. For example, female literacy rates vary between 20 percent in Rajasthan and 86 percent in Kerala. Literacy is almost universal in urban Kerala, but particularly low among scheduled caste women in rural Rajasthan. Furthermore, the Gini coefficient for inequality in education is quite high in India with a value of 0.56.[7]Higher education, on the other hand, is a federal matter, with the central government bearing most of the responsibility and most of the financial costs.

India has one of the world's largest gender gaps in education. According to the 1998 Human Development Report only five countries have a greater gender gap than India: Bhutan, Syria, Togo, Malawi and Mozambique.[8]The reasons for this gender gap are partly related to economic reasons such as lower benefits from female education as females have lower participation in paid work, and partly related to the lower social status of females. In fact, Kingdon and Unni (2001) find that education has a U-shaped relationship with paid work participation and that only schooling beyond the junior/middle level enhances their paid work participation. Women who participate in the labour market appear to have higher returns to education than men, but in contrast Kingdon (1998) found that the returns to education for women are lower than those for men. However, this earlier study was based on a relatively small sample of around 200 women and 1000 men from a single city.

Education is divided into pre-primary, primary (children aged 6-11, in classes 1-5), middle or upper primary (children aged 11-14, in classes 6-8), and high school (children aged 14-16 in secondary school, and aged 16-18 in higher secondary school).[9]Higher education includes technical schools, colleges and universities. In most of the villages there is a government primary school (94% of the rural population live within 1 km. of a primary school). Government schools charge negligible fees, and never refuse to enrol a child. There are no Board examinations until class 5, but primary schools can sometimes ask a child to repeat a particular class. The supply problem is more pronounced in middle schools as only 57% of the population live within 1 km. of a middle school.

The low quality of schools mainly in the rural areas has been noted by many authors. A famous example is the PROBE survey, which collected data on primary schools in

[7]World Development Report, 2006 reporting data for 2000.

[8]The 2003 report shows some improvement in India's gender gap, but it remains relatively high compared with other developing countries.

[9]This was formerly one school with all children aged 14-18, in classes 9-12.

randomly selected villages in north India. It finds that only one quarter of the sample schools have at least two teachers, two all-weather classrooms, and some teaching aids. In addition, the report indicates an insufficient number of teachers and classes so that if all children at primary school age were enrolled in a government primary school, there would be more than 100 pupils per classroom on average, and 68 pupils for each teacher. The school facilities are also in poor condition. Only 41% of all primary schools had drinking water and only 40% had non-leaking roofs. Teachers' absence is a serious concern, and at the time of the team visit, around 50% of schools had no teaching activity.

1.3 Literature Review

Human Capital Externalities

There are many theoretical reasons to expect social benefits from education to exceed private benefits. One key channel is through the significant role education plays in determining economic growth. Aghion and Howitt (1988) divide the role of human capital in endogenous growth models into two broad groups. According to the first group, a high level of human capital increases innovation and adoption of new technology (Romer (1990), Nelson and Phelps (1966)). The second group claims that social interaction in the same industry creates learning opportunities which enhance productivity (Lucas (1988)). While the first approach attributes growth to an existing stock of human capital, the second approach emphasises the need for ongoing accumulation of human capital to sustain economic growth.

In both of these groups, externalities are generally built into the production technology without full description of the micro-foundations generating the externalities. But Acemoglu (1996) provides a model which explores the micro-foundation of externalities. His model states that since human capital and physical capital are complementary in production, investment in physical capital depends on the expected level of human capital. Because labour markets are imperfect and workers and firms are not perfectly matched, firms choose to locate themselves in more educated areas. As a result, workers in educated areas have a higher probability of working for a capital-intensive firm and therefore their productivity tends to be higher.

Other theoretical approaches emphasise the role of human capital in affecting non-market outcomes. For example, Lochner (2004) suggests that education increases the opportunity cost of crime, and a more educated society therefore endures lower levels of

crime. Others have suggested that educated citizens can make better-informed political decisions and take greater part in communities and political activities (for example, Milligan, Moretti and Oreopoulos (2004)). Similarly it is believed that a traditional society with more educated women might also benefit from lower fertility rates and better nutrition and health.

The macro-level evidence on the contribution of education to growth is mixed and inconclusive.[10]The main drawback of the literature is that unobserved characteristics could affect both growth and schooling. It is difficult to isolate changes in education from other policies enhancing growth. Moreover the direction of causality is not clear, as a society which expects the economy to expand and generate more opportunities for skilled labour would invest more in education. In addition, measurement errors and differences in the quality of education between countries make it even harder to reach a clear verdict on the contribution of education to economic growth.

Micro-level evidence on the effect of human capital on innovation and technology adoption is more promising. For example, Foster and Rosenzweig (1995) show that the adoption rate of high-yielding seed varieties (HYVs) in India is affected by neighbours' experience. They also discover that farmers with experienced neighbours are significantly more profitable than those with inexperienced neighbours. Thus, the neighbour effect is not just peer pressure or imitation but an actual learning process. Similarly Knight and Weir (2004) find that educated farmers are the first to adopt new technology in rural Ethiopia. They also show that uneducated farmers imitate them and adopt the new technology in a later stage.

The evidence on social non-market outcomes is limited. While there is much evidence that schooling is correlated with better aggregate economic outcomes and with a variety of social outcomes, the magnitude of the effect and the interpretation of the results are unclear. Only a small number of existing papers deal with causality issues and unobservable variables in a convincing way. For example, societies with a lower crime level might attract a better-educated population, rather than the other way around. Alternatively, societies with good opportunities for skilled labour may encourage people to learn instead of performing a criminal activity.

One direct approach to estimate human capital externalities is to focus on estimating the effect of aggregated human capital on a firm's productivity. For example, Knight and Weir (2006) estimate the effect of local average education on farm productivity in rural Ethiopia. They show that average community education significantly

[10]See Krueger and Lindahl (2001) for a review.

increases farm productivity as well as reducing inefficiencies in the production process. An alternative empirical approach toward estimating market externalities is to estimate the effect of local education level on an individual's wage.[11]The basic premise here is that in the absence of externalities local achievements should not affect individuals.

The second chapter of this thesis attempts to assess human capital externalities in India employing this approach. It follows three existing papers: Acemoglu and Angrist (2000), Ciccone and Peri (2000), and Moretti (1998), which are discussed in more detail in that chapter.

Credit Market Imperfections in Developing Countries

Households in developing countries are exposed to significant fluctuations in income due to the important role played by agriculture. Fluctuations in weather and in world commodity prices translate into income shocks. In addition poor sanitation and health services increase the risk of infectious diseases. Market opportunities to deal with these risks are limited due to information problems, absence of collateral, and enforcement difficulties, causing households to engage in alternative methods to smooth their consumption and income.

Extensive literature analyses the level of success such informal mechanisms have in smoothing consumption.[12]The key benchmark of the empirical tests is that with perfect markets transitory income should not have an effect on consumption. Thus the coefficient of transitory income in a regression of household consumption should be close to zero. Alternatively, the coefficient of transitory income in a regression of household saving should be near to one. Similarly, if informal mechanisms to share risk are successful the coefficient of idiosyncratic income shock in a consumption regression should again be close to zero. The general finding of these approaches is that households are reasonably successful in smoothing consumption.

However, this type of analysis does not recognise the cost of strategies employed by households to limit their exposure to risk. This is most often achieved by making conservative but less profitable production decisions and engaging in a more diverse economic activity. In other words, when financial markets are incomplete household production and consumption choices are no longer separable, leading to costly and inefficient adjustments to production and investment decisions. Research focusing on the almost full consumption smoothing achieved by informal mechanisms gives only a very partial picture of the full costs of credit market inefficiencies.

[11]For a full literature review see Moretti (2004).

[12]For a literature review see Besley (1994), Conning and Udry (2005), Morduch (1995).

Substantial literature provides evidence of such costly activity to mitigate risk. For example, Binswanger and Rozenzweig (1993) examine the impact of risk on farm input composition. They show liquid assets and draught animals both being used to smooth consumption. Furthermore, as the environment becomes riskier, as measured by changes in weather conditions, vulnerable households shift production toward more conservative input composition. Analysis by wealth suggests that wealthy households do not significantly change the composition of their inputs, whereas income smoothing costs poor households and households of median wealth 35% and 15% of farm profits, respectively. These results indicate that market imperfection can also result in escalating income inequality.

A few theoretical papers examine the role of imperfect credit markets in the persistence of income discrepancies. Galor and Zeira (1993) show that if an indivisible investment is required for acquiring education, then initial income inequalities will translate to unequal accumulation of human capital and hence to the persistence of wealth inequalities across generations. Banerjee and Newman (1993) generate similar predictions with occupational choice depending on initial wealth when credit markets are imperfect. An important implication of these predictions is that income distribution can have an effect on growth.[13]

Some research directly tests the effect of credit market imperfection on investment in human capital, including health and education. For example, Foster (1995) investigates the fluctuations in child weight after major floods in rural Bangladesh. He shows that in the presence of credit constraints income shocks harm child growth. However landowner households with collateral made use of sufficient credit and did not suffer these effects. Similarly, in investigating child health in South India, Behrman (1988) finds that because households are unable to smooth consumption, the health of children suffers during the season before the main harvest when food is scarce. Furthermore, girls' health suffers more due to gender discrimination. Jacoby and Skoufias's (1997) research, discussed in more details in the main text of the thesis, finds that children in rural South India are often taken out of school in response to adverse income shocks.

The third chapter of the thesis forms part of the research examining the effect of credit constraints on educational outcomes. However unlike other literature I employ a direct measure of credit availability, which allows me to directly estimate the marginal effect of credit availability on educational outcomes.

[13] For a survey of empirical literature testing the effect of inequality on growth, see Banerjee and Duflo (2003).

Political Institutions and Education in Developing Countries

Traditional public economics assumed the objective function of governments to be a weighted average of social welfare. In other words it was assumed that the government acted as a central planner which tried to achieve optimal allocation of resources for the benefits of society. Political economy has shifted the focus of analysis to conflict of interests between various groups in the society and between politicians and the public. Today it is clear that in order to be able to say more about the appropriate form and level of government involvement it is important to focus considerable attention on government incentives. Moreover, it is clear that government incentives are influenced not only by the composition and preferences of the voters but also by the institutional structure.[14]

There is clear evidence that government investment in education in developing countries is motivated to some degree by political considerations rather than mere considerations of efficiency. For example, Pritchett and Filmer (1999) in their meta-literature review show that most papers find that the marginal product per dollar of inputs not directly valued by teachers is commonly 10 to 100 times higher than that of inputs valued by teachers. These results are in contrast with optimisation models that predict equal marginal returns and could be explained by a stronger political influence of teachers on policymakers compared with limited influence of parents. Kingdon and Muzammil (2003) provide a systematic study of teachers' political influence in Uttar Pradesh in India. They find that teachers not only influence the appointment of teachers but also influence legislation on education and the level of supervision over teachers. Reinikka and Svensson (2004) show that corruption is a major issue with financing of education in Uganda, with only 20% of transfers from central government actually reaching schools in 1995.

One of the main political institutions affecting education is the level of centralisation of funds and monitoring. In recent years many developing countries have undergone financial decentralisation in response to the failures of centralised school systems. In addition many communities were given more responsibility in monitoring providers of public goods. Local communities are supposed to have the best knowledge about the needs of their children, strong incentives to monitor the performance of teachers and headmasters, and a comparative advantage in conducting this monitoring.

Evidence for the effectiveness of decentralisation is inconclusive. For example,

[14] A large section of the literature also analyses the efficiency of government schools compared with private schools, however this is beyond the scope of this thesis.

Jimenez and Sawada (1999) and Jimenez and Paqueo (1996) find that enhanced community and parental involvement in schools has improved student skills, diminished student absences, and reduced school costs. Reinikka and Svensson (2003) show that better information enabling the community to monitor local officials increased the percentage of funds from the central government which actually reached the schools from 20 percent in 1995 to 80 percent in 2001.

However, it is not clear that decentralisation and empowerment of local communities is always beneficial. One danger is that in a country with a heterogeneous population, decentralisation will enable strong elite groups to intimidate disadvantaged weaker segments of society and to ignore their needs. A second danger is that the local demand for education is too weak and thus a better ability to monitor the education providers is not sufficient to improve educational outcomes. For example, Kremer and Vermeersch (2005) find no effect of granting school committees with monitoring power on teachers' absence and performance in Kenya. Simlarly, Chaudhury etc. (2005) find little evidence that attempting to strengthen local community ties in India reduces absence.

Miguel and Gugerty (2005) illustrate how ethnic diversity limits the ability to impose social sanctions in diverse communities, resulting in reduced primary school funding and poorer school facilities in western Kenya. Kremer, Moulin and Namunyu (2003) analyse decentralised school responsibility in Kenya allowing local communities freedom to establish new schools. They show that this situation created incentives for local communities to build too many small schools, to spend too much on teachers relative to non-teacher inputs, and to set school fees above those preferred by the median voter with the result that many children are prevented from attending school.

Another important political structure in developing countries is political reservation. Many countries have experimented with political reservation in order to increase minority representation in the political process. Pande (2003) shows in a theoretical model that when candidate entry is mediated by political parties with policy preferences independent of their candidate's identity, changes in legislator identity brought about by political reservation can only affect policy in the absence of full policy commitment. Moreover, such changes may not be significant unless every legislator has a voice in the policymaking process.

The empirical evidence supports the incompleteness of political commitments and shows that politician identity does affect policy outcomes. The main finding of Pande (2003) is that political reservation in Indian states has increased redistribution of re-

sources in favour of scheduled castes and scheduled tribes which benefit from political reservation. She also finds that scheduled caste and scheduled tribe reservation reduces overall resources toward education. Chattopadhyay and Duflo (2004) study the effect of reservation for women in local government in two districts in India and show that women invest more in goods which are relevant to the needs of local women and invest less in education.

The fourth chapter of the thesis examines the effect of a special political institution in India on educational outcomes. Teachers in India have a unique representation in the Indian parliament, and I check how this representation affects educational inputs and outcomes.

1.4 Description of the Thesis

This thesis attempts to assist the debate on government intervention in the provision of education. Its first part examines two important market failures in India which are used to justify government involvement in the educational market. Specifically, it assesses human capital externalities and the influence of credit constraints. Its second part studies political considerations affecting the provision of education in India. This part questions the assumption that governments are motivated purely by efficiency considerations. It shows that political considerations play an important role in the government decision-making process.

The second chapter employs two large representative employment and unemployment surveys in India to assess human capital externalities. Following existing literature regarding the USA,[15]the identification method is based on examining the coefficient of local educational achievements in individuals' wage equations. The key assumption is that in the absence of externalities an individual's wage is affected only by the individual's characteristics. The analysis starts by examining the effect of regional average schooling years on individuals' wages, and then focuses the analysis down to the district level.

The chapter offers two main contributions to the existing literature. One is that it is one of a very few attempts to assess human capital externalities in a developing country where average educational achievements are particularly low.[16]It is not certain that externalities should be larger under these circumstances. However, since there are

[15] Acemoglu and Angrist (2000), Ciccone and Peri (2000), and Moretti (1998).

[16] A discussion of the relevant papers on developing countries is provided under chapter 2 below.

good theoretical reasons to believe so, it is beneficial to test this possibility empirically. Secondly, the chapter presents a theoretical model underlining the differences between general equilibrium models and human externalities models.[17]In doing so it contributes to the discussion in the existing literature regarding the method for assessing human capital externalities.

Special attention is paid to the key identification issues of unobserved variables and endogenous education. The main concern is that regions with better infrastructure would have better educational achievements but also better salaries for reasons other than education. Therefore, fixed effects are included to account for unchanged local characteristics. Control variables are also included to account for unfixed characteristics. A second important concern is that causality might be reversed as wealthier regions have higher demand for education. While potential endogeneity concerns cannot be completely eliminated the chapter provides various robustness checks to mitigate the endogeneity concerns.

A third econometric concern is the ability to separate between a general equilibrium effect and an externalities effect. I develop a simple theoretical model which illustrates the effect human capital externalities have on wages. It highlights the differences between an externalities model and a general equilibrium model and generates testable predictions regarding a human capital externalities effect, in order to distinguish it from a general equilibrium effect. I then test these predictions empirically.

The findings demonstrate a strong effect of local (regional/district) average educational achievements on household wages, even after accounting for the household's own education level and a regional fixed effect. For example, a one-year increase in average regional schooling years raises household wages by 3.8%. These results suggest the existence of human capital externalities. Moreover, in line with theoretical predictions for human capital externalities, I find the above effect to be more pronounced at the district level rather than the regional level. Contrary to the general equilibrium interpretation of the effect, educated workers are not disadvantaged by an increase in the supply of educated workers. In fact, educated workers benefit from living in an educated society. Furthermore, externalities are much more pronounced in non-primary industries, where human capital plays a more significant role in production.

The third chapter employs a large and representative employment and unemployment survey combined with information on bank availability to study the effect of credit constraints on educational outcomes in India. It starts by examining the effect

[17]Independently, a similar model was developed by Moretti (2004b).

of credit on school attendance and incidences of lagging behind at school. Then it makes use of panel data to explore the effect on school attendance of changes in family circumstances. It also tests the extent that credit availability helps to mitigate adverse income shocks. Finally, the chapter considers the effect of bank availability on school completion.

This chapter offers three main contributions to the existing literature. Most importantly, it focuses on identifying the marginal effect of credit constraints instead of merely verifying the existence of these constraints. Secondly, it employs a direct and relatively exogenous measure of credit availability, compared with most research which employs variables that are clearly endogenous, such as access to land or other assets. Thirdly, it focuses on a developing country where there are clear credit constraints and in doing so, contributes to understanding the widespread costs of credit imperfections.

The main finding of the empirical analysis is that the availability of banks significantly increases the probability of attending school and reduces the probability of lagging behind at school. Moreover the availability of banks mitigates wealth inequalities in school outcomes. The panel data analysis reveals that adverse temporary income shocks, such as unemployment, reduces the probability of attending school. However access to credit helps to mitigate the negative effects of these shocks. Finally, the results are not limited to short-term outcomes, such as school attendance, but are also present in actual school completion rates.

The fourth chapter focuses on the supply side and the effect of certain political considerations on the provision of education in India. Specifically, it examines the effect of teachers' political power on education. Teachers in India are granted a special representation in the state Upper Houses, which no other profession enjoys. Middle and secondary school teachers enjoy political representation, but primary school teachers are excluded from the political process. The parliamentary setting varies across time and across the Indian states, with some states having a fixed bicameral parliamentary system, some a fixed unicameral system, and some switching from bicameralism to unicameralism. The chapter makes use of this unique time-state variation to study the effect of teachers' political power within bicameralism on the provision of educational inputs and on educational outputs.

The chapter offers a notable contribution to the political economy literature. Upper Houses are widespread political institutions but little is known about their effect. It is generally very difficult to empirically assess the effect of political institutions, and Upper Houses in particular, because there is very limited country-time variation in their

existence. As a result, identification is based upon cross-country regressions, and in this framework it is hard to account for unobserved variables. In addition, differences in structure and powers of political institutions across countries make it even more difficult to interpret the results. The institutional setting in India provides a unique opportunity to examine the role of Upper Houses.

India is a federal country and educational matters are principally the responsibility of state governments. As a result, educational achievements are highly uneven across India. Specifically, literacy rates and school completion rates vary significantly across the country. By examining some of the political considerations of state governments, this chapter makes a contribution toward understanding the substantial inequalities in education. Moreover, the discrimination of primary school teachers in the political system helps to partially explain the major inequalities in educational resources and outcomes between primary schools and secondary schools.

Political influence of teachers in the parliament is limited by the fact that Upper Houses do not have a power of veto. I develop a simple theoretical model to examine the political effect of Upper Houses with no veto power. The main assumption of the model is that there is political cost associated with disagreement between the Upper House and the government. Therefore the government compromises its preferences and adjusts the budget toward that preferred by the Upper House. This theoretical section predicts that the representation of middle school and secondary school teachers will shift resources toward these teachers.

I test the effect of teacher representation within bicameralism by employing both micro-level regressions and state-level regressions. Accordingly, the data used includes both household data and state panel data for the main sixteen Indian states combined with information on the existence of Upper Houses.

A simple comparison between bicameral states and unicameral states reveals that the allocation of Upper Houses is not random. Larger states, more populated states and richer states were allocated Upper Houses. Therefore a comparison of educational achievements between those types of states is misleading. Consequently, the identification is based on a differences in differences method. I compare between two groups, the first comprising all states with a fixed parliamentary regime, and the second comprising all states which switched from bicameralism to unicameralism. Then I compare the rate of change from before the regime change to after the change.

The assumption of the differences in differences method is that in the absence of change education would have evolved similarly in the two groups of states. I therefore

test this assumption in this chapter. I also discuss the possibility of endogenous change of the parliamentary regime. Finally I examine the option of a general political change affecting both the parliamentary regime and educational policy.

The main finding of the chapter is that teacher representation increases the level of teacher employment in represented schools and reduces employment in unrepresented schools. This shift in resources influences actual educational outcomes. Upper Houses are associated with a 3% increase in middle and secondary school completion rates and a corresponding drop in primary school completion rates. Contrary to the intended objectives of teacher representation, teachers have used their political power to shift resources in their favour.

Overall, the thesis demonstrates strong evidence of market failures affecting the education sector. The existence of externalities implies that individuals under-invest in education. Limited access to credit results in insufficient human capital investment, so households sacrifice potential future benefits from education in favour of short-term benefits from child labour.[18]Taken together, the evidence presented in the first part of the thesis indicates that existing investment in education in India is insufficiently low and that an intervention in the market could be beneficial. However, the second part of the thesis suggests that government intervention should be assessed and implemented carefully. Even though an intervention in the market is valuable it is not clear that the government would achieve optimal allocation of resources, since governments are driven by political considerations and not just by considerations of economic efficiency. In the debate over government role, more attention should be devoted to government incentives.

[18]The economic returns for literacy and for primary school in India are found to be quite low and estimated to be lower than 2%. (See Duraisamy (2002), Kingdon (1998), Kingdon and Unni (2001) and Vasudeva-Dutta (2006)). However, the returns rise with the level of schooling and are estimated to be around 5-10% for each year of secondary school studies and around 15-20% for each year of university studies.

REFERENCES

Acemoglu, Daron, (1996), "A Microfoundation for Social Increasing Returns in Human Capital Accumulation", *The Quarterly Journal of Economics*, 111, 778-804.

Acemoglu, Daron and Joshua Angrist, (2000), "How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws", MIT, Department of Economics, Mimeo.

Aghion, Philippe and Peter Howitt, (1998), "Endogenous Growth Theory", MIT Press, Cambridge MA.

Banerjee, Abhijit V. and Esther Duflo, (2003), "Inequality And Growth: What Can The Data Say?", *Journal of Economic Growth*, 267-299.

Banerjee, Abhijit V. and Andrew F. Newman, (1993), "Occupational Choice and the Process of Development", *Journal of Political Economy*, 101 (2), 274-298.

Becker, G. (1964), "Human Capital", Colombia University Press, New York.

Behrman, Jere, (1988), "Interhousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?", *Oxford Economic Papers*, 40 (1), 32-54.

Besley, Timothy, (1994), "Savings, Credit and Insurance", in Handbook of Development Economics, edited by J. Behrman and T.N. Srinivasan, Vol. 3, Amsterdam, North Holland.

Binswanger, Hans and Mark Rosenzweig, (1993), "Wealth, Weather and the Composition and Profitability of Agriculture Investments", *Economic Journal*, 103, 56-78.

Card, David, (1999), "Education in the Labour Market", in Handbook of Labour Economics, edited by O. Ashenfelter and D. Card, Amsterdam, North Holland.

Chattopadhyay, Raghabendra and Esther Duflo, (2003), "The Impact of Reservation in the Panchayatu Raj: Evidence From A Nationwide Randomized Experiment", MIT ,Mimeo.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan and F. Halsey Rogers, (2005), "Teacher Absence in India: A Snapshot", forthcoming in *Journal of the European Economic Association.*

Ciccone, Antonio and Giovanni Peri, (2000), "Human Capital Externalities in Cities", CEPR Discussion Paper 2599.

Conning, Jonathan and Christopher Udry, (2005), "Rural Financial Markets in Developing Countries", forthcoming in The Handbook of Agricultural Economics, Vol. 3, edited by Evenson, R.E., P. Pingali, and T. P. Schultz.

Drèze, Jean and Jackie Loh, (1995), "Literacy in India and China", Centre of Development Economics, Delhi School of Economics, Working Paper 29.

Duraisamy, P. (2002), "Changes in Returns to Education in India, 1983–94: by Gender, Age-Cohort and Location", *Economics of Education Review*, 21(6), 609–622.

Foster, Andrew, (1995), "Prices, Credit Constraints, and Child Growth in Low-Income Rural Areas", *Economic Journal*, 105(430), 551-570.

Foster, Andrew and Mark R. Rosenzweig, (1995), "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture", *Journal of Political Economy*, 1176-1209.

Galor, Oded and Joseph Zeira, (1993), "Income Distribution and Macroeconomics", *Review of Economic Studies*, 60 (1), 35-52.

Glewwe, Paul and Michael Kremer, (2005), "Schools, Teachers, and Education Outcomes in Developing Countries", forthcoming in Handbook on the Economics of Education, edited by Eric A. Hanushek.

Hanushek, Eric A. (2002), "Publicly Provided Education", in Handbook of Public Economics, Vol. 4, edited by A.J. Auerbach and M. Feldstein, Amsterdam, North Holland.

Human Development Report, (1998),

http://hdr.undp.org/reports/view_reports.cfm

Human Development Report, (2002),

http://hdr.undp.org/reports/view_reports.cfm

Jacoby, Hanan and Emanuel Skofias, (1997), "Risk Financial Markets and Human Capital in a Developing Country", *Review of economic Studies*, 64(3), 311-35

Jimenez, Emmanuel, and Vicente Paqueo, (1996), "Do Local Contributions Affect the Efficiency of Public Primary Schools?", *Economics of Education Review*, 15(4), 377-386.

Jimenez, Emmanuel and Yasuyuki Sawada, (1999), "Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program", *World Bank Economic Review*, 13(3), 415-441.

Kingdon, Geeta G. (1998), "Does the Labour Market Explain Lower Female Schooling in India?", *Journal of Development Studies*, 35(1), 39-65.

Kingdon, Geeta G. and Mohd. Muzammil, (2003), "The Political Economy of Education in India: Teacher Politics in Uttar Pradesh", Oxford University Press, New Delhi.

Kingdon, Geeta G. and Jeemol Unni, (2001), "Education and Women's Labour Market Outcomes in India", *Education Economics*, 9(2), 173-195.

Knight, John and Sharada Weir, (2004), "Externality Effects of Education: Dynamics of the Adoption and Diffusion of an Innovation in Rural Ethiopia", *Economic Development and Cultural Change*, 53 (1), 93-113.

Knight, John and Sharada Weir, (2006), "Production externalities of education: Evidence from rural Ethiopia", forthcoming in *Journal of African Economies.*

Kremer, Michael, Sylvie Moulin, and Robert Namunyu, (2003) "Decentralization: A Cautionary Tale", Work in Progress, Harvard University.

Kremer, Michael and Christel Vermeesh, (2005), "School Committee Empowerment: Preliminary Notes", Harvard University, Mimeo.

Krueger, Alan B. and Mikael Lindahl, (2001), "Education for Growth: Why and For Whom?", *Journal of Economic Literature*, 39, 1101-1136.

Lochner, Lance, (2004), "Education, Work, and Crime: A Human Capital Approach", *International Economic Review*, 45 (3), 811-843.

Lucas, Robert E. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22(1), 3-42.

Miguel, Edward and Mary Kay Gugerty, (2005), "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya", *Journal of Public Economics*, 89 (11-12), 2325-2368.

Milligan, Kevin, Enrico Moretti and Philip Oreopoulos, (2004), "Does Education Improve Citizenship? Evidence from the U.S. and the U.K.", *Journal of Public Economics*, 88, 1667-1695.

Morduch, Jonathan, (1995), "Income Smoothing and Consumption Smoothing", *Journal of Economic Perspectives*, 9 (3), 103-114.

Moretti, Enrico, (1998), "Social Returns to Education and Human Capital Externalities: Evidence from Cities", UC Berkley Department of Economics, Mimeo.

Moretti, Enrico, (2004), "Human Capital Externalities in Cities", in Handbook of Regional and Urban Economics, Amsterdam, North Holland.

Moretti, Enrico, (2004b), "Estimating the Social Return to Higher Education: Evidence From Longitudinal and Repeated Cross-Sectional Data", *Journal of Econometrics,* 121(1-2).

Nelson, Richard R. and Edmund S. Phelps, (1966), "Investment in Humans, Technological Diffusion, and Economic Growth", *American Economic Review,* 61, 69-75.

Pande, Rohini, (2003), "Can Mandated Political Representation Provide Disadvantaged Minorities Policy Influence? Theory and Evidence from India", *American Economic Review,* 93 (4), 1132-1151.

Pritchett, Lant and Deon Filmer, (1999), "What Education Production Functions Really Show: A Positive Theory of Education Expenditures", *Economics of Education Review,* 18, 223-239.

The PROBE Team, (1999), Public Report on Basic Education in India, Oxford University Press, New Delhi.

Reinikka, Ritva and Jakob Svensson, (2003), "The Power of Information: Evidence from a Campaign to Reduce Capture", World Bank, Mimeo.

Reinikka, Ritva and Jakob Svensson, (2004), "Local Capture: Evidence from a Central Government Transfer Program in Uganda", *Quarterly Journal of Economics,* 119(2), 679-705.

Romer, Paul M. (1990), "Endogenous Technological Change", *Journal of Political Economy,* 98(5), 71-102.

UNESCO, (2002), "Education for All: Is the World On Track?", UNESCO Publishing. http://portal.unesco.org/education/en/ev.php

UNESCO, (2003), "Gender and Education for All: The Leap to Equality", UNESCO Publishing. http://portal.unesco.org/education/en/ev.php

Vasudeva-Dutta, P. (2006), "The structure of wages in India, 1983-1999", forthcoming in *Education Economics.*

World Development Report, (1999/2000), "Entering the 21st Century", http://econ.worldbank.org

World Development Report, (2006), "Equity and Development", http://econ.worldbank.org

Chapter 2

Assessing Human Capital Externalities in India

2.1 Introduction

Human capital externalities have significant implications for both endogenous growth models and educational policies. In particular, externalities are crucial to one of the key questions regarding educational policies: the question of whether governments should subsidise education or not. The case for subsidising human capital on a wide scale rests on the existence of a significant divergence between private and social benefits. If private benefits from education are much lower than social benefits, then individuals under-accumulate human capital. This under-accumulation generates the possibility of a classic market failure in which, on efficiency grounds, there is room for government intervention. But in the absence of this or other market failures, government provision for education would for the most part only result in reducing private investment.

Many economists believe that human capital externalities play an important role in explaining the substantial variation in growth rates existing between various countries. Accordingly, growth models emphasise the central role of human capital externalities in enhancing technological innovation and technological diffusion, which result in faster economic growth (see for example, Acemoglu and Angrist (2000)).[1]

[1]In contrast, there are important papers that view factor accumulation as the main source of differences between countries. Young (1995) in his influential paper uses growth accounting to show that much of East Asian growth can be attributed to factor accumulation. Mankiw Romer and Weil (1992) claim that the Solow model with human capital explains most of the differences between countries.

There are substantial theoretical reasons to believe in the existence of externalities. One approach stresses that countries with proportionally well-educated populations generate greater innovation and technological diffusion due to efficient exchange of ideas.[2] A second approach is based on imperfect matching between firms and workers, causing firms to invest more physical capital in localities where the expected human capital is higher. This results in an increase in productivity of all types of workers.[3] A third approach emphasises the non-market benefits of education, such as reducing crime, increasing voter participation, and reducing birth rates.[4]

This chapter takes an important step at estimating human capital externalities in developing countries and offers two central contributions. First, there is very limited literature estimating human capital externalities outside of the USA and, in particular, in developing countries. Moreover, this chapter offers much more convincing methods to deal with econometric concerns over unobserved variables and endogeneity than the methods employed by existing literature on developing countries. Developing countries provide a better prospect for assessing human capital externalities because of the low levels of literacy and schooling in many of these countries. Empirical evidence shows that private benefits from education are decreasing, i.e. the financial benefits from primary education are much higher than the financial benefits from secondary school or graduate studies.[5] If social returns from education also diminish with increasing average education, we can expect externalities to matter more in developing countries.[6]

The second contribution of this chapter is that it helps to resolve the controversy in existing literature regarding methods for assessing human capital externalities. Ciccone and Peri (2000) showed that if workers with different levels of human capital were complementary in production then an increase in the supply of educated workers increased the marginal productivity of uneducated workers and thus raised their wages, even in the absence of externalities. Hence, in contrast to the conventional methodology of identifying externalities, a positive effect of average education level on workers' wages cannot necessarily be interpreted as evidence of externalities. This chapter presents

[2] Aghion and Howit (1998), Lucas (1988), Romer (1990), and Shumpeter (1934).

[3] Acemoglu (1996).

[4] See chapter 1 for a full discussion.

[5] Krueger and Lindahl (2001).

[6] It is not obvious that social benefits and private benefits behave in the same way. It appears that having a few brilliant researchers can contribute more to a society than an egalitarian increase of an additional year of schooling (Murphy, Shleifer and Vishny (1991)). Diminishing social benefits from education are more of an empirical question. However, a comparison of the benefits to society of first degrees, and the benefits to society from an increase in literacy rates, is likely to be in favour of the latter. Therefore it is well worth pursuing this possibility empirically.

a simple empirical test to distinguish between externalities and a general equilibrium effect. The central idea is to show that an increase in the supply of educated workers has diverse effects on different types of workers. While Ciccone and Peri focused on the positive effect of an increase in the supply of educated workers on uneducated workers' wages, I show that in the absence of externalities the effect on educated workers' wages is negative.

India provides an ideal setting given low educational achievements, significant cross-state variation, and high quality data. Using two NSS (National Sample Survey) rounds of employment and unemployment surveys from India, this chapter examines the effect of average local education level on household expenditure, used as a proxy for workers' wages. It tests the effect at the district level and also at the larger regional level, adjusting for household characteristics including level of education. In addition, this chapter empirically examines the question of what affects externalities, and which types of people benefit from living in an educated society. By assessing the effect of average education on different types of workers in different industries, we can get a better understanding of the mechanism underlying externalities.

The main finding of my empirical analysis is that average education in any locality (district or region) has a strong positive effect on household expenditure over and above the household's own education, even after accounting for the regional fixed effect and after applying different measures of average education. The magnitude of the externalities is approximately half as large as that of private returns. For example, a one-year increase in average regional schooling raises household expenditure by 3.8%, while the private benefit from an additional schooling year is 7.8%.

The rest of the chapter is organised as follows: section 2 discusses the previous literature; section 3 outlines a theoretical framework and highlights the differences between two alternative interpretations of the empirical estimation; section 4 reviews the economic strategy; section 5 presents the results; section 6 concludes.

2.2 Literature Review

In spite of the important implications toward educational policies and growth models, most of the empirical microeconomic literature focuses on the benefits to individuals. Using data on individuals' education and income, the existing research provides strong evidence of a substantial payoff from investment in education, especially in primary schools (see Card (1999) or Krueger and Lindahl (2001) for a review). However, from

this microeconomic evidence it is not clear whether the aggregate benefits to society from schooling exceed the private benefits. Discussing of the effect of education in the macroeconomic literature does not provide any better evidence (see Venniker (2001) or Krueger and Lindahl (2001) for a review). Considering the difficulties in obtaining accurate country-level data on changes in educational attainments, it is not surprising that the cross-country literature generates inconclusive evidence on aggregate educational effects. I take an alternative approach, using cross-state data within a large country.

The empirical literature on human capital externalities is quite limited. Earlier attempts to measure human capital externalities, using data on cities, suffered from serious econometric problems, such as the effect of unobserved variables and the issue of endogenous city education levels, mainly due to selective inter-city migration (e.g. Rauch (1993)). However, three important attempts using data on cities in the USA were made to measure human capital externalities and deal with concerns over endogenous variables and unobserved variables. These papers used a repeated cross-section which allowed them to introduce city fixed effects and time fixed effects, in order to reduce concerns over unobserved city characteristics.

However, these papers employed very different estimation methods and so it is not surprising that they yielded mixed results. Angrist and Acemoglu (2000) included city average schooling years in the individual wage equation as their measure for externalities. To solve the concerns over endogenous variables, they employed compulsory schooling laws as instrument variables for average education. However, they did not discuss the existence of an alternative interpretation of their results as a general equilibrium effect, and the identified externalities of around 1% were not statistically significant.

Moretti (1998) applied a different method for estimating human capital externalities. In the first stage he computed the specific representative wage of group j in city c at time t, after accounting for observed individuals' characteristics including education level. In the second stage he estimated the effect of city average education on the average representative wage, accounting for the city's ethnic structure. As an instrument variable for city average education in 1980 and 1990 he used the city's demographic structure in 1970 and 1980 respectively. Moretti found significant externalities of around 2%. However, his claim that the city demographic structure does not affect wages directly is debatable. In addition, like Angrist and Acemoglu, he did not discuss the existence of an alternative interpretation of his results as a general

equilibrium effect.[7]

Ciccone and Peri (2000) applied three different methods for identifying externalities. To eliminate the general equilibrium interpretation of the results they kept the composition of workers fixed. Like Moretti they computed the representative wage of group j in city c at time t, but when computing the average wage they kept the weights assigned to each group fixed at the level they were at the earliest data point. They found no evidence of externalities.

Literature on human capital externalities in developing countries is limited and inconclusive. Various papers have estimated the effect of education on adoption of new technologies. For example, Knight and Weir (2004) examined the role of education in the adoption and diffusion of fertiliser in rural Ethiopia. They found that at the early stages of technology adoption educated people were more likely to adopt the technology. In later stages they were imitated by other non-educated farmers and education became less important in determining the usage of the technology. While this research approach is interesting and demonstrates the existence of a learning process in agriculture, it is hard to generalise the results as they are quite limited to a specific area and to a very specific context.

A more comprehensive approach is to estimate the effect of average education on productivity or on wages. Unlike similar literature focusing on the USA, literature focusing on externalities in developing countries did not deal convincingly with econometric issues of unobserved variables, endogeneity and general equilibrium effect.[8]One of the main problems is that none of the papers employed a panel dataset, which would have allowed it to account for unobserved local characteristics.

Knight and Weir (2006) estimated the effect of average community education on productivity in rural Ethiopia. They found the effect of average education on productivity to be positive, significant and large. However, once village variables were included the effect was reduced to a third of its original size. Moreover, the lack of panel data did not allow them to include community fixed effects. Consequently a third unobserved factor such as agricultural or economic conditions could have been partially responsible for the results by causing farmers to invest more in both education

[7]Subsequent to completion of this chapter Moretti (2004) has now provided a similar theoretical model to distinguish between externalities and general equilibrium. He also employed individual fixed effect and dealt in a more convincing way with migration issues. In this later version he used an additional instrument variable: the presence of land-granted colleges. The magnitude of the results is slightly reduced.

[8]Duflo's paper (2000) while dealing convincingly with econometric concerns touch on the subject, but did not focus on human capital externalities.

and productivity-enhancing farm inputs.

Employing data from the Kenyan welfare monitoring survey Kimenyi, Manda and Mwabu (2004) estimated human capital externalities in Kenya. Following the literature on the USA they introduced local average years of education into Mincer's wage equation. They found that district-level average education for males and females had a positive and statistically-significant effect on earnings for all workers in urban areas. However, they made only a very limited attempt to account for unobserved district characteristics or for endogeneity. They included teacher-pupil ratio as a proxy for school quality and province fixed effects, which account for unobserved province differences but not for district characteristics.

Michaud and Vencatachellum (2003) estimated the effect of average village education of the four main racial groups on wages in South Africa. They found that the impact of human capital externalities on wages differs by race. More specifically, they found that the average human capital of blacks in South Africa had a positive effect on the wages of white workers, whereas the aggregate human capital of whites had a negative impact on the wages of black workers, so the demand for black workers fell when there were more skilled white workers. As in the papers discussed above, their data was not a panel dataset and therefore their ability to account for unobserved characteristics was limited.

Kochar (2003) estimated separately the effect of average education of the poor and of the rich on farm profitability in India. Employing government expenditure on schooling as an IV he found that schooling of the poor had a negative effect on farm profits because it increased the wages paid for farm workers. In contrast schooling of the rich had a positive effect on profitability. These results are in line with Ciccone and Peri (2000) and the general equilibrium model, and could easily be explained by labour supply and demand, so they do not imply negative or positive externalities. Moreover, like other papers discussed above the authors did not employ a panel dataset and therefore had limited ability to account for unobserved variables.

2.3 Theoretical Framework

This section presents a theoretical model which illustrates how human capital externalities affect wages and expenditures. It also provides a useful organisational device for discussing the empirical assessment of human capital externalities. The aim is to provide foundations for the empirical tests and a structured scheme for interpreting the

results. This section also outlines a unifying framework in the subsequent discussion of the literature. In particular the model examines criticism raised by Ciccone and Peri (2000) regarding previous identification methods, and underlines the differences between predictions generated by general equilibrium models and those generated by human capital externalities models.

However, this theoretical section does not attempt to explain what generates externalities or identify their micro-foundations. Thus the notion that an exchange of ideas can increase productivity is modelled by allowing the total factor productivity (TFP) to depend on average education.

Micro-economic theories, including that of Mincer (1974), predict no effect of average education on individuals' wages. Therefore the common empirical strategy to identify externalities has been to estimate the effect of average education on wages. However, Ciccone and Peri (2000) show that average education affects wages if we apply a general equilibrium framework and assume different levels of human capital to be complementary in production. Therefore interpreting the effect of average education as evidence of human capital externalities is potentially unsafe. In the following model I demonstrate the effect of average education on wages both in the presence and in the absence of externalities.

Let the economy consist of one production sector[9] and two types of agents, who can supply either 1 unit of raw labour or h_i units of human capital. These two groups are distinguished by their different levels of disutility from studies $\frac{1}{\delta_i}$.[10] In the first group, of size L, each individual generates very high disutility from studies $\frac{1}{\delta_1}$, and therefore chooses to supply only 1 unit of raw labour.[11] In the second group, of size $(x-1)L$,[12] each individual generates a lower disutility from studies $\frac{1}{\delta_2}$ and therefore chooses to study and supply h_i units of human capital according to his utility:

$$v(c_i, h_i) = c_i - \frac{h_i^{(1+\tau)}}{\delta_2(1+\tau)} \tag{2.1}$$

where c_i is consumption and τ is a parameter measuring the convexity of the costs

[9] An extension of the model to a two-sector economy may be found in Appendix 2.7.

[10] Alternatively, one can interpret the distribution of δ_i as different abilities or distribution of wealth, when there are credit constraints.

[11] The optimal level of h_1 is too small, therefore workers choose to supply raw labour. For simplicity, I assume that the labour supply is inelastic.

[12] $x > 1$. The proportion of educated people is therefore $\frac{x-1}{x}$ or $1 - \frac{1}{x}$, which is positively depended on x.

function.

For purposes of simplicity, I assume that there are only two factors of production: human capital (h) and raw labour (l). Consistent with the argument by Ciccone and Peri, I assume complementary production factors[13]and more specifically, a Cobb-Douglas production function. Thus, the production function of firm ij, with h_i educated workers and l_j uneducated workers, is:

$$y_{ij}(l_j, h_i) = A l_j^{\alpha} h_i^{(1-\alpha)} \tag{2.2}$$

where $A = A_0 \overline{H}^{\gamma}$is the TFP with γ being a parameter measuring externalities. Externalities can be either positive or negative and in the absence of externalities $\gamma = 0$. $\overline{H}$ is the average education level and therefore can be rewritten as the aggregate education divided by the overall population size, $\overline{H} = \frac{H}{xL}$. Thus the demand functions of firm ij for workers are:

$$w(l) = \frac{\partial y_{ij}(l_j, h_i)}{\partial l_j} = \alpha A_0 \overline{H}^{\gamma} l_j^{(\alpha-1)} h_i^{(1-\alpha)} \tag{2.3}$$

$$w(h) = \frac{\partial y_{ij}(l_j, h_i)}{\partial h_i} = (1-\alpha) A_0 \overline{H}^{\gamma} l_j^{\alpha} h_i^{-\alpha} \tag{2.4}$$

Dividing equation (2.3) by equation (2.4), yields the following result:

$$\frac{h_i}{l_j} = \frac{(1-\alpha) w(l)}{\alpha w(h)} \tag{2.5}$$

Hence, all firms have the same ratio of educated labour to raw labour $\frac{H}{L}$, and this ratio depends negatively on the wage ratio $\frac{w(h)}{w(l)}$.

The aggregate supply of raw labour is the number of uneducated people, L. The supply of educated labour depends on the optimisation decision of individuals from the second group. Assuming that there is no saving, we can substitute c_i by $w_i = w(h)h_i$, and solve $\frac{\partial v(c_i, h_i)}{\partial h_i} = 0$ for h_i, to find that the supply of individual i from group H is:

$$\frac{h_i^{\tau}}{\delta_2} = w(h) \tag{2.6}$$

Hence, all individuals with low disutility of studying $\frac{1}{\delta_2}$ choose the same level of edu-

[13]Ciccone and Peri (2000) show that the identification problems arise only when workers are imperfect substitutes. If workers are perfect substitutes then the effect of average education on individuals' wages can be attributed uniquely to externalities and the identification problem is eliminated.

cation h_2. Solving for equilibrium generates the following results:[14]

$$w(l)^* = \alpha A_0 \overline{H}^{(\gamma+1-\alpha)} x^{(1-\alpha)} = \alpha A_0 h_2^{(\gamma+1-\alpha)} x^{-\gamma} (x-1)^{(\gamma+1-\alpha)} \tag{2.7}$$

$$w(h)^* = (1-\alpha) A_0 \overline{H}^{(\gamma-\alpha)} x^{-\alpha} = (1-\alpha) A_0 h_2^{(\gamma-\alpha)} x^{-\gamma} (x-1)^{(\gamma-\alpha)} \tag{2.8}$$

Equations (2.7) and (2.8) characterise equilibrium. To determine the effect of an increase in human capital on wages and to correspond to the empirical section we can analyse two cases: first, the case of an increase in the average human capital ($\overline{H}$) regardless of the reason for such increase, and second the case of an increase in the proportion of educated workers in the population (x).

Proposition 1 ***The Effect of Human Capital on Wages***

- *If there are no externalities ($\gamma = 0$) then the wages of uneducated workers increase if there are more educated people ($\Delta x > 0$) or if average education increases ($\Delta h_2 > 0$ or $\Delta \overline{H} > 0$). The wages of educated workers will however fall.*
- *If externalities are sufficiently large ($\alpha < \gamma < \alpha x$), then wages of both educated and uneducated workers increase when educated workers accumulate more human capital.*
- *If externalities are even larger ($\gamma > \alpha x$), then wages of both groups of workers increase when educated workers accumulate more human capital and/or a larger proportion of the population becomes educated.*

Proof.

- $\frac{\partial w(l)^*}{\partial \overline{H}} = (\gamma + 1 - \alpha)\alpha A_0 \overline{H}^{(\gamma-\alpha)} x^{(1-\alpha)}$, so $\frac{\partial w(l)^*}{\partial \overline{H}} > 0$ if and only if $\gamma + 1 - \alpha > 0$. In particular, if $\gamma = 0$ then $\frac{\partial w(l)^*}{\partial \overline{H}}$ must be positive because $\alpha < 1$.
- $\frac{\partial w(l)^*}{\partial x} = (\gamma + x - \alpha x)\, \alpha A_0 h_2^{(\gamma+1-\alpha)} x^{(-\gamma-1)} (x-1)^{(\gamma-\alpha)}$, so $\frac{\partial w(l)^*}{\partial x} > 0$ if and only if $(\gamma + x - \alpha x) > 0$. In particular, if $\gamma = 0$ then $\frac{\partial w(l)^*}{\partial x}$ must be positive as $\alpha < 1$ and $x > 0$.
- $\frac{\partial w(h)^*}{\partial \overline{H}} = (\gamma - \alpha)(1-\alpha) A_0 \overline{H}^{(\gamma-\alpha-1)} x^{(-\alpha)}$, so $\frac{\partial w(h)^*}{\partial \overline{H}} > 0$ if and only if $\gamma > \alpha$. In particular, if $\gamma = 0$ then $\frac{\partial w(h)^*}{\partial \overline{H}}$ must be negative because $\alpha > 0$.

[14] Note that $\frac{h_i}{l_j} = \frac{H}{L} = \overline{H}X = h_2(x-1)$

- $\frac{\partial w(h)^*}{\partial x} = (1-\alpha)\, A_0 h_2^{(\gamma-\alpha)} x^{(-\gamma-1)}(x-1)^{(\gamma-\alpha-1)}(-\gamma(x-1)+x(\gamma-\alpha))$, so $\frac{\partial w(h)^*}{\partial x} > 0$ if and only if $\gamma > \alpha x$. In particular, if $\gamma = 0$ the sign of $\frac{\partial w(h)^*}{\partial x}$ must be negative, as $\alpha > 0$ and $x > 0$.

■

In order to assess human capital externalities, I focus first on the arguments of Ciccone and Peri. The results in the absence of externalities ($\gamma = 0$), specifically equations (2.3) and (2.7) and part 1 of proposition 1, indicate that an increase in the supply of educated labour, either in the form of a greater proportion of the population being educated or in the form of each individual being more educated, will increase the marginal productivity of raw labour (MPL). Therefore, unless the supply of workers is fully elastic, an increase in the supply of educated workers will also raise the uneducated workers' wages. The essential point is that even in the absence of externalities an increase in average education raises some of the workers' wages, so it is impossible to assess human capital externalities simply by looking at the coefficient of average education in the wage equation. Therefore, Ciccone and Peri challenge the common estimation method of externalities, and claim that it is impossible to rule out that the underlying reason for the average education effect may be purely a general equilibrium mechanism.

However, turning to the effect of human capital on educated workers' wages yields very different results. While the prediction of a general equilibrium effect is that wages of uneducated workers rise as a result of an increase in the supply of educated workers, the prediction for wages of educated workers is now a fall instead of a rise. Only if externalities are sufficiently large ($\gamma > \alpha$ or even $\gamma > x\alpha$) will we observe a positive effect of average education on educated workers' wages.

These results generate significant implications for the identification strategy. To be more precise, in order to assess human capital externalities it is important to look separately at the effect of average education on educated workers. If the empirical analysis shows that educated workers' wages depend positively on average education, then the general equilibrium effect alone cannot explain the results and hence the existence of human capital externalities can be inferred.

2.4 Data and Estimation Method

The chapter aims to identify the effect of local educational achievements on the productivity of individuals. Subsequent to Ciccone and Peri (2000) challenging previous identification methods, considerable attention has been devoted in this section to discussing the differing predictions of the externalities model and the general equilibrium model.

My assessment of human capital externalities is based on two representative household employment and unemployment surveys of 365 districts[15]within 74 different regions of India, conducted in 1987-1988 (round 43) and 1993-1994 (round 50), by the National Sample Survey (NSS) organisation. Each round covers approximately 40,000 urban households and 80,000 rural households, and contains various details about their socio-economic status, educational attainments, and demographic composition. I have focused on households where the head of the household is of working age, i.e. between 20 and 64.

Table 2.1 contains a limited list of descriptive statistics categorised by round and sector. For all education levels the achievements are seen to be significantly better in urban areas. For instance, while literacy rates are below 50% in rural areas they are above 75% in urban areas. Over time there is a small improvement in all educational attainments, especially literacy rates and primary school completion rates. Examination of less aggregated measures of educational attainments in regions reveals substantial variation across regions at any specific point in time, and also substantial variation in the change of educational attainments over time. For example, the literacy rates vary significantly between regions, from 95-97% in urban regions in Karnataka, Kerala and some Union territories, to 20-30% in rural areas of Rajasthan. Table 2.2 presents average expenditure by education level. In general, expenditure increases with education and over time there is an increase in expenditure in both rural and urban areas.

The main identification strategy is based on the absence of local educational attainments in the individual's wage equation according to Mincer (1974). In other words, according to standard theory an individual's wage is determined only by the individual's own characteristics including their level of education, but not by average education level in the relevant geographical area. Therefore, a key part of the identification strategy is to add local average education into the household's wage equation.

[15]Some regions of Jammu and Kashmir were not surveyed.

Thus the basic estimation equation is:

$$\ln(E_{ilt}) = \beta_0 + \beta x_{ilt} + \lambda h_{ilt} + \delta \overline{H}_{lt} + \gamma_0 S_{lt} + R + T + U + \epsilon_{ilt} \tag{2.9}$$

which makes use of three kinds of variables. The first type is the dependent variable $\ln(E_{ilt})$ being the log expenditure per capita of household i in locality l at time t. The second type $\overline{H}_{lt}$, is the educational achievement in locality l at time t. The third type are control variables x_{ilt}, h_{ilt}, and S_{lt}. In addition each regression includes regional fixed effect R,[16]time fixed effect T, and sector dummy U (urban/rural).

The dependent variable $\ln(E_{ilt})$ is a function of expenditure rather than wage. While both the theoretical model and previous empirical literature discuss human capital externalities in terms of education affecting wages, in developing countries a great share of income is generated by home production, which is by and large agricultural activity. Moreover, since a considerable part of the wage is provided in goods and kinds instead of cash, it becomes difficult to estimate correctly the value of the income. Furthermore, self-employment in small businesses makes it even harder to generate a reliable measure of income. Therefore assuming that human capital externalities affect all parts of income, including home production,[17]it seems more appropriate to use expenditure data instead of wage figures, even if saving is not a constant share of income.

The main variable of interest is local educational attainment $\overline{H}_{lt}$, but the data allows me to measure educational achievements only in terms of education levels rather than schooling years. Therefore I employ three broad types of educational attainment measures. The first type is the proportion of people in a certain geographical area who achieved a specific level of educational attainment, for example, the regional literacy rate, or the regional rate of primary school completion. In the theory section I note it as $(\frac{x}{x-1})$, which is positively depended on x. The second type is the average education level, where each level of educational attainment receives a ranked value.[18]In the theory section I note it as $\overline{H}$. The third type is a proxy for average schooling years, which in the theory section I note as h_i. I assign the number of schooling years to individuals based on the number of years required to complete each educational level under the education

[16]Regression size limitation made it impossible to include district fixed effect.

[17]I discuss the empirical effect of home production in Appendix 2.8.

[18]1 for illiterate, 2 for literate, 3 for completing primary school, 4 for completing middle school, 5 for completing secondary school, and 6 for graduate. The exact formula is detailed in the Data Appendix.

system in India. To be precise, I assign 5 years of schooling to individuals completing primary school, 8 years of schooling to individuals completing middle school, 12 years of schooling to individuals completing secondary school, and 15 years to individuals completing higher education. I also assign one year of studies for literate individuals.

The measurement of the latter is inaccurate as it does not fully take into account schooling years of people who dropped out of school midway through a particular education level. For example, a person who studied for ten years appears in my data as someone who completed middle school but not secondary school and therefore accounts only for eight years of primary and middle school. However, I find it useful to employ this measure as it is easier to interpret the coefficient on average schooling years than to interpret the coefficients on the other two measures of educational achievements.

The relevant geographical unit is of two types: regions and districts. Each state in India comprises a number of regions, between one and seven depending on the state's size, with each region covering several districts.

The control variables can be classified into two distinct categories. The first type is a vector of household characteristics x_{ilt}, including social group, religion, age,[19]number of males/females working, size of the household,[20]and a vector of household education h_{ilt}.[21]For the household education levels I have calculated the number of people at each household with each level of education (illiterate, literate, primary school, middle school, secondary school, graduate).[22]The second type of control variable is a vector of state variables S_{lt}, including GDP per capita, length of roads, population and taxes. To obtain these state control variables I employed a second source of data, the Indian states panel dataset, which includes annual economic and political variables for the sixteen major Indian states between 1958 and 1994.[23]

Following previous papers, in the first phase I identify externalities by examining δ, the coefficient of local average education. This identification is based on the assumption that in the absence of externalities and accounting for the individual's own education and other characteristics, this effect is expected to be zero. However, the theoretical

[19]Square age was not significant.

[20]The regressions generate similar results if instead of including the size of the household and the number of household members who work, they include the size of the household and the square size of the household.

[21]The group omitted is the number of household members who completed primary school. Where appropriate the regressions include schooling years instead.

[22]The results are robust when using dummy variables for the education level of the head of the household instead of the education levels of all of the household members.

[23]The dataset is described in Besley and Burgess (2000).

discussion, based on Ciccone and Peri (2000), demonstrated that even in the absence of externalities an increase in the supply of educated workers in the city raises the wages of uneducated workers. Therefore, if we employ the regression only for uneducated workers, then the coefficient on average education, δ, is equal to $\gamma + 1 - \alpha$. This is a positive expression even in the absence of externalities, as $1 - \alpha > 0$. Employing the regression for all types of workers, δ becomes a weighted average of $\gamma + 1 - \alpha$ and $\gamma - \alpha$. Consequently at least part of δ in equation (2.9) is due to a general equilibrium effect rather than externalities.

However the theoretical section showed that in the absence of externalities an increase in either the supply of educated workers or the average schooling years affects educated workers negatively, so only in the presence of sufficiently large externalities can the coefficient for educated people be positive. Consequently, an important part of the empirical strategy is based on analysing the expenditures equation of the educated workers, where finding a positive effect on educated workers can be interpreted as evidence for the existence of human capital externalities. In view of that, I have introduced education interaction terms. This helps me to examine empirically whether externalities affect educated workers differently from uneducated workers, or if all workers are equally affected, and in particular whether educated workers are positively affected by local educational attainments.

The amended estimation equation is as follows:

$$\ln(E_{ilt}) = \beta_0 + \beta x_{ilt} + \lambda h_{ilt} + \delta \overline{H}_{lt} + \varphi \overline{H}_{lt} C_{ilt} + \gamma_0 S_{lt} + R + T + U + \epsilon_{ilt} \qquad (2.10)$$

where all the definitions are the same as before, and $\overline{H}_{lt} C_{ilt}$ is the education interaction term. $C_{ilt} = 1$ if the household is educated and zero otherwise. For robustness checks I have performed regressions varying the level of education for a household to be deemed educated. Specifically, I employed three main definitions for educated households. Under the first definition a household was classified as educated if the head of the household was literate. Under the second definition a household was classified as educated if the head of the household had completed primary school. Under the third definition a household was classified as educated if the head of the household had completed middle school. To test the hypothesis that educated households are disadvantaged by living in an educated society, I have tested whether $\varphi + \delta \leqslant 0$. Observing $\varphi > 0$ implies that educated workers' gains are higher than those of uneducated work-

ers. Yet even if φ is negative, which implies that educated workers' benefits are lower than uneducated workers' benefits, then as long as $\varphi + \delta$ is positive we can interpret the results as evidence of externalities.

In general, the main econometric concerns with the above approach are unobserved variables which could affect both regional average education and household expenditure, and potential endogeneity of average education level.

Unobserved variables could cause people living in regions with higher levels of education to have higher income due to other reasons than average education in the region; for example better bureaucracy in the region can improve both the functioning of markets and the education system, leading to an increase in both the education levels in the region and individuals' earnings. Another example could be credit constraints: limited borrowing possibilities can affect both investment in physical capital and investment in human capital, therefore creating a correlation between productivity and human capital.

In order to minimise the issue of unobserved variables I have used a repeated cross-section to introduce a regional fixed effect and time fixed effect. However, using the regional fixed effect solves only the problem of unobserved fixed characteristics of the region but not the problem of changes in unobserved variables. If average education is a proxy for local GDP or if changes in unobserved local characteristics, such as growth rates, are correlated to the changes in the levels of education, it can result in a correlation between the error term and the dependent variable. Therefore I have added to the regressions measures of GDP (per capita state deflated income), infrastructure (length of state highways), rural and urban population, and proxies for government actions (state taxes, state expenditures).

Endogeneity can be an issue if educated workers move to richer regions. Therefore, I have examined migration trends employing a third source of data, the Indian district database created by Maryland University which combines district-level data from census and agricultural sources between 1961 and 1991. It turns out that migration rates are small and fairly constant over time, implying that the latter concern can be dealt with by the regional fixed effect so long as the educational achievements of the immigrants are fixed over time.

While the above argument does mitigate the problem, endogeneity remains a concern and I still need to perform robustness checks to provide greater confidence in the results. Therefore to further mitigate concerns over endogeneity and unobserved variables I have introduced industry interaction terms. The key principle is that finding

a significant difference in the coefficients on local educational achievements between industries makes it less likely that household expenditure have increased the local educational achievements. To claim for such reverse causality one would have to explain why current household expenditure of workers in one industry has affected the local educational achievements while the expenditure of workers in another industry does not have the same effect.

In addition to the general concerns regarding unobserved variables and endogeneity, a specific concern exists with regard to interpretation of the education interaction term. If being educated is correlated with other variables it can result in the education interaction term not measuring the gap in externalities between educated workers and uneducated workers. In particular, if educated people work in separate industries, then the interaction term will measure the gap in externalities between different types of industries. Thus to separate the two effects of industry and education, I have added two more interaction terms into the following regression:

$$\ln(E_{ilt}) = \beta_0 + \beta x_{ilt} + \lambda h_{ilt} + \delta \overline{H}_{lt} + \varphi \overline{H}_{lt} C_{ilt} + \psi \overline{H}_{lt} I_{ilt} + \theta \overline{H}_{lt} C_{ilt} I_{ilt} + \gamma_0 S_{lt} + R + T + U + \epsilon_{ilt} \quad (2.11)$$

in which all of the definitions are as in equations (2.9) and (2.10), apart from the additional non-primary industry interaction term $\overline{H}_{lt} I_{ilt}$, and the educated people in non-primary industry interaction term $\overline{H}_{lt} C_{ilt} I_{ilt}$, which captures the difference in the gap between educated and uneducated workers in the two different industries.

Introducing local averages into micro-regressions raises yet another econometric concern. As Moulton (1986, 1990) shows, if the error terms are correlated within the regions then the standard errors from OLS estimation can be biased. Therefore all the regressions are clustered by year and by locality (region or district), and the standard errors are corrected accordingly.

2.5 Results

2.5.1 Base Results

As explained in the methodology section, I have measured externalities as the effect on household expenditure of local average education, over and above the household's

own education level, with equation (2.9) providing the basic estimation.

The upper part of Table 2.3 reports basic results from my data, with the dependent variable in all regressions being the natural logarithm of household expenditure per capita. Throughout my analysis I account for other factors affecting household expenditure by using state and time fixed effects and some household characteristics including religion, social group, main industry, age, size of household, number of workers in the household, and most importantly household level of education.

In column 1 of table 2.3 local educational achievements are represented by regional average schooling years, and a positive and significant association between local educational achievements and household expenditure is clearly seen. Columns 2-6 confirm that these results are not sensitive to the construction of schooling years from grouped information on education.

Each of the columns contains a different measure of regional educational achievements. For example in column 2 regional educational achievements are measured in terms of education level completed, while in column 3 regional educational achievements are measured in terms of regional literacy rates. While the measures of literacy rates and education level are more precise given the limitations of the data, they are more difficult to interpret and therefore later on I focus on the measure of schooling years.

The magnitude of the effect is quite large, half as large as the size of private returns. In column 1 an increase of one year in the regional average schooling is associated with an increase of household expenditure by 3.8%, while an increase in household schooling years is associated with 7.8%. This means that living in a region where average schooling is two years higher is as good for the household members as obtaining an additional schooling year themselves. Column 2 reveals very similar results: a one unit increase in education level, which is equivalent on average to approximately three schooling years, raises household expenditure by 12.4%.

The lower part of the table presents the effect of local educational achievements when the reference area is a district rather than a region. The results turn out to be much more pronounced when they refer to a smaller local area, so the effect of local educational achievements is seen to diminish with size. Theories on human capital externalities suggest such a pattern, where the importance of local educational achievements decreases with the size of area. For example the probability of exchanging ideas for improvement of productivity or innovation falls with increasing distance between people. Similarly, applying the imperfect matching theory of Acemoglu, companies

making investment decisions take into account the expected level of human capital in their local labour market rather than the human capital in the whole country.

Taken together, the results demonstrate a consistent picture. Local educational achievements are associated with higher expenditures, even after taking into account the private educational achievements. The magnitude of the effect is impressive and counts for as much as half of the private benefits from education, and in line with theory the externalities are even more pronounced in smaller geographical units.

2.5.2 Robustness Checks

While these results show a clear and positive correlation between local educational achievements and household expenditure, three significant issues remain outstanding. First, they make no effort to separate the externalities from the general equilibrium effect. Secondly, other unobserved local characteristics could affect both local educational achievements and household expenditure. Finally, local educational achievements could be endogenous and responding to household expenditure. In the next sub-section I address these concerns.

A significant concern regarding the interpretation of equation (2.9) arises from the complementarity in production between educated workers and uneducated workers. When different types of workers are complementary the results cannot be automatically interpreted as evidence for human capital externalities, as they could be due to a general equilibrium effect. However, a suitable way to distinguish between the two interpretations is to assess the effect of average education on educated workers. If general equilibrium were the only driving force behind the results above, we would expect to find a negative supply effect on educated workers' wages. Therefore, as described in equation (2.10) I have added an education interaction term[24]to test whether an educated environment disadvantages educated workers.

The results are reported in table 2.4. In order to simplify the interpretation of the local educational coefficient, local educational achievements are measured in terms of local average schooling years in all of the regressions. In the first part of the table the local area refers to regions, and in the second half it refers to districts. For purposes of robustness I employ two different definitions of educated workers. Therefore, in columns 1, 2, 5 and 6 educated households are defined as households with the head being literate, while in the rest of the table they are defined as households with the

[24]Set to 1 if the household is above a specific level of education.

head completing at least primary school. As the table shows the interaction terms are negative and significant. The results are robust with respect to a change in the local area from region to district, to inclusion of additional control variables, and to different definitions of educated workers. The benefits for educated people from an increase in the supply of educated people are smaller than those of uneducated people as a result of the downward slopping demand function for educated workers. Nonetheless they are seen to be positively affected, implying that educational externalities are present and large enough to fully offset the complementary effect between educated workers and uneducated workers.

Table 2.5 and the even columns of table 2.4 incorporate an additional set of specific time-state control variables. All regressions now include GDP per capita, population size, taxes, and length of roads. Together these variables help to control for alternative factors which could theoretically affect both educational achievements and household expenditure. The table shows that the results are not sensitive to the inclusion of state controls.

Tables 2.6, 2.7 and 2.8 report the results with a non-primary industry interaction term. In all of the tables local educational achievements are measured as local average schooling years. Table 2.6 includes the industry interaction term on its own, while table 2.7 combines it with the two other interaction terms described in equation (2.11): the education interaction term, and the educated in non-primary industry interaction term.[25]The last interaction term captures the different size of gap between educated and uneducated workers in different industries.

The non-primary industry interaction terms are between 1% and 2% and significant, revealing that externalities are much more evident in non-primary industries where human capital plays a more significant role. The theoretical reasons for expecting greater externalities in non-primary industries are directly linked to the micro-foundations of externalities. If externalities are a result of an exchange of ideas, we will observe greater externalities where ideas contribute more to production. If externalities are a result of capital investment by firms anticipating the local level of human capital, as in the Acemoglu model, then we will observe greater externalities where investment decisions are more dependent upon human capital. Appendix 2.7 extends the theoretical model to allow for two types of industries with greater externalities in the non-primary sector.

Table 2.8 shows explicitly the sum of the relevant coefficients from table 2.7. For example, in part B of table 2.8 educated workers are defined as primary school gradu-

[25]Set to 1 for educated household if its main industry is not mining or agriculture.

ates. Primary school graduates benefit less than workers who did not complete primary school. The difference between the columns is also significant, implying that externalities exist more significantly in non-primary industries.

Furthermore, these results make it less likely that the driving force behind the estimation is solely unobserved variables or endogeneity, as otherwise we should expect to find no difference between primary and non-primary sectors. It is also important to note that these additional results strengthen the previous ones, and that living in a more educated society does not disadvantage educated workers as the sum of the coefficients of educated workers is positive.

Comparing my results with existing ones, there are four key distinctions in variable definitions and datasets. The first is that I compose a measure of schooling years out of grouped educational data rather than having a direct measure of schooling years. The second difference is that I employ expenditure data rather than wage data. The third difference is that the geographical unit of my analysis is either a region or district rather than a city. Finally, I analyse at the household level rather than the individual. However, all of these differences are fairly inconsequential. The most important distinction between my analysis and previous papers is its focus on a developing country where educational achievements are particularly low, and therefore local educational achievements are more likely to affect household expenditure. Indeed the magnitude of my results turns out to be considerably greater than those found in the previous papers, detailed in section 2.2, which focused on the USA.

To conclude, these results support the existence of externalities and suggest that externalities are greater in non-primary industries where human capital plays a more important role in production. Secondly, my focus on a developing country with particularly low educational achievements generates more significant results than those found in previous papers.

Finally, the chapter estimates human capital externalities to be between 3% and 5%, varying with the industry and education of individuals. This is a large effect, especially when compared with the very low levels of private returns to primary school estimated in previous papers.[26]The magnitude of the effect gives rise to important policy implications. It shows that there are large benefits to education in India. However, since the perceived private returns to education are quite low, especially at the primary school level, and at the literacy level, individuals under-invest in education. To help

[26]The estimated average private returns to a year of schooling in India is around 7%, and the private returns to a year of primary school are only around 1%-2%.

individuals internalise the effect of their education the government of India should try to reduce the direct and indirect cost of schooling and to improve the private returns from schooling.

2.6 Conclusion

It has long been argued that education can generate important consequences by contributing to the economic growth and elimination of poverty in developing countries. It was considered that private benefits from education would be of a different scale in developing countries; however Duflo (2001) established in the Indonesian case that private benefits from primary school in developing countries are of the same magnitude as those in developing countries. Therefore, the argument for distinctive treatment of educational attainments in developing countries leans mainly on either credit constraints or a significant divergence between the social benefits from education and the private benefits.

However, although human capital externalities have an important implication on both policy decisions and growth theory, most of the empirical literature is focused on private returns. Since there are also positive social returns from education, private returns underestimate the total economic value of schooling. The few attempts to measure human capital externalities have been concentrated mainly on the USA, but less developed countries provide an interesting case for measuring human capital externalities due to their low levels of education. India is characterised by particularly low levels of educational achievements even compared to other developing countries, and also by a large variation in education levels across different regions. These two characteristics combine with low constant migration, very high quality data, and large samples, to make India an excellent place for trying to reach a clearer conclusion regarding the existence of externalities.

The main contribution of the chapter is its attempt to provide rigorous estimates of human capital externalities in developing countries and to deal convincingly with concerns over unobserved variables and endogeneity. Apart from an attempt to get more conclusive evidence on externalities, this chapter takes the existing literature a step further. It examines what affects human capital externalities, and who benefits from an educated society. It also provides an answer to criticism by Ciccone and Peri of measures of externalities. By analysing the different predictions of a general equilibrium model and of an externalities model, the chapter provides an empirical test

to distinguish between the two models.

The results show a strong effect of average local education on household expenditure even after accounting for the household's own education level and other household characteristics. The effect is robust to different measures of average education. More importantly, the effect is also significant when regional fixed effects are included.

The combined evidence suggests that an educated environment positively affects individuals' achievements. Educated people also benefit from living in an educated society, and externalities are seen to be greater in industries where human capital plays a more important role in production. Furthermore, externalities are more pronounced when the locality is smaller, so educational achievements at the district level are more significant than at the larger regional level. The magnitude of the effect suggests that individuals under-invest in education. Therefore the government of India should help individuals to internalise the full benefits of education by reducing the direct and indirect cost of schooling. However, in order to generate policy recommendations there are a few questions that still need to be addressed. In particular, more detailed research is required toward establishing the level of education which is most important in generating significant social benefits. Finally, future research is also required to measure the effect of average education on non-market outcomes, such as voting rates, crime, and birth rates.

2.7 Appendix: Model For Two-Sector Economy

I assume here that the economy consists of two production sectors: primary industries and non-primary industries. I will indicate the sector as subscript $m \in [1, 2]$, where $m = 1$ is the sector of primary industries, and $m = 2$ is the sector of non-primary industries. The main difference between the two sectors is in the role of human capital. In the non-primary sector, the human capital externalities are greater. Similar to Acemoglu (1996), I assume some frictions in the labour market, and that workers have a certain probability p of being employed in the non-primary sector.[27] These assumptions have three implications. The first is that different wages may exist between the two sectors. The second implication is that the size of each sector is exogenous, and the third is that the ratio between human capital and raw labour is the same in both sectors.

[27] For simplicity, I assign the same probability of being in the non-primary sector to an educated worker and to a non-educated worker.

All assumptions regarding factors of production are as before. The production function of firm ij in sector m, with h_i educated workers and l_j uneducated workers, is:

$$y_{ijm}(l_j, h_i) = A_m l_j^{\alpha} h_i^{(1-\alpha)} \tag{2.12}$$

where $A_m = A_0 \overline{H}^{\gamma_m}$ is the TFP, with γ_m being a parameter measuring externalities in sector m. Externalities are greater in the non-primary sector, i.e. $\gamma_2 > \gamma_1$. $\overline{H}$ is the average education and therefore can be written also as $\overline{H} = \frac{H}{xL}$. Accordingly the demand functions of firm ij in sector m for workers are:

$$w_m(l) = \frac{\partial y_{ijm}(l_j, h_i)}{\partial l_j} = \alpha A_0 \overline{H}^{\gamma_m} l_j^{(\alpha-1)} h_i^{(1-\alpha)} \tag{2.13}$$

$$w_m(h) = \frac{\partial y_{ijm}(l_j, h_i)}{\partial h_i} = (1-\alpha) A_0 \overline{H}^{\gamma_m} l_j^{\alpha} h_i^{-\alpha} \tag{2.14}$$

Dividing equation (2.13) by equation (2.14) yields the following result:

$$\frac{h_{im}}{l_{jm}} = \frac{(1-\alpha) w_m(l)}{\alpha w_m(h)} \tag{2.15}$$

Hence, all firms in sector m have the same ratio of educated labour to raw labour $[\frac{H}{L}]_m$, and this ratio depends negatively on the wage ratio $\frac{w_m(h)}{w_m(l)}$.

The aggregate supply of raw labour is the number of uneducated workers, L. Nevertheless, the supply of educated labour depends on the optimisation decision of individuals from the second group. Substituting c_i by w_i and maximising the expected utility $p[w_2(h)h_i - \frac{h_i^{(1+\tau)}}{\delta_2(1+\tau)}] + (1-p)[w_1(h)h_i - \frac{h_i^{(1+\tau)}}{\delta_2(1+\tau)}]$, the supply h_i of individual i from the second group is:

$$\frac{h_i^{\tau}}{\delta_2} = (1-p) w_1(h) + p w_2(h) \tag{2.16}$$

Hence, all individuals with low cost of studying $\frac{1}{\delta_2}$ choose the same level of education h_2. Solving for equilibrium generates the following results:

$$w_m(l)^* = \alpha A_0 \overline{H}^{(\gamma_m+1-\alpha)} x^{(1-\alpha)} = \alpha A_0 h_2^{(\gamma_m+1-\alpha)} x^{-\gamma_m} (x-1)^{(\gamma_m+1-\alpha)} \tag{2.17}$$

$$w_m(h)^* = (1-\alpha)A_0\overline{H}^{(\gamma_m-\alpha)}x^{-\alpha} = (1-\alpha)A_0 h_2^{(\gamma_m-\alpha)}x^{-\gamma_m}(x-1)^{(\gamma_m-\alpha)} \qquad (2.18)$$

It is easy to see that the model yields very similar results to those in the main theoretical section above, with one additional insight that the effect of average education on wages is greater in the non-primary sector:

Proposition 2 ***Comparison of the Effect of Average Education***

For each type of worker, the effect of average education is greater in the non-primary sector if externalities in the non-primary sector are greater than corresponding externalities in the primary sector.

Proof. 1. $\frac{\partial w_2(h)^*}{\partial h_2} - \frac{\partial w_1(h)^*}{\partial h_2} > 0$ if
$(\gamma_2-\alpha)h_2^{(\gamma_2-\alpha-1)}x^{(-\gamma_2)}(x-1)^{(\gamma_2-\alpha)} - (\gamma_1-\alpha)h_2^{(\gamma_1-\alpha-1)}x^{(-\gamma_1)}(x-1)^{(\gamma_1-\alpha)} > 0.$
Since $(\gamma_2-\alpha) > (\gamma_1-\alpha)$ it is enough to show that
$h_2^{(\gamma_2-\alpha-1)}x^{(-\gamma_2)}(x-1)^{(\gamma_2-\alpha)} - h_2^{(\gamma_1-\alpha-1)}x^{(-\gamma_1)}(x-1)^{(\gamma_1-\alpha)} \geqslant 0,$
which is true as long as $w_2(h)^* \geqslant w_1(h)^*$, and this is true as long as $\gamma_2 > \gamma_1$.

2. $\frac{\partial w_2(h)^*}{\partial x} - \frac{\partial w_1(h)^*}{\partial x} > 0$ if
$h_2^{(\gamma_2-\alpha)}x^{(-\gamma_2-1)}(x-1)^{(\gamma_2-\alpha-1)}(\gamma_2-\alpha x) - h_2^{(\gamma_1-\alpha)}x^{(-\gamma_1-1)}(x-1)^{(\gamma_1-\alpha-1)}(\gamma_1-\alpha x) > 0.$
Since $(\gamma_2-\alpha x) > (\gamma_1-\alpha x)$ it is enough to show that
$h_2^{(\gamma_2-\alpha)}x^{(-\gamma_2-1)}(x-1)^{(\gamma_2-\alpha-1)} - h_2^{(\gamma_1-\alpha)}x^{(-\gamma_1-1)}(x-1)^{(\gamma_1-\alpha-1)} \geqslant 0,$
which is true as long as $w_2(h)^* \geqslant w_1(h)^*$, and this is true as long as $\gamma_2 > \gamma_1$.

3. $\frac{\partial w_2(l)^*}{\partial h_2} - \frac{\partial w_1(l)^*}{\partial h_2} > 0$ if
$(\gamma_2+1-\alpha)x^{(-\gamma_2)}\,(x-1)^{(\gamma_2+1-\alpha)}\,h_2^{(\gamma_2-\alpha)} - (\gamma_1+1-\alpha)x^{(-\gamma_1)}\,(x-1)^{(\gamma_1+1-\alpha)}\,h_2^{(\gamma_1-\alpha)} > 0.$
Since $(\gamma_2+1-\alpha) > (\gamma_1+1-\alpha)$ it is enough to show that
$x^{(-\gamma_2)}\,(x-1)^{(\gamma_2+1-\alpha)}\,h_2^{(\gamma_2-\alpha)} - x^{(-\gamma_1)}\,(x-1)^{(\gamma_1+1-\alpha)}\,h_2^{(\gamma_1-\alpha)} > 0,$
which is true as long as $w_2(l)^* \geqslant w_1(l)^*$, and this is true as long as $\gamma_2 > \gamma_1$.

4. $\frac{\partial w_2(l)^*}{\partial x} - \frac{\partial w_1(l)^*}{\partial x} > 0$ if $(\gamma_2+x-\alpha x)\,h_2^{(\gamma_2+1-\alpha)}x^{(-\gamma_2-1)}\,(x-1)^{(\gamma_2-\alpha)}$
$-(\gamma_1+x-\alpha x)\,h_2^{(\gamma_1+1-\alpha)}x^{-\gamma_1-1}\,(x-1)^{(\gamma_1-\alpha)} > 0.$
Since $(\gamma_2+x-\alpha x) > (\gamma_1+x-\alpha x)$ it is enough to show that
$h_2^{(\gamma_2+1-\alpha)}x^{(-\gamma_2-1)}\,(x-1)^{(\gamma_2-\alpha)} - h_2^{(\gamma_1+1-\alpha)}x^{(-\gamma_1-1)}\,(x-1)^{(\gamma_1-\alpha)} > 0,$
which is true as long as $w_2(l)^* \geqslant w_1(l)^*$, and this is true as long as $\gamma_2 > \gamma_1$. ■

2.8 Appendix: Expenditures versus Wages

This appendix describes in a simple framework the effect of employing expenditure data rather than wage data for the estimation of human capital externalities in developing countries. As I mention in the main text, although the theory and previous empirical literature discuss human capital externalities in terms of education affecting wages, in developing countries a great share of income is generated by home production, which is by and large an agricultural activity. In theory, human capital externalities might affect home agricultural production as well as wages. Moreover, since a considerable part of wages in developing countries is provided in goods and kinds instead of cash it becomes more difficult to estimate reliably the monetary value of the income, and consequently employing wage data reduces the dataset significantly.

The following model illustrates how human capital externalities in home production might affect the estimation. I assume that household i in locality l has two sources of income: home production f_{il} and money wage, w_{il}. In addition it is assumed that there is no saving and therefore the household's aggregate income $w_{il}+f_{il}$ is equal to its expenditure, EX_{il}. Both wage and home production are functions of the household's characteristics as well as regional educational attainments, and can be described by the following equations:

$$w_{il} = w(z_i) + \rho \overline{H}_l \tag{2.19}$$

$$f_{il} = f(z_i, L) + \sigma \overline{H}_l \tag{2.20}$$

where z_i is a vector of household characteristics and L is the size of its land. ρ measures the externalities in the market and σ measures the externalities in home production. As long as $\rho = \sigma$ it does not make a difference whether I employ expenditure data or wage data. However if $\rho \neq \sigma$ then my estimation of human capital externalities is an average of the two effects. In the absence of reliable wage data the way to separate these two effects is to add a land interaction term, as a proxy for the amount of home production. Thus, I estimate the following regression separately for non-primary industries and for primary industries:[28]

$$\ln(E_{ilt}) = \beta_0 + \beta x_{ilt} + \lambda h_{ilt} + \eta \overline{H}_{lt} + \upsilon \overline{H}_{lt} L_{ilt} + \gamma_0 S_{lt} + R + T + U + \epsilon_{ilt} \tag{2.21}$$

[28] The results are robust to the inclusion of an education interaction term.

in which all of the definitions are as in the main text, and $\overline{H}_{lt}L_{ilt}$ is a land interaction term.

The results indicate that the land interaction coefficient υ is negative, but the total externalities effect is still positive i.e. $\eta + \upsilon > 0$, implying that externalities in home agriculture activity are smaller but positive nonetheless. Consequently, employing expenditure data instead of wages provides a lower bound for the effect of externalities on wages.

However, employing expenditure data also has some limitations. One limitation of the consumption data is that it is household-level data rather than individual-level data. As a result I observe the average externalities at the household level, so my estimation is inaccurate to the extent that externalities may affect individuals within the household differently. Moreover, this empirical test cannot distinguish human capital externalities within the household from private returns to education. A second limitation is that expenditures are affected to some extent by household wealth and assets. As a result the observed effect of education on expenditures might actually capture the returns to physical capital. However, this issue is not of a large concern, because human capital externalities are shown above to actually be smaller for households possessing land, contrary to the expected effect if the coefficient of human capital externalities were capturing the returns to physical capital.

2.9 Data Appendix

This appendix documents the data sources and describes the various calculations and adjustments which have been made for the econometric analysis. As I mention in the main text the data used in this chapter comes principally from the NSS employment and unemployment surveys, which are combined with other sources for purposes of robustness.[29]

The National Sample Survey (NSS), initiated in the year 1950, is a nationwide, large-scale, continuous survey operation conducted in the form of successive rounds. In 1970 the NSS was reorganised and all aspects of its work were brought under a single government organisation, namely the National Sample Survey Organisation (NSSO) under the overall direction of a Governing Council to impart objectivity and autonomy in the matter of collection, processing and publication of the NSS data. The Governing

[29] I have added state controls from Besley and Burgess (2000) who collected data on GDP, taxes, and road infrastructure for the sixteen major states of India.

Council consists of 18 experts from within and outside the Government and is headed by a recognised economist or statistician.

The NSS employment and unemployment survey is conducted once every five years along with the annual consumer expenditures survey. The survey is a large and representative sample, which covers the whole of India except some districts of Jammu and Kashmir, some interior remote villages of Nagaland, and some inaccessible villages of the Andaman and Nicobar Islands.

The sample households are drawn based on a two-stage stratified random sampling procedure. In the first stage villages and urban blocks are randomly selected from a list of villages and urban blocks based on the previous census. In the second stage, the households are arranged by means of livelihood (main occupation), and area of landholding in rural areas, or monthly per capita consumption expenditure in urban areas.

The survey period is divided into four sub-rounds, each covering a period of three months. Within a particular sub-round, the fieldwork is spread out uniformly over different weeks/months as far as possible.

The survey is carried out by trained professional teams. The questionnaires and collection methods are identical across India as far as possible. As illiteracy is prevalent in India the questionnaires are completed by the interviewers. Hence non-response bias is unlikely to be correlated with literacy. In fact, Datt and Ravallion (2002) examined the representation of the poor in the NSS surveys and found it to accurately represent the poor households. They compared the NSS consumption data and the national accounts data on consumption of food staples, items which figured prominently in the budgets of the poor. The authors found little disparity between the two sources of data, suggesting that the NSS data was likely to accurately represent the poor households and their consumption.

The **local (regional/district) educational achievements** are calculated in three ways based on the NSS data, which includes information on the highest levels of general and technical education attained by the members in the household. The first way is to use the percentage of people in the population who achieved at least a certain education level, for example the percentage of literate people or the percentage of people completing at least primary school. More formally, if n_i is the number of people in household i, w_i is the weight assigned to the household, and d_{ki} is a dummy variable which takes the value 1 if the kth member of household i achieved at least a certain education level, then the local average educational attainment would be: $\frac{\sum_{i=1}^{I}\sum_{k=1}^{n_i} w_i d_{ki}}{\sum_{i=1}^{I} w_i n_i}$.

The second way is to calculate the weighted average of the ranked values assigned to each education level. More formally, v_{ki} is the ranked value assigned to the specific education level achieved by the kth member of household i, with $v_{ki} = 1$ for being illiterate, $v_{ki} = 2$ for being literate, $v_{ki} = 3$ for completing primary school, $v_{ki} = 4$ for completing middle school, $v_{ki} = 5$ for completing secondary school, and $v_{ki} = 6$ for being a graduate. If all other variables are defined as before, than the local average ranked education level would be: $\frac{\sum_{i=1}^{I}\sum_{k=1}^{n_i} w_i v_{ki}}{\sum_{i=1}^{I} w_i n_i}$.

The third way is to calculate the local average years of schooling. More formally, s_{ki} is the number of schooling years attributed to the kth member of household i, with $s_{ki} = 1$ for literate, $s_{ki} = 5$ for completing primary school, $s_{ki} = 8$ for completing middle school, $s_{ki} = 12$ for completing secondary school and $s_{ki} = 15$ for completing higher education. If all other variables are defined as before, than the local average number of schooling years would be: $\frac{\sum_{i=1}^{I}\sum_{k=1}^{n_i} w_i s_{ki}}{\sum_{i=1}^{I} w_i n_i}$.

The information on **monthly per capita expenditure** (in Rupees) is taken directly from the NSS employment and unemployment survey. The figures are based on a detailed questionnaire in the consumer expenditures survey, which is carried out simultaneously with the employment and unemployment survey.

In general, the **household characteristics** are based on the characteristics of the head of the household, although for purposes of robustness I also check if the results are affected by including an average of household member characteristics. The breakdown of industries into **non-primary industry and primary industry** is done under the National Industrial Classification system (NIC), with agriculture and mining being defined as primary industry. **Educational achievements of the household** are defined as the number of people in the household who achieved each specific level of education, namely the number of literate people, the number of people completing primary school etc.,[30]or alternatively as household average schooling years, as appropriate.

[30] All the results are repeated with education level being defined as the household head's education level.

REFERENCES

Acemoglu, Daron, (1996), "A Microfoundation for Social Increasing Returns in Human Capital Accumulation", *The Quarterly Journal of Economics*, 111, 778-804.

Acemoglu, Daron and Joshua Angrist, (2000), "How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws", MIT, Department of Economics, Mimeo.

Aghion, Philippe and Peter Howitt, (1998), "Endogenous Growth Theory", MIT Press, Cambridge MA.

Becker, G. (1964), "Human Capital", Colombia University Press, New York.

Besley, Timothy and Robin Burgess, (2000), "Land Reform, Poverty and Growth: Evidence from India", *Quarterly Journal of Economics*, 115 (2), 389-430.

Card, David, (1999), "Education in the Labour Market", in Handbook of Labour Economics, edited by O. Ashenfelter and D. Card, Amsterdam, North Holland.

Ciccone, Antonio and Giovanni Peri, (2000), "Human Capital Externalities in Cities", CEPR Discussion Paper 2599.

Datt, Gauravand and Martin Ravallion, (2002), "Is India's Economic Growth Leaving the Poor Behind?", *Journal of Economic Perspective*, 16(3), 89-108.

Duflo, Esther, (2000), "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia", MIT Department of Economics, Mimeo.

Duflo, Esther, (2001), "Schooling and Labour Market Consequences of School Construction in Indonesia: Evidence From An Unusual Policy Experiment", *American Economic Review,* 91 (4), 795-813.

Greenwald, Bruce C. (1983), "A General Analysis in the Estimated Standard Errors of Least Squares Coefficients", *Journal of Econometrics,* 22(3), 323-338.

Harbison, F. and Myers C. (1965), "Manpower and Education", xi, McGraw-Hill Vintage Books, New York.

Heckman, James J. (2000), "Policies to Foster Human Capital", paper prepared for the Aaron Wildavsky Forum, University of California, Berkely.

Jacobs, J. (1970), "The Economy of Cities", Vintage Books, New York.

Kimenyi, Mwangi S., Damiano K. Manda and Germano Mwabu, (2004), "Human capital externalities and private returns to education in Kenya", Department of Economics, Connecticut University, Working Paper Series 2004-08.

Kloek, Teun, (1981), "OLS Estimation in a Model Where a Microvariable is Explained by Aggregated and Contemporaneous Disturbances Are Equicorrelated", *Econometrica,* 49(1), 205-207.

Knight, John and Sharada Weir, (2004), "Externality Effects of Education: Dynamics of the Adoption and Diffusion of an Innovation in Rural Ethiopia", *Economic Development and Cultural Change*, 53 (1), 93-113.

Knight, John and Sharada Weir, (2006), "Production externalities of education: Evidence from rural Ethiopia", forthcoming in *Journal of African Economies.*

Kochar, Anjini, (2003), "Schooling, Wages and Profits: Some Negative Pecuniary Externalities from the Spread of Schooling in Rural India and their Consequences for Schooling Investments", Department of Economics, Stanford University, Mimeo.

Krueger, Alan B. and Mikael Lindahl, (2001), "Education for Growth: Why and For Whom?", *Journal of Economic Literature*, 39, 1101-1136.

Lucas, Robert E. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22(1), 3-42.

Mankiw, N. Gregory, David Romer and David N. Weil, (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics,* 107(2), 407-437.

Manski, Charles F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem", *The Review of Economic Studies*, 60(3), 531-542.

Michaud, Pierre-Carl and D. Vencatachellum, (2003), "Human capital externalities in South Africa", *Economic Development and Cultural Change*, 51(3), 603-28.

Mincer, Jacob, (1974), "Schooling, Experience and Earnings", National Bureau of Economic Research, New York.

Moretti, Enrico, (1998), "Social Returns to Education and Human Capital Externalities: Evidence from Cities", UC Berkley Department of Economics, Mimeo.

Moretti, Enrico, (2004), "Estimating the Social Return to Higher Education: Evidence From Longitudinal and Repeated Cross-Sectional Data", *Journal of Econometrics,* 121(1-2).

Moulton, Brent R. (1986), "Random Group Effects and The Precision of Regression Estimates", *Journal of Econometrics,* 32, 385-397.

Moulton, Brent R. and William C. Randolph, (1989), "Alternative Tests of Error Components Model", *Econometrica,* 57(3), 685-693.

Moulton, Brent R. (1990), "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units", *The Review of Economics and Statistics,* 72(2), 334-338.

Murphy, Kevin M., Andrei Shleifer and Robert W. Vishny, (1991), "The Allocation of Talent: The Implications For Growth", *Quarterly Journal of Economics,* 106(2), 503-530.

Psacharopoulos, George and Harry A. Patrinos, (2002), "Returns to Investment in Education: A Further Update", World Bank Policy Research Paper 2881.

Rauch, James E. (1993), "Productivity Gains from Geographic Concentration of Human Capital: Evidence From the Cities", *Journal of Urban Economics,* 34, 380-400.

Romer, Paul M. (1990), "Endogenous Technological Change", *Journal of Political Economy,* 98(5), 71-102.

Schumpeter, J. A. (1934), "The Theory of Economic Development", Harvard University Press, Cambridge MA.

Venniker, Richard, (2000), "Social Returns to Education: A Survey of Recent Literature on Human Capital Externalities", CPB Report 00/1 http://www.cpb.nl/eng/cpbreport/2000_1/s3_4.html

Young, Alwyn, (1995), "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience", *Quarterly Journal of Economics, 110(3), 641-680.*

TABLE 2.1

DATA DESCRIPTION

	ROUND 43 (1987-8)		ROUND 50 (1993-4)	
	Rural	Urban	Rural	Urban
Monthly Expenditure Per Capita (Rupees.)	161.70	274.56	290.85	506.64
	(96.41)	(196.36)	(157.6)	(333.87)
Schooling (Years)	1.92	4.41	2.92	5.43
	(0.81)	(0.86)	(0.81)	(0.75)
Education Level (Average Ranked Value)	1.93	3.26	2.08	3.47
	(1.309)	(1.73)	(1.41)	(1.77)
Literacy Rates*	0.43	0.76	0.47	0.78
	(0.49)	(0.43)	(0.50)	(0.41)
Primary School Completion Rates*	0.27	0.62	0.31	0.65
	(0.45)	(0.49)	(0.46)	(0.48)
Middle School Completion Rates*	0.14	0.45	0.19	0.52
	(0.35)	(0.50)	(0.39)	(0.50)
Secondary School Completion Rates*	0.07	0.32	0.09	0.37
	(0.25)	(0.47)	(0.29)	(0.48)
Graduation Rates*	0.01	0.11	0.02	0.14
	(0.12)	(0.31)	(0.13)	(0.35)
Non-Primary Industry	0.28	0.91	0.28	0.91
	(0.45)	(0.29)	(0.45)	(0.29)
Women Head of Household	0.10	0.10	0.10	0.10
	(0.30)	(0.30)	(0.30)	(0.31)
Scheduled Caste and Tribes	0.31	0.16	0.32	0.16
	(0.46)	(0.36)	(0.47)	(0.36)
Size of Household (Number)	5.10	4.70	4.90	4.46
	(2.58)	(2.60)	(2.40)	(2.33)
Hindu	0.85	0.79	0.86	0.80
	(0.35)	(0.40)	(0.35)	(0.40)
Muslim	0.10	0.14	0.09	0.14
	(0.30)	(0.35)	(0.29)	(0.34)
Age of Head of Household (Years)	43.43	41.59	43.80	41.91
	(14.05)	(13.85)	(13.70)	(13.71)
Self-Employed	0.12	0.34	0.13	0.34
	(0.33)	(0.47)	(0.33)	(0.47)
Number of Observations	80033	43705	66521	44364

Notes: Standard deviations are in parentheses. All values are proportions unless other units are indicated. *Education variables represent the proportion of people who completed each relevant education level or higher. See the Data Appendix for details on the construction and sources of the variables.

TABLE 2.2

AVERAGE EXPENDITURE PER CAPITA (RS.) BY EDUCATION LEVEL

	ROUND 43 (1987-8)		ROUND 50 (1993-4)	
	Rural	Urban	Rural	Urban
Illiterate	**152.36**	**193.94**	**268.06**	**345.48**
	(86.14)	(127.09)	(137.96)	(208.99)
	42013	10008	32115	8861
Literate	**174.90**	**220.55**	**306.80**	**409.20**
	(97.12)	(138.03)	(154.07)	(253.17)
	12758	5726	10595	5307
Completed Primary School	**193.87**	**243.79**	**334.43**	**436.75**
	(108.24)	(160.14)	(171.18)	(267.76)
	11215	7088	8617	5791
Completed Middle School	**212.04**	**272.11**	**359.86**	**483.78**
	(118.38)	(181.82)	(185.39)	(298.19)
	7086	5915	7185	6695
Completed Secondary School	**255.35**	**374.23**	**430.95**	**638.27**
	(138.75)	(246.58)	(220.55)	(388.49)
	5394	9299	6183	10659
Graduate and Above	**311.23**	**529.56**	**511.65**	**910.90**
	(158.54)	(325.31)	(249.03)	(495.16)
	1518	5655	1815	7031

Notes: Standard deviations are in parentheses, with the number of observations below the SD. Education level refers to the education of the head of the household. See the Data Appendix for details on the construction and sources of the variables.

TABLE 2.3
HUMAN CAPITAL EXTERNALITIES – BASIC RESULTS

	(1)	(2)	(3)	(4)	(5)	(6)
Local Education[1]	**Years**	**Level**	**Literacy**	**Primary**	**Middle**	**Secondary**
Regional Education	**0.038**	**0.124**	**0.242**	**0.41**	**0.564**	**0.659**
	(3.27)**	**(3.74)****	**(2.61)****	**(3.35)****	**(4.33)****	**(3.71)****
Household Schooling Years	0.078					
	(85.01)**					
Illiterate HHM[2]		-0.055	-0.055	-0.055	-0.056	-0.056
		(18.19)**	(18.16)**	(18.11)**	(18.74)**	(18.32)**
Literate HHM[2]		-0.017	-0.017	-0.017	-0.017	-0.017
		(6.02)**	(6.07)**	(5.90)**	(6.19)**	(5.98)**
Middle school HHM[2]		0.025	0.025	0.025	0.024	0.025
		(9.14)**	(9.21)**	(9.16)**	(9.01)**	(9.09)**
Secondary school HHM[2]		0.076	0.076	0.076	0.075	0.075
		(20.04)**	(20.22)**	(20.13)**	(19.91)**	(19.96)**
Graduate HHM[2]		0.211	0.212	0.212	0.211	0.21
		(32.68)**	(32.82)**	(32.88)**	(32.75)**	(32.55)**
Observations	222502	222563	222563	222563	222563	222563
R-squared	0.59	0.56	0.56	0.56	0.56	0.56
District Education	**0.047**	**0.155**	**0.417**	**0.428**	**0.57**	**0.874**
	(6.00)**	**(7.46)****	**(7.03)****	**(5.66)****	**(6.33)****	**(7.77)****
Household Schooling Years	0.077					
	(88.72)**					
Illiterate HHM[2]		-0.057	-0.057	-0.057	-0.059	-0.06
		(27.43)**	(27.25)**	(27.74)**	(27.82)**	(28.02)**
Literate HHM[2]		-0.017	-0.018	-0.016	-0.018	-0.018
		(7.28)**	(7.75)**	(6.84)**	(7.37)**	(7.54)**
Middle school HHM[2]		0.025	0.025	0.025	0.023	0.025
		(10.03)**	(10.37)**	(10.28)**	(9.47)**	(10.32)**
Secondary school HHM[2]		0.074	0.076	0.076	0.074	0.072
		(24.24)**	(24.70)**	(24.70)**	(23.91)**	(23.85)**
Graduate HHM[2]		0.193	0.195	0.195	0.193	0.19
		(31.35)**	(31.28)**	(31.35)**	(31.47)**	(31.72)**
Observations	160699	160746	160746	160746	160746	160746
R-squared	0.55	0.52	0.52	0.52	0.52	0.52

Notes: * significant at 5%; ** significant at 1%. The table reports OLS coefficients. Absolute t-statistics calculated using standard errors clustered by region/district are in parentheses. The dependent variable is monthly per capita expenditure. 1. Local education in column 1 is measured in terms of local average schooling years, in column 2 in terms of average education level, and in columns 3, 4, 5, and 6 in terms of literacy rate, primary school completion rate, middle school completion rate, and secondary school completion rate respectively. 2. HHM stands for number of household members. All regressions include regional fixed effect, time and sector fixed effects, with all additional variables detailed in table 2.4. The number of regions is 74, while the number of districts is 365.

TABLE 2.4
EDUCATION INTERACTION TERM

	REGIONS				DISTRICTS			
Education Interaction Term:	Literate		Primary School		Literate		Primary School	
Regional/District Schooling	**0.043**	**0.052**	**0.042**	**0.051**	**0.056**	**0.055**	**0.054**	**0.054**
	(3.57)**	**(3.78)****	**(3.53)****	**(3.73)****	**(6.89)****	**(6.33)****	**(6.76)****	**(6.23)****
Education Interaction	**-0.008**	**-0.008**	**-0.008**	**-0.008**	**-0.016**	**-0.015**	**-0.015**	**-0.013**
	(2.41)*	**(2.20)***	**(2.38)***	**(2.18)***	**(4.12)****	**(3.59)****	**(3.92)****	**(3.31)****
Dummy for Educated Household	0.095	0.098	0.054	0.055	0.121	0.118	0.079	0.073
	(8.48)**	(8.21)**	(4.49)**	(4.34)**	(10.40)**	(9.90)**	(6.45)**	(5.85)**
Household Schooling Years	0.073	0.075	0.076	0.077	0.071	0.073	0.074	0.076
	(78.12)**	(76.25)**	(79.80)**	(79.25)**	(79.45)**	(82.11)**	(78.28)**	(81.84)**
Time and Sector Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	No	Yes	No	Yes	No	Yes	No	Yes
Number of Regions/Districts	74	58	74	58	365	323	365	323
Observations	222441	175083	222441	175083	160655	135655	160655	135655
R-squared	0.59	0.59	0.59	0.59	0.55	0.53	0.55	0.53

Notes: * significant at 5%; ** significant at 1%. The table reports OLS coefficients. Absolute t-statistics calculated using standard errors clustered by region/district are in parentheses. The dependent variable is monthly per capita expenditure. All regressions include time and sector fixed effects, regional fixed effect, and control variables for age of head of household, size of the household, number of male/female workers, dummy for female head of household, dummy for scheduled castes and tribes, various industry dummies, and dummy for being self-employed. In addition, the even columns include state income per capita, state length of roads, state total taxations, and state population. Average education is measured in terms of average schooling years. Educated status is defined as being literate in columns 1,2,5, and 6, and as having completed at least primary school in other columns.

TABLE 2.5
ROBUSTNESS CHECK: STATE CONTROLS

	(1)	(2)	(3)	(4)	(5)	(6)
Regional Education	**0.047**	**0.139**	**0.227**	**0.471**	**0.699**	**0.814**
	(3.51)**	**(3.61)****	**(2.10)***	**(3.34)****	**(4.41)****	**(3.90)****
Household Schooling	**0.08**					
	(81.75)**					
Illiterate HHM[1]		-0.058	-0.058	-0.058	-0.059	-0.059
		(18.40)**	(18.44)**	(18.33)**	(18.65)**	(18.75)**
Literate HHM[1]		-0.018	-0.018	-0.018	-0.018	-0.018
		(6.43)**	(6.47)**	(6.30)**	(6.46)**	(6.50)**
Middle school HHM[1]		0.023	0.024	0.023	0.023	0.023
		(8.37)**	(8.48)**	(8.42)**	(8.18)**	(8.32)**
Secondary school HHM[1]		0.074	0.074	0.074	0.074	0.073
		(17.93)**	(18.07)**	(18.01)**	(17.86)**	(17.95)**
Graduate HHM[1]		0.209	0.21	0.21	0.209	0.208
		(29.68)**	(29.77)**	(29.85)**	(29.73)**	(29.61)**
Observations	175110	175143	175143	175143	175143	175143
R-squared	0.59	0.56	0.56	0.56	0.56	0.56
District Education	**0.048**	**0.148**	**0.385**	**0.405**	**0.563**	**0.947**
	(5.56)**	**(6.76)****	**(6.64)****	**(5.15)****	**(5.39)****	**(7.63)****
Household Schooling	0.079					
	(93.54)**					
Illiterate HHM[1]		-0.062	-0.061	-0.062	-0.063	-0.064
		(27.43)**	(27.36)**	(27.53)**	(27.95)**	(28.12)**
Literate HHM[1]		-0.02	-0.021	-0.019	-0.02	-0.02
		(7.81)**	(8.28)**	(7.41)**	(7.85)**	(8.07)**
Middle school HHM[1]		0.025	0.025	0.025	0.023	0.025
		(9.02)**	(9.25)**	(9.15)**	(8.59)**	(9.32)**
Secondary school HHM[1]		0.075	0.075	0.076	0.074	0.073
		(23.22)**	(23.43)**	(23.45)**	(23.02)**	(22.80)**
Graduate HHM[1]		0.194	0.196	0.196	0.194	0.191
		(30.26)**	(29.96)**	(30.09)**	(30.49)**	(30.46)**
Observations	135670	135699	135699	135699	135699	135699
R-squared	0.53	0.51	0.51	0.51	0.51	0.51

Notes: * significant at 5%; ** significant at 1%. The table reports OLS coefficients. Absolute t-statistics calculated using standard errors clustered by region/district are in parentheses. The dependent variable is monthly per capita expenditure. 1. HHM stands for number of household members. All regressions include regional fixed effect, time and sector fixed effects, and all additional variables including state controls are as in table 2.4. Local education in column 1 is measured in terms of schooling years, in column 2 in terms of education level, and in columns 3, 4, 5, and 6 in terms of literacy rate, primary school completion rate, middle school completion rate, and secondary school completion rate respectively. The number of regions is 74, while the number of districts is 365.

TABLE 2.6
NON-PRIMARY INDUSTRY INTERACTION TERM

	REGIONS		DISTRICTS	
Regional/District Schooling	**0.031**	**0.042**	**0.043**	**0.043**
	(2.72)**	**(3.20)****	**(5.18)****	**(4.96)****
NPI Interaction[1]	**0.015**	**0.015**	**0.011**	**0.013**
	(3.72)**	**(3.19)****	**(2.69)****	**(2.66)****
NPI Dummy[1]	-0.016	-0.015	-0.001	-0.005
	(1.04)	(0.91)	(0.05)	(0.37)
Household Schooling	0.079	0.081	0.078	0.08
	(87.38)**	(85.01)**	(90.05)**	(94.41)**
Time and Sector Fixed Effect	Yes	Yes	Yes	Yes
Regional Fixed Effect	Yes	Yes	Yes	Yes
State Controls	No	Yes	No	Yes
Number of Regions/Districts	74	58	365	323
Observations	222441	182898	166558	140932
R-squared	0.59	0.59	0.54	0.53

Notes: * significant at 5%; ** significant at 1%. The table reports OLS coefficients. Absolute t-statistics calculated using standard errors clustered by region/district are in parentheses. The dependent variable is monthly per capita expenditure.
1. Non primary industry (NPI) equal to 1 if the main industry of the household is not agriculture or mining. All additional control variables are as in table 2.4. Average education is measured in terms of average schooling years.

TABLE 2.7

EDUCATION AND NON-PRIMARY INDUSTRY INTERACTION TERMS

	REGIONS				DISTRICTS			
	Literate		Primary School		Literate		Primary School	
Regional/District Schooling	**0.037**	**0.045**	**0.038**	**0.046**	**0.052**	**0.049**	**0.051**	**0.049**
	(3.07)**	**(3.26)****	**(3.22)****	**(3.43)****	**(5.99)****	**(5.33)****	**(5.91)****	**(5.38)****
Education Interaction	**-0.013**	**-0.009**	**-0.015**	**-0.012**	**-0.022**	**-0.017**	**-0.021**	**-0.016**
	(3.11)**	**(1.88)**	**(3.76)****	**(2.60)***	**(5.23)****	**(3.71)****	**(4.95)****	**(3.52)****
NPI Interaction[1]	**0.017**	**0.02**	**0.014**	**0.016**	**0.014**	**0.018**	**0.012**	**0.016**
	(3.81)**	**(3.89)****	**(3.33)****	**(3.36)****	**(2.99)****	**(3.65)****	**(2.81)****	**(3.19)****
Education and NPI Interaction[1]	**0.003**	**-0.001**	**0.006**	**0.002**	**0.005**	**0.001**	**0.006**	**0.002**
	(1.32)	**(0.37)**	**(2.68)****	**(0.82)**	**(2.36)***	**(0.29)**	**(2.72)****	**(0.86)**
NPI Dummy[1]	-0.035	-0.035	-0.026	-0.026	-0.024	-0.029	-0.016	-0.019
	(2.31)*	(2.09)*	(1.7)	(1.54)	(1.82)	(2.08)*	(1.18)	(1.4)
Dummy for Education	0.109	0.104	0.07	0.065	0.133	0.124	0.09	0.079
	(9.34)**	(8.17)**	(5.71)**	(4.89)**	(11.43)**	(10.24)**	(7.25)**	(6.20)**
Household Schooling	0.074	0.076	0.077	0.078	0.073	0.075	0.075	0.077
	(79.59)**	(79.35)**	(80.78)**	(81.36)**	(79.09)**	(81.36)**	(77.29)**	(79.62)**
State Controls	No	Yes	No	Yes	No	Yes	No	Yes
Number of Regions/Districts	74	58	74	58	365	323	365	323
Observations	222441	175083	222441	175083	160655	135655	160655	135655
R-squared	0.59	0.59	0.59	0.59	0.55	0.53	0.54	0.53

Notes: * significant at 5%; ** significant at 1%. The table reports OLS coefficients. Absolute t-statistics calculated using standard errors clustered by region/district are in parentheses. The dependent variable is monthly per capita expenditure. All regressions include regional fixed effect, time and sector fixed effects, and all additional variables are as in table 2.4. Average education is measured in terms of average schooling years. 1. Non primary industry (NPI) equal to 1 if household main industry is not agriculture or mining.

TABLE 2.8

SUMMARY OF EDUCATION AND INDUSTRY EFFECTS

8.A	Primary Industry	Non-Primary[1]
Illiterate[2]	3.7	5.4
Literate	2.4	4.1

8.B	Primary Industry	Non-Primary[1]
Below Primary[2]	3.8	5.2
Primary + [3]	2.3	4.3

8.C	Primary Industry	Non-Primary[1]
Below Middle	3.4	4.6
Middle + [3]	3	4.8

Notes: The table reports the sum of the appropriate coefficients from regression (2.11) with average regional education. The control variables are the same as in table 2.7. 1. The difference between non-primary industries and primary industries is significant at 5%. 2. The difference between educated and uneducated households is significant at 5%. 3. The gap between educated and uneducated households is significantly different in primary and non-primary industries.

Chapter 3

Education and Credit Constraints in India

3.1 Introduction

This chapter analyses the effect of credit constraints on children's schooling in India. The widely-cited justification for government intervention to finance education, including considerable loans and generous subsidies, is the observation that people are unable to mortgage their future income to finance their studies. Thus in many countries, and particularly in the USA and the UK, policy on education is based on the assumption of credit constraints. Moreover, one of the key explanations for the relatively low educational attainments in developing countries in spite of high potential returns is households' limited access to credit.[1]

But the empirical evidence for credit constraints affecting education is indirect and inconclusive. In developed countries the main focus of the empirical literature is to test for credit constraints affecting schooling decisions. To identify these constraints researchers take one of two key empirical approaches. The first approach builds on the separation between investment decisions and income. With perfect credit markets educational decisions depend only on interest rates and future returns but not on income. Therefore the observed correlation between family income and schooling achievements has been taken as an indicator of credit constraints.[2]The second approach is based on research estimating the returns to education. Research, summarised in Card (1999,

[1]Other explanations include low quality of schools, insufficient market for skilled labour, and culture barriers.

[2]See Carneiro and Heckman (2003) for a critical review of the literature.

2001), has found that in the majority of cases the Instrument Variable (IV) coefficients have been significantly larger than the OLS coefficients. IV coefficients indicate the potential returns from a marginal increase in schooling, while the OLS coefficients evaluate current returns to schooling. The fact that potential returns to schooling are greater than actual returns implies that investment is not being fully exploited, due to some form of constraint.

However Carneiro and Heckman (2002, 2003) point out that there are serious interpretation problems with the existing evidence and that limited credit access is only one possible explanation of existing results. Since current income is highly correlated with permanent income the correlation between income and education can indicate a long-term effect of income on ability and motivation. In their papers they show that once ability is taken into account the effect of current income on American college enrolment is in fact limited. Furthermore, heterogeneous returns to education can cause the IV coefficient to be greater than the OLS coefficient even in the absence of credit constraints. Carneiro and Heckman estimate that no more than 8% of the US population are prevented by credit constraints from attending college, and therefore claim that further policies targeted to deal with credit constraints are less likely to eliminate the extensive wealth gap in American college enrolment. Given the relative importance of early childhood education they recommend to devote more resources toward intervention at an early age. Moreover, education is not just an investment but also a consumption good. The demand for education increases with income because education has a status value. Hence a positive relationship between parental income and child schooling could simply be due to education being a normal good rather than as a result of credit constraints.

Surprisingly, in spite of the clear evidence of credit constraints in developing countries, the literature on their effect on education is quite limited. More importantly, the existing literature focuses mainly on inferring from schooling decisions the structure of the credit markets. It shows that schooling decisions respond to income shocks and takes this to indicate incomplete credit markets. However it does not evaluate the marginal effect of credit availability on school achievements. For example, Edmonds (2005) studies the timing effect of anticipated large cash transfers on child labour and schooling in South Africa. He concludes that households do not borrow against anticipated pensions to finance child education and takes this to indicate the existence of credit constraints. Similarly, Jacoby and Skofias (1997) employ data on variation in rainfall from six Indian villages as a proxy for unexpected income shocks. Based

on the response of school attendance to idiosyncratic shocks they reject the existence of complete intra-village credit and insurance markets and conclude that with limited credit markets parents draw on child labour to smooth their income.

A second key drawback of the existing literature is that it employs various measures of credit constraints which are likely to be endogenous and to directly affect returns from education. For example, employing data for Peruvian children Jacoby (1994) studies the relationship between credit constraints and the probability of lagging behind at school. As an indicator of credit constraints he employs household assets and concludes that children in households with credit constraints leave school earlier than those with access to credit. However, a different explanation of the results could be that richer households care more about education. Beegle, Dehejia and Gatti (2003) also find that transitory income shocks in Tanzania increase child labour and that having collateral mitigates these effects, but again the authors cannot rule out the possibility that richer households have higher returns from schooling. Likewise, Guarcello etc. (2002) identify a sub-sample of credit-constrained households in Guatemala using self-reported information on denial of credit and why families were not able to apply for credit. They find that income shocks increase child labour and reduce school attendance, and that limited access to credit causes similar effects. Once more, assumption of causality is problematic as it is possible that unmotivated households are excluded from the credit market and also care less for education.

This chapter takes a different approach and offers three main contributions to the existing literature. Most importantly, it focuses on the key question of identifying the marginal effect of credit constraints instead of merely verifying their existence. It does so by adopting a more direct way of estimating credit constraints. Unlike previous research I directly observe credit constraints and therefore do not count on the correlation between income and schooling, which could have some alternative explanations. Secondly, while some literature employs land and assets as an indicator of collateral, both of which are clearly endogenous and could directly affect the returns to education, I employ bank availability as an indicator of credit availability. While I cannot claim that bank availability is a completely exogenous variable it is a significant improvement on employing household assets, and on the spectrum ranging from completely exogenous to completely endogenous it is closer to the exogenous end. Thirdly, by focusing on a developing country the chapter contributes to a better understanding of the indirect costs of credit constraints in developing countries.

Households in developing countries are more exposed to risk for various reasons. For

example, agricultural activity is an importance source of income, and poor sanitation increases the risk of infectious diseases. Furthermore market opportunities to deal with risk are limited due to information problems, absence of collateral, and enforcement difficulties. The restricted formal mechanism to deal with risk causes households to engage in expensive and less efficient methods to smooth their consumption and income.[3]Research reveals that households choose less risky but less advanced productive technology (for example Binswanger and Rosenzweig (1993), Morduch (1994)). This chapter shows that this effect is not limited to production decisions but also extends to human capital investments. Thus households surrender future benefits from education in favour of short-term benefits from child labour when access to credit is limited.

India provides an ideal setting given the low educational achievements, significant time-state variation in bank availability, and high quality data. Employing the National Sample Survey (NSS) employment and unemployment round 55 combined with information on bank availability this chapter studies the effect of credit constraints on level of school attendance and lagging behind at school. In addition, employing panel data this chapter explores the effect of changes in family circumstances on schooling decisions. Finally, the chapter studies the effect of bank availability on school completion.

The main finding of my empirical analysis is that bank availability significantly increases the probability of attending school and reduces the probability of lagging behind at school. Moreover access to credit significantly mitigates the importance of income in determining school outcomes. Using panel data the chapter reveals that if one of the parents becomes unemployed the probability of continuing schooling drops by 8 percentage points, but credit availability almost completely mitigates the effect. These results are also evident in actual school completion results with bank availability during schooling years increasing primary school completion by 3 percentage points.

The chapter is organised as follows: section 2 describes the relevant background and the data; section 3 presents the econometric method and results; section 4 concludes.

[3]Informal mechanisms to deal with risk help to smooth consumption. However, existing research shows that these informal mechanisms are far from perfect.

3.2 Background and Data

3.2.1 Credit Availability

Credit and insurance markets in India, as in many other developing countries, suffer from enforceability and information problems. Consequently access to credit from the formal sector is relatively limited. The share of formal or institutional sources in total debt was only 61.2% in the 1990s.[4]For insurance, more than 85 percent of the population relies primarily on extended family networks and informal saving arrangements for old-age security and insurance.

Yet India compares favourably with other developing countries in terms of the distribution of financial services. The average population served per commercial bank branch in India was around 15,000 in 2002, or if including branches of rural cooperative banks, 12,800, close to levels in Indonesia and Mexico.[5]

Furthermore, following Indira Gandhi's bank nationalisation drive launched in 1969, bank branches are relatively available and accessible even in rural areas. These government regulations undertook to affect bank location and lending practices to favour the poor. As part of this effort, in 1977 they imposed a 1:4 branch licence policy which required banks to open four branches in rural locations without banks for every branch opened in locations already covered by an existing bank.

This legislation resulted in substantial achievements in enhancing access to credit in rural areas. In 1947, the first survey of rural indebtedness documented that moneylenders and other informal lenders met more than 90% of rural credit needs. The share of banks in particular was only about 1% of total rural household debt. Following bank legislation, the share of banks in rural household debt increased to about 29% in 1991 and the share of formal or institutional sources in total debt reached 61.2%.[6]Correspondingly, the share of moneylenders declined steadily over these four decades. At the present time, according to official surveys, the median distance to the nearest financial institution ranges from 2 km (post office branches) to 5 km (commercial banks, cooperative banks) and the median time taken to travel to the nearest commercial bank, cooperative or rural bank is 30 minutes.

The ambitious bank branch expansion program sought to expand the rural bank

[4]Basu and Srivastava (2005).

[5]In addition, the share of financial assets in GDP in India is about 93%, compared to 81% in Argentina and 68% in Mexico. (Basu and Srivastava (2005)).

[6]Basu and Srivastava (2005), Burgess and Pande (2005).

branch network and also equalise individual access to banks across Indian states. While it was highly successful in increasing bank branch availability and in reducing inequality across Indian states, there remains a substantial variation in the availability of bank branches across India. For instance, in 1996 there was one branch serving an average of 10,000 people in Punjab compared with almost twice as many people per branch in Assam.

On the insurance side, unemployment benefits and other forms of income insurance are available only to a small minority of the labour force, being those employed in the formal sector.[7]One of the key schemes is provident funds, a retirement saving tool available mainly to formal sector employees with only 12% of the working population being covered.[8]Provident funds, besides providing an old-age income, are used to mitigate income shocks and their holders are entitled to withdraw their funds in event of unemployment or job change. In addition, members of provident funds are automatically members of the Employees Deposit Linked Insurance, which pays compensation to the legal heirs upon death of a member.

3.2.2 Data

The chapter seeks to identify the effect of credit availability on schooling decisions. Following criticism by Carneiro and Heckman (2002, 2003) of the circumstantial evidence of credit constraints from the USA, considerable attention has been devoted towards obtaining a more direct and exogenous measure of credit constraints.

My empirical analysis makes use of household data, combined with state-level variables on the availability of banks. The household data is the National Sample Survey (NSS) employment and unemployment round 55 conducted in 1999. A panel dataset was created for a sub-sample of households which were revisited three months later. After restricting the sample to children aged 6-18 who were revisited, the data has around 31,000 observations for rural areas and 19,000 observations for urban areas. The state-level variables come from Burgess and Pande (2005) and Besley and Burgess (2000).[9]

[7]The government initiated some programs under which those in need and without regular employment are obliged to undertake work-related activity in return for state income transfers. Some of these programs are aimed at the unemployed in the non-formal sector. However, the programs helping the non-formal sector are limited and oblige people to undertake work in return for transfers.

[8]Another important scheme is the Employees' State Insurance Act (ESI), which provides health insurance.

[9]Description of the state-level data can be found in the Data Appendix.

The NSS data contains detailed information on household characteristics including religion, social group, household composition, and Monthly Per Capita Expenditure (MPCE). In addition detailed information on the parents is available, including employment status, principal activity, education, age, and whether they are members of a provident fund. The information on each child includes not only gender and age but also information on the principal activity of the child, i.e. whether he/she works, does household chores, or attends school. It also includes information on educational status, i.e. whether the child currently attends school, has dropped out of school, or never attended school. Furthermore, if the child attends school the data contains information on the school level (primary, middle or secondary).

The key dependent variable is a child school attendance dummy, based on whether school is the child's main activity. In addition to this, based on child age and the type of school attended, I construct a variable indicating whether a child is lagging behind his age group at school. For example if a child is 16 but still attends middle school this variable equals 1. As I have no information on the precise class attended by each individual but only on the type of school I can capture only some of the cases of lagging behind at school and my measure is most accurate for children at the age of starting a new school level (i.e. 12 or 15).[10]

To measure the extent of credit availability I employ information on the number of bank branches. Although this is a crude measure which ignores other formal and informal sources of credit, I believe it successfully captures the cost of borrowing as banks account for a substantial amount of household credit, and also as other informal arrangements are much more expensive sources of credit which only emerge as a response to a fault in the formal sector.[11]I make use of the substantial variation between states in the availability of bank branches to estimate the effect of credit on schooling outcomes. I also make use of the change in credit availability over time as a result of the bank expansion program to obtain supplementary evidence regarding credit constraints and actual school level achieved.

I employ the information on provident fund membership as an indicator of family access to insurance. In my data, a provident fund is available for 12% of the children. The

[10]The main argument against using a school lagging variable is that it may be the result of data errors. My data contains two visits for each individual and allows me to compare the age recorded in each visit. I found only a small percentage of people with contradicting ages and these were omitted from my analysis.

[11]The bank branch expansion caused a dramatic rise in the importance of banks as a source of household credit, and bank loans now account for more than 30% of rural household debt.

major drawback of this variable is that provident fund membership is available mainly for the formal sector and is therefore liable to be correlated with other unobserved family characteristics. Consequently I report results for provident fund membership but keep the main focus of this chapter on the effect of bank availability.

Table 3.1 presents a summary of the main variables classified by rural and urban location. School attendance is lower in rural areas with 64% of children of school age attending school, compared with 78% in urban areas. Lagging behind at school is more prevalent in rural areas. In rural areas 30% of children who are supposed to be starting middle and secondary schools areas are still attending primary and middle schools respectively; in urban areas the equivalent figure is only 24%. The mean monthly expenditure is 416 Rs. in rural areas compared with more than 50% higher in urban areas.

Table 3.2 provides state-by-state information on school attendance and bank availability. School attendance ranges from as low as 47% in Bihar to 85% in Kerala. The number of people per bank branch varies between 9,700 and 20,000. Generally speaking, states with reduced bank availability have lower school attendance figures.

3.3 Econometric Analysis

I am interested in the effect of credit constraints on schooling in India. In the following four sub-sections I present my econometric approach and results. The first two sub-sections present some basic results. They examine the correlation between expenditure and school achievements, and the extent to which bank availability affects these relationships. The last two sub-sections present the key results based on panel data estimations. I first examine the effect of income shocks on school attendance and then examine the long-term effect of credit on school outcomes.

3.3.1 Household Characteristics and School Outcomes

As a first stage I simply explore the correlation between household income[12] and school attendance. Thus employing the NSS employment and unemployment data, I run the following probit regression:

[12] I employ household expenditure as a proxy for family income. In developing countries it is difficult to obtain a reliable measure of income because a large share of income is generated by home production, and a considerable part of the wage is received in goods and kinds.

$$h_{is} = \alpha + \beta X_{is} + \gamma I_{is} + S + \epsilon_{is} \tag{3.1}$$

where h_{is} is a dummy variable indicating whether child i currently attends school, X_{is} are child and family characteristics such as demographic composition, land ownership, religion, gender, age, and household composition, S is state fixed effect and I_{is} is family income proxied by monthly per capita expenditure measured both as a continuous variable and as a quartile-based dummy variable allowing for non-linear income effects. I estimate equation (3.1) separately for each type of school to allow for different effects of income on primary, middle, and secondary schools. To account for possible serial correlation in the error terms I cluster the standard errors by state.

Table 3.3 reports the basic results on the probability of attending school. Unsurprisingly, the correlation between income and school attendance is positive and significant. As expected parents' education has a very significant and positive effect on the probability of attending school. While the level of maternal education over and above basic literacy does not have a notable effect, paternal education has a diminishing positive marginal effect. The demographic, ethnic and religious effects are also as anticipated, with the existence of young siblings reducing the probability of attending school by 2 to 4 percentage points, females being 12 percentage points less likely to attend school than males, and a reduced level of school attendance being reported for Muslim children and traditionally disadvantaged groups. The effect of land is insignificant, possibly as a result of two opposing effects: on the one hand land can be used as collateral and therefore can reduce credit constraints, while on the other hand with incomplete labour markets land increases the returns for child labour.[13]

To give some idea of the magnitude involved Column 1 shows that moving from the lowest income quartile to the second quartile increases the probability of attending school by 6.6 percentage points. The effect of moving from the second quartile to the third quartile is smaller with a further 3-4 percentage points increase in probability, and moving to the final quartile has a similar marginal effect. Column 2 employs a continuous income variable and reveals results of a slightly greater magnitude. A rise of 300 Rs. in monthly per capita income[14]increases the probability of attending school by 9 percentage points.[15]Columns 3-5 present the results separately for each age group.

[13]Bhalotra (2000b).

[14]Monthly income of 300 Rs. per capita is around the 25-percentile level.

[15]The small differences between columns 1 and 2 are probably a result of omitting outliers in the continuous regressions.

The effect of income on the probability of attending school is much higher for children of secondary school age than for children of primary and middle school age.

Following Jacoby (1994) I next estimate equation (3.1) employing school lagging as the dependent variable for all children currently attending school.[16]Table 3.4 displays the results for school lagging conditional on school attendance. Column 1 presents the results for children aged 12 and 15 and indicates a negative effect of income on school lagging.[17]An increase of 300 Rs. in monthly per capita income reduces school lagging by 2.4 percentage points. Separating these results by age group, columns 2 and 3 reveal the effect for children of middle school age to be not significant, but the effect for children of secondary school age to be negative and significant. Column 4 introduces a rural interaction term for children of secondary school age.[18]Households in rural areas have a greater income dependency on agriculture. In addition, poor infrastructure and sanitation increase the exposure to health risks in rural areas. Therefore I expect credit availability to be of greater importance in rural areas. As predicted, the effect of income on school lagging becomes insignificant for urban areas, but remains negative and significant in rural areas where credit constraints are more apparent.

Overall the above results are in line with expectations and demonstrate a large and statistically significant correlation between income and school achievement. Since educational decisions are not dependent on income when credit markets are perfect, a possible interpretation of these results is that households are credit-constrained. However, while this correlation may be consistent with credit constraints it is by no means an exclusive explanation. As pointed out by Carneiro and Heckman (2002, 2003), unobserved family characteristics such as motivation and ability could increase both school attendance and family income and consequently create the observed correlation. Moreover as mentioned earlier, part of the demand for education is related to its status value rather than investment value, especially in countries like India, where the returns to primary school education are quite low. In that case the demand for education will increase with income and create a positive correlation between education and income regardless of credit constraints. A more explicit and direct approach is required for estimating the importance of credit constraints and for quantifying the influence of

[16]Jacoby (1994) argues that using a school lagging variable mitigates the issue of unobserved family characteristics because it is not clear why poor children should lag behind at school as opposed to leaving school altogether. However, this argument is problematic because if poor children come from less motivated families then they can exhibit irregular school attendance and repeat classes as a result.

[17]Children start middle school at the age of 11 and start secondary school at the age of 14.

[18]I do not report the results for middle school children because the results are not significant even without an interaction term.

credit availability on schooling decisions.

3.3.2 Credit Availability

Now I turn to a more direct approach to estimate borrowing constraints and the influence of credit availability. The main variables of interest are credit availability and its interaction with household expenditure. My main measure of credit availability is based on the number of people per bank branch in each state. If in the child's state of residence 15,000 people or less are served by one bank I define bank availability to be 1.[19]In an alternative specification, I also consider provident fund membership which allows households access to money at times of economic distress as an indicator of credit accessibility. As provident fund membership is more likely to be correlated with unobserved household characteristics, I limit the use of this variable. I estimate the following probit regression:

$$h_{is} = \alpha + \beta X_{is} + \gamma I_{is} + \eta B_{is} I_{is} + S + \epsilon_{is} \tag{3.2}$$

where $B_{is}I_{is}$ is an income-credit interaction term, and all other variables are defined as in equation (3.1). To account for possible serial correlation in the error terms I again cluster the standard errors by state. As before, I estimate equation (3.2) multiple times using a dummy for school attendance as one dependent variable and a school lagging variable as a second dependent variable.

The identification takes advantage of the variation in bank availability. However, bank branch placement is not random. In general, banks choose to locate themselves in relatively rich states with substantial economic activity. If these states have more banks and they also provide better education or more opportunities for educated people the coefficients on bank availability would be an overestimate of the true impact of access to banks on education.[20]Therefore I include state-fixed effect to account for unobserved state characteristics.

For the interaction term I employ a dummy variable for banks serving no more than 15,000 people.[21]The coefficient of interest is η. Coefficient η measures the interaction between credit availability and family income and therefore is less likely to be

[19]15,000 is the mean number of people served per branch in India; for a specific list of states see table 2. When I change the threshold, results do not change significantly.

[20]The government of India has attempted to equalise access to credit across the Indian states, however it is still the case that relatively rich states have better access to credit.

[21]If instead I employ the number of people per branch the results stay qualitatively the same.

correlated with unobserved state characteristics. I expect credit availability to mitigate the importance of income in determining schooling outcomes, so η is expected to be negative in school attendance regressions and positive in school lagging regressions.

Table 3.5 reports the results for credit availability. Section A presents the results for bank availability. Income increases the probability of attending school but this effect is significantly mitigated when banks are widely available. Column 2 discloses the results for lagging behind at school. Once more bank availability mitigates the effect of income. Section B presents the results for provident fund membership. Access to a form of insurance mitigates the importance of income; however this effect is only significant for the school attendance regression.

The magnitude of the effect is quite impressive, especially considering that the interaction term is roughly half the size of the income effect. For example in column 1 a rise of 300 Rs. in monthly per capita income increases the probability of attending school by 13 percentage points. However, if banks are widely available the effect is reduced to 5 percentage points. Column 2 displays an even greater influence on school lagging regressions. The bank interaction term is the same magnitude as the income effect, so where banks are available income does not affect the probability of lagging behind at school, compared with an adverse effect of 6 percentage points where banks are not widely available.

Overall, the above results establish a large and statistically significant influence of credit availability on school achievement. Access to credit not only improves educational outcomes but also reduces wealth inequality in school achievement. However, I cannot completely rule out that some unobserved state characteristics could be correlated with bank availability and affect rich people differently to poor people. Therefore now I turn to a more comprehensive analysis based on panel data estimation.

3.3.3 Income Shocks

One of the significant implications of credit constraints is that temporary income shocks can force households to engage in expensive and less efficient methods to smooth their consumption and income. Therefore, households might choose to rely on child labour in spite of its heavy long-term cost. To study the effect of shocks on schooling I construct a panel dataset by merging the data from the first visit with data collected three months later. Thus the panel data has among other variables information on changes in employment and school activity in the three months between the two visits. In the

three months between the two visits around 5% of children dropped out of school, and 4% re-entered school. 5% of the families had a new child, 3% lost a member of the family, and around 4% of parents lost their job.

I estimate the effect of changes in family circumstances on schooling decisions for all children who attended school in the first visit. The regression is:

$$h_{2is} = \alpha + \beta \Delta X_{is} + \gamma \Delta Z_{is} + \eta B_{is} \Delta X_{is} + S + \epsilon_{is} \tag{3.3}$$

where h_{2is} is a dummy variable equal to 1 if the child attends school at the time of the second visit. Child and family characteristics which never change are dropped from the regression. ΔX_{is} and ΔZ_{is} are two types of changes in family circumstances. The first type is a job loss by one or more of the parents. The second type is a change in family size due to a birth, or a household member leaving the house.[22] $B_{is}\Delta X_{is}$ is an interaction term between bank availability and job loss.

Out of these two types of shocks the first one is more likely to represent a transitory shock. Permanent income shocks are likely to affect schooling decisions regardless of borrowing opportunities, whereas temporary income shocks can be overcome through temporary money borrowing if the opportunity is available. Therefore while I expect both β and γ to be negative, inferring from γ on the credit market is less obvious. The interaction term is between bank availability and loss of employment. The prediction is that $\eta > 0$ because the adverse effects of income shocks should be mitigated where credit is easily available.

In general a decision to drop out of school is likely to be determined by the long-run returns to education. Hence if returns to education vary over time and are also correlated with credit constraints it can impose a bias on the estimation. However, my analysis focuses on the decisions to drop out of school over a period of three months. It is unlikely that the returns to education would dramatically change in such a short period.

Table 3.6 displays the panel data results. The left part of the table shows the basic results for each age group. On average the effect of a departure in the family or a birth is between 2 to 3 percentage points. It is interesting to note that older children are more vulnerable to a change in the family condition, probably because they can help by finding a job or by doing household chores. The effect of job loss is

[22] I cannot identify the reason for leaving the household and therefore cannot separate between death, marriage and other reasons for leaving the household.

considerable and reduces the probability of staying in school by 8 percentage points. The second part of the table introduces an interaction term between credit and job loss. Bank availability reduces the adverse effect of job loss but the coefficient is significant only for older children. This means that credit can help to overcome job loss without taking older children out of school. This could be because parents are more willing to borrow to smooth school attendance at secondary school level due to the high returns to secondary education in the Indian labour market. However, parents may not be willing to similarly borrow to smooth primary school or middle school attendance as the returns to education are low at these levels.[23]

Overall these results show that credit constraints limit families' ability to smooth school attendance. While these results are somewhat weaker compared with those presented above, they are much more compelling as I employ a panel dataset and account for unobserved family characteristics.

3.3.4 School Completion

The above results establish a negative effect of credit constraints on current school outcomes. It is interesting to next check whether these results present themselves in actual long-term educational outcomes or whether these results merely indicate a temporary withdrawal from school with limited long-term effects.

I now therefore turn to examining the effect of bank availability on obtained education levels. The analysis here focuses on the education obtained by individuals who were adults in 1999 and were six years old sometime between 1960 and 1990. It examines how bank availability at the time those people went to school has affected their educational outcomes. The regression for individual i from state s and cohort c is:

$$h_{ics} = \alpha + \beta X_{ics} + \delta B_{ics} + S + C + \epsilon_{ics} \tag{3.4}$$

where h_{ics} is an educational outcome variable, the variable of interest B_{ics} is a variable measuring the number of schooling years during which banks were widely available, and X_{ics} is a vector of exogenous variables determining educational outcomes.[24] S is state fixed effect and C is cohort fixed effect, which account for unobserved state

[23] The economic returns for the first 8 years of school in India are quite low and estimated to be lower than 2%. (See Duraisamy (2002), Kingdon (1998), Kingdon and Unni (2001) and Vasudeva-Dutta (2006)). However, the returns are much higher for secondary school.

[24] Individual controls including gender, social group, rural residency, and other state controls from the year when each individual started school.

and time characteristics.

The data allows me to measure educational outcomes only in terms of education levels rather than schooling years. Therefore I employ two types of educational outcomes. The first one is a dummy variable for completing primary school. The second type is a proxy for average schooling years. The measurement of the latter is not accurate since it does not fully take into account schooling years of people who dropped out of school before completing a particular education level. For example, a person who studied for ten years appears in my data as someone who completed middle school but not secondary school and therefore accounts only for eight years of primary and middle school. However, I find it useful to employ the schooling measure as it is easier to interpret the coefficient on average schooling years.

Table 3.7 reports the results. The left part of the table describes the marginal effects from probit estimation on the probability of completing primary school. Each additional school year under a widespread banking system increases the probability of primary school completion by 0.6 percentage points. Hence the probability of primary school completion grows by 3 percentage points where banks are available throughout all five years of school. Column 6 shows that each year of bank availability increases the number of years in school by 0.05 years.

An important concern with the specification in equation (3.4) is the possibility of alternative local characteristics which could enhance both bank expansion and educational achievements. In the absence of a policy requirement on bank branch placement I would expect greater branch expansion in richer states, and if richer states provide better education then I would overestimate the true effect of bank availability. On the other hand, if the Indian central bank was successful in forcing banks to open relatively more branches in poorer states then the above logic suggests that I would underestimate the true effect of bank availability.

The inclusion of time fixed effect and state fixed effect accounts only for fixed unobserved characteristics, but does not perfectly eliminate the prospect of specific time and state unobserved variables. Therefore I include a set of specific time-state control variables, including political and economic variables. Political variables account for a change in political atmosphere affecting educational policies. Economic controls account for economic development enabling more investment in education. Columns 2 and 7 incorporate an additional set of control variables: all regressions now include educational inputs as well. Together these variables help to control for alternative factors which could theoretically affect schooling decisions. Furthermore as a robustness

check columns 3 and 8 provide the results with additional state-specific time trend. The table shows that the results are generally not sensitive to the inclusion of state controls, although a state-specific time trend reduces the magnitude and the significance of the marginal effect reported in column 3.

Credit constraints are expected to have a larger effect in rural areas because income is largely dependent on agricultural activity and because health risks are also higher. Therefore I lastly estimate equation (3.4) separately for rural and urban areas. I expect the effect of credit availability to have greater significance in rural areas. Columns 4, 5, 9 and 10 show that the effect of credit availability on educational outcomes is only significant in rural areas. Given poverty levels in India it is hard to believe that credit constraints have no effect in urban areas of India. One explanation for these results is that the returns to primary education are much higher in rural areas compared with urban areas (Duraisamy (2002)), so people might be less willing to borrow money to finance primary education in urban areas. In addition it is important to notice that these results refer to long-term effect on school outcomes. According to my findings primary school completion rates are not affected by credit constraints in urban areas but school attendance is affected. It is possible that even if income shocks force children to withdraw from school temporarily they can go back to school when the situation improves.

Overall these results suggest a significant positive effect of bank availability on education levels, which is more pronounced in rural credit-constrained areas. They show that the effect of imperfect credit is not limited to a short-term withdrawal from school but also has a long-term impact.

3.4 Conclusion

In this chapter I have tested the influence of credit constraints on school outcomes in India. The policy on education in many countries, including government loans and subsidies for education, is based on the assumption of borrowing constraints. These policies, together with the substantial private returns of education and the contribution of education to economic growth, make this issue of considerable interest.

I have chosen to focus on a developing country where households are considerably exposed to risk and where official mechanisms to mitigate such risk are limited. Existing literature has established the adverse effect of limited borrowing opportunities on investment in physical capital. This chapter shows that the effect extends also to in-

vestment in human capital. In the absence of formal mechanisms for income smoothing, households rely on child labour and sacrifice long-term benefits of education. Looking specifically at India is valuable given the low educational achievements and substantial state variation in both educational outcomes and credit availability.

The main contribution of this chapter is the use of a direct and relatively exogenous measure of credit constraints. Unlike previous literature I do not rely on the correlation between income and education to establish the existence of credit constraints and do not employ household assets as an indicator of collateral. Instead I make use of information on bank availability.

My findings suggest that credit constraints play an important role in educational outcomes in India. Furthermore credit constraints explain a substantial part of existing wealth inequality in education. Where banks are widely available income plays only a small role in determining educational achievements. However, when access to credit is limited and expensive, income has a large and statistically significant effect on education. Credit also helps to mitigate the adverse effects of transitory income shocks. Finally, I find that the negative effect of credit constraints in rural areas is not limited to short-term outcomes but is seen also in long-term outcomes of school completion rates and schooling years. This is an important finding as more than 70% of the population lives in rural areas.

From the policy perspective my results suggest that programs to expand schools without understanding the risks and constraints faced by households might be unsuccessful. In addition, compulsory school laws could damage household welfare. On the other hand, programs to provide cheap credit and insurance to credit-constrained households could have an indirect effect on human capital decisions. Financial development might reduce wealth inequality in access to education.

3.5 Data Appendix

This appendix describes the main data sources employed and the various calculations and adjustments which have been made for the econometric analysis. As I mention in the main text the data used in this chapter comes principally from three sources. The first source is the 1999 employment and unemployment round of the National Sample Survey (NSS). This is a household-level dataset, which contains among other things information on the main activity, age, and education of each individual in the household, and their last state of residency if they came from another state. In all

school completion regressions I employ the last state of residency as a proxy for the state of birth. The second data source is Indian state panel data, collected by Besley and Burgess (2000) from a wide variety of original sources. This data covers the main sixteen Indian states for the period 1960-1992. Haryana split from the state of Punjab and enters the sample in 1965. The third source is bank data collected by Burgess and Pande (2005) for the main sixteen Indian States for the period 1960-1992.

Bank Availability: Burgess and Pande (2005) collected bank branch data from the Reserve Bank of India (2000). Based on bank and population data I define a dummy variable B_{is} equal to 1 if each bank branch serves on average 15,000 people or less. Alternatively, I employ the continuous variable of number of people served per bank branch. For school completion regressions I construct a variable B_{ics} measuring the number of schooling years out of the maximum potential 5 years of primary school during which banks were widely available. So for individuals who were born in a state which never had widely available banks this number is equal to 0, and for individuals living in a state with widely available banks throughout the period this number is equal to 5. Otherwise, if z defines the first calendar year after banks became widely available, and b defines individual birth year, then $B_{ics} = (b+11) - z$ subject to maximum value of 5 and minimum value of 0.

Educational Variables: Based on the child's main activity I construct a school-attendance dummy equal to 1 if the child's main activity is school. Based on child age and the type of school he/she attends I define a school lagging variable. Specifically, if the child is 12 or older but still attends primary school, or if the child is 15 or older but still attends middle school, this variable is equal to 1. The number of schooling years is constructed in the same way as in chapter 2.

Economic Shocks: A sub-sample of households was revisited three months after the first visit, allowing me to construct household panel data. I use changes in household composition and economic activity between the two visits to construct variables measuring economic shocks. In particular, I define a variable indicating a new child, a variable indicating a departure from the household, and a variable indicating job loss by one or more of the parents.

The information on **MPCE** (monthly per capita expenditure in Rupees) is taken directly from the NSS employment and unemployment survey. The figures are based on a detailed questionnaire in the consumer expenditures survey, which is carried out simultaneously with the employment and unemployment survey. In general, the household characteristics are based on the characteristics of the parents of the children.

State-Level Variables include the size of the population, political variables, and economic and infrastructure variables. They are taken from Besley and Burgess (2000) and are described in more detail in the next chapter.

REFERENCES

Baland, Jean-Marie and James A. Robinson, (2000), "Is Child Labor Inefficient?", *Journal of Political Economy*, 108 (4), 663-79.

Basu, Kaushik, (1999), "Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standards", *Journal of Economic Literature*, 37 (3), 1083-119.

Basu, Priya and Pradeep Srivastava, (2005), "Scaling-up Microfinance for India's Rural Poor", World Bank Policy Research Working Paper 3646.

Beegle, Kathleen, Rajeev H. Dehejia and Roberta Gatti, (2003), "Child Labor, Income Shocks, and Access to Credit", World Bank Policy Research Working Paper 3075.

Becker, Gary S. and Nigel Tomes, (1986), Human Capital and the Rise and the Fall of Families", *Journal of Labor Economics*, 4(3(2)), S1-S39.

Besley, Timothy and Robin Burgess, (2000), "Land Reform, Poverty and Growth: Evidence from India", *Quarterly Journal of Economics,* 115 (2), 389-430.

Bhalotra, Sonia, (2000), "Is Child Labour Necessary?", STICERD working paper.

Bhalotra, Sonia, (2000b), "Child Farm Labour: Theory and Evidence", STICERD working paper.

Binswanger, Hans and Mark Rosenzweig, (1993), "Wealth, Weather and the Composition and Profitability of Agriculture Investments", *Economic Journal*, 103, 56-78.

Burgess, Robin and Rohini Pande, (2005), "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment", *American Economic Review*, 95(3), 780-795.

Burgess Robin, Rohini Pande and Grace Wong, (2005), "Banking for the Poor: Evidence from India", *Journal of the European Economic Association*, 3(2-3), 2005.

Cameron, Stephen V. and Christopher Taber, (2001), "Borrowing Constraints and the Returns to Schooling", NBER, Working Paper w7761.

Card, David, (1999), "The Causal Effect of Education on Earnings", in Handbook of Labor Economics edited by O. Ashenfelter, and D. Card, Vol. 3A. Amsterdam, North-Holland.

Card, David, (2001), "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems', *Econometrica*, 69(5), 1127-60.

Carneiro, Pedro and James J. Heckman, (2002), "The Evidence on Credit Constraints in Post-Secondary Schooling", Discussion Paper Series, Institute for the Study of Labor.

Carneiro, Pedro and James J. Heckman, (2003), "Human Capital Policy", NBER Working Paper, w9495.

Duraisamy, P. (2002), "Changes in Returns to Education in India, 1983–94: by Gender, Age-Cohort and Location", *Economics of Education Review*, 21(6), 609–622.

Edmonds, Eric V. (2005), "Child Labor and Schooling Responses to Anticipated Income in South Africa", forthcoming in the *Journal of Development Economics.*

Guarcello, Lorenzo, Fabrizia Mealli and Furio C. Rosati, (2002), "Household Vulnerability and Child Labor: The Effect of Shocks, Credit Rationing and Insurance", Manuscript, Research Center, UNICEF.

Jacoby, Hanan, (1994), "Borrowing Constraints and Progress Through School: Evidence from Peru", *The Review of Economics and Statistics*, 76, 151-160.

Jacoby, Hanan and Emanuel Skofias, (1997), "Risk Financial Markets and Human Capital in a Developing Country", *Review of economic Studies*, 64(3), 311-35.

Keane, P. Michael, Kenneth I. Wolpin, (2001), "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment", *International Economic Review*, 42(4), 1051-1103

Kingdon, Geeta G. (1998), "Does the Labour Market Explain Lower Female Schooling in India?", *Journal of Development Studies*, 35(1), 39-65.

Kingdon, Geeta G. and Jeemol Unni, (2001), "Education and Women's Labour Market Outcomes in India", *Education Economics*, 9(2), 173-195.

Morduch, Jonathan, (1995), "Income Smoothing and Consumption Smoothing", *Journal of Economic Perspectives*, 9 (3), 103-114.

Reserve Bank of India, (2000), "Directory of Commercial Bank Offices in India", Reserve Bank of India, Mumbai.

Vasudeva-Dutta, P. (2006), "The structure of wages in India, 1983-1999", forthcoming in *Education Economics.*

TABLE 3.1
DATA DESCRIPTION

	Rural	Urban
School Attendance	0.64	0.78
	(0.48)	(0.41)
Lagging at School	0.30	0.24
	(0.46)	(0.43)
Monthly Per Capita Expenditure (Rupees)	415.99	680.72
	(324.59)	(497.30)
Child Age (Years)	11.51	12.00
	(3.72)	(3.72)
Female Child	0.46	0.47
	(0.50)	(0.50)
Land Less Than 0.01 Hectares	0.19	0.43
	(0.39)	(0.49)
Land 0.01-1 Hectare	0.49	0.20
	(0.50)	(0.40)
Land More Than 1 Hectare	0.27	0.05
	(0.45)	(0.22)
Household Size (Number)	6.61	6.30
	(2.69)	(2.63)
Females in Household (Number)	3.24	3.08
	(1.70)	(1.67)
Scheduled Castes and Scheduled Tribes	0.69	0.53
	(0.46)	(0.50)
Muslim	0.13	0.19
	(0.33)	(0.39)
Christian	0.02	0.02
	(0.14)	(0.15)
Other Religions	0.03	0.03
	(0.17)	(0.18)
Mother Doesn't Work	0.60	0.82
	(0.49)	(0.39)
Mother Self-Employed (or family business)	0.21	0.08
	(0.40)	(0.27)
Mother Employed	0.19	0.10
	(0.39)	(0.30)
Father Doesn't Work	0.18	0.17
	(0.38)	(0.37)
Father Self-Employed (or family business)	0.45	0.37
	(0.50)	(0.48)
Father Employed	0.37	0.46
	(0.48)	(0.50)
Female Siblings Aged 0-5 (Number)	0.39	0.47
	(0.68)	(0.50)
Male Siblings Aged 0-5 (Number)	0.39	0.25
	(0.67)	(0.56)
Observations	38545	21806

Notes: standard deviations are in parentheses. All values are proportions unless other units are indicated. The omitted religious group is Hindus. The omitted land-ownership group is landless. See the Data Appendix for details on the construction and sources of the variables.

TABLE 3.2
BANK AVAILABILITY AND SCHOOLING STATISTICS

	Population Per Branch	School Attendance	School Rank
Punjab	9716	0.77	5
Kerala	9890	0.85	1
Karnataka	10922	0.69	8
Jammu & Kashmir	11040	0.78	4
Tamil Nadu	12811	0.76	6
Gujarat	12968	0.69	9
Haryana	13616	0.78	3
Andhra Pradesh	14598	0.66	10
Maharashtra	14867	0.80	2
Rajasthan	15557	0.65	12
Orissa	15783	0.63	15
Madhya Pradesh	16520	0.64	13
West Bengal	17175	0.64	14
Uttar Pradesh	17903	0.65	11
Bihar	19658	0.47	16
Assam	19771	0.75	7
All of India	15047	0.67	

Notes: School attendance refers to children aged 6-18 and is based on whether school is defined as the child's main activity.

TABLE 3.3
SCHOOL ATTENDANCE

	(1)	(2)	(3)	(4)	(5)
Age Group	All	All	6-11	12-14	15-18
25%-50% MPCE	**0.066252**				
	(4.75)* **				
50%-75% MPCE	**0.099**				
	(7.57)* **				
75%-100% MPCE	**0.143**				
	(13.94)* **				
MPCE (Impact Per Rupee)		**0.0003**	**0.000225**	**0.000284**	**0.000486**
		(13.49)* **	**(7.96)*** **	**(6.36)*** **	**(14.30)*** **
Mother is Literate or more	0.136	0.132	0.09	0.15	0.168
	(16.19)***	(14.84)***	(10.55)***	(10.34)***	(8.89)***
Literate Unschooled Father	0.086	0.088	0.073	0.084	0.058
	(7.86)***	(7.73)***	(6.66)***	(5.41)***	(2.53)**
Primary School Father	0.122	0.125	0.093	0.115	0.14
	(9.90)***	(9.85)***	(7.83)***	(9.57)***	(7.82)***
Middle School Father	0.156	0.158	0.109	0.157	0.197
	(15.78)***	(15.71)***	(14.33)***	(10.21)***	(8.12)***
Secondary School Father (or higher)	0.226	0.218	0.13	0.207	0.342
	(14.83)***	(15.41)***	(11.78)***	(9.12)***	(11.01)***
Female Siblings Aged 0-5	-0.032	-0.036	-0.022	-0.035	-0.032
	(4.21)***	(4.69)***	(4.01)***	(3.38)***	(2.04)**
Male Siblings Aged 0-5	-0.018	-0.02	-0.008	-0.023	-0.03
	(1.81)*	(1.98)**	(1.10)	(1.20)	(2.54)**
Mother Doesn't Work	0.022	0.026	0.019	0.055	0.041
	(1.76)*	(1.84)*	(1.28)	(2.72)***	(1.94)*
Mother Self-Employed	-0.001	0.003	0	0.017	0.017
	(0.04)	(0.14)	(0.01)	(0.70)	(0.82)
Father Doesn't Work	-0.103	-0.099	-0.058	-0.139	-0.103
	(5.96)***	(5.69)***	(2.80)***	(2.80)***	(3.51)***
Father Self-Employed	0.017	0.021	0.016	0.027	0.004
	(1.02)	(1.18)	(1.32)	(1.01)	(0.18)
Rural	-0.073	-0.022	-0.013	-0.046	-0.012
	(6.87)***	(2.36)**	(0.84)	(2.34)**	(0.63)
Female	-0.122	-0.124	-0.087	-0.158	-0.133
	(6.06)***	(6.16)***	(7.17)***	(6.23)***	(3.50)***
Scheduled Castes/Tribes	-0.053	-0.05	-0.037	-0.053	-0.045
	(7.46)***	(7.26)***	(4.28)***	(4.76)***	(3.49)***
Muslim	-0.127	-0.125	-0.076	-0.154	-0.148
	(10.75)***	(10.29)***	(9.06)***	(9.67)***	(7.72)***
Christian	0.042	0.038	0.018	0.012	0.123
	(1.39)	(1.29)	(0.57)	(0.25)	(2.23)**
State Fixed Effect	Yes	Yes	Yes	Yes	Yes
Observations	50242	49313	24692	11869	12752

Notes: * significant at 10%; ** significant at 5%, *** significant at 1%. The table reports the marginal effect from probit regressions, using monthly per capita expenditure (MPCE) as a proxy for income level. Absolute t-statistics calculated using standard errors clustered by state are reported in parentheses. Columns 1 and 2 include also school age dummy. All regressions include state fixed effect and are weighted. Variables covering households with more than one married child, number of females, number of males, being the eldest child, and age are also included in the regressions but are not reported due to lack of space.

TABLE 3.4
LAGGING AT SCHOOL

	(1)	(2)	(3)	(4)
Age Group	12 or 15	12	15	15
MPCE (Impact Per Rupee)	**-0.00008**	**-0.000047**	**-0.00014**	**-0.000069**
	(1.79)*	**(0.62)**	**(1.84)***	**(0.84)**
Interaction Rural and MPCE				**-0.000257**
				(1.80)*
Mother is Literate or more	-0.045	-0.066	-0.035	-0.03
	(1.93)*	(2.48)**	(0.96)	(0.78)
Literate Unschooled Father	-0.009	-0.052	0.082	0.083
	(0.22)	(1.17)	(1.27)	(1.30)
Primary School Father	-0.085	-0.062	-0.118	-0.117
	(2.82)***	(1.28)	(2.49)**	(2.43)**
Middle School Father	-0.136	-0.161	-0.087	-0.083
	(6.01)***	(3.99)***	(1.48)	(1.44)
Secondary School Father (or higher)	-0.203	-0.183	-0.237	-0.239
	(6.97)***	(3.82)***	(4.90)***	(4.86)***
Female Siblings Aged 0-5	0.064	0.077	0.038	0.031
	(4.28)***	(5.66)***	(1.18)	(0.98)
Male Siblings Aged 0-5	0.042	0.058	-0.021	-0.026
	(2.69)***	(3.46)***	(0.87)	(1.05)
Mother Doesn't Work	-0.049	-0.036	-0.067	-0.055
	(2.76)***	(1.36)	(1.11)	(0.92)
Mother Self-Employed	-0.03	-0.007	-0.064	-0.057
	(1.21)	(0.29)	(0.95)	(0.80)
Father Doesn't Work	0.086	0.082	0.105	0.112
	(1.14)	(1.28)	(0.93)	(1.01)
Father Self-Employed	-0.018	0.013	-0.07	-0.068
	(1.05)	(0.85)	(2.24)**	(2.13)**
Rural	0.022	0.026	0.015	0.151
	(1.00)	(1.07)	(0.36)	(2.36)**
Female	0.027	0.036	0.013	0.016
	(0.91)	(0.91)	(0.31)	(0.36)
Scheduled Castes/Tribes	0.018	0.007	0.042	0.044
	(0.95)	(0.27)	(1.38)	(1.43)
Muslim	0.07	0.112	0.015	0.021
	(2.98)***	(4.06)***	(0.29)	(0.42)
Christian	-0.028	-0.008	-0.079	-0.088
	(0.43)	(0.09)	(0.91)	(1.01)
State Fixed Effect	Yes	Yes	Yes	Yes
Observations	6072	3818	2239	2239

Notes: * significant at 10%; ** significant at 5%, *** significant at 1%. The table reports marginal effect from probit regressions, using monthly per capita expenditure (MPCE) as a proxy for income level. Absolute t-statistics calculated using standard errors clustered by state are reported in parentheses. Columns 1 and 2 include also school age dummy. Variables covering households with more than one married child, number of females, number of males, being the eldest child, and age are also included in the regressions but are not reported due to lack of space. All regressions include state fixed effect and are weighted.

TABLE 3.5
SCHOOL AND CREDIT AVAILABILITY

	A. Banks		B. Provident Funds	
Dependent Variable	Attending School	Lagging	Attending School	Lagging
Age	6-18	12 or 15	6-18	12 or 15
	(1)	(2)	(3)	(4)
MPCE (Impact Per Rupee)	0.00043	-0.00021	0.00023	-0.00012
	(6.65)***	(7.26)***	(9.87)***	(2.00)**
Interaction Term	**-0.00026**	**0.00023**	**-0.00013**	**0.00017**
	(3.38)* **	**(5.07)*** **	**(7.31)*** **	**(1.4)**
State Fixed Effect	Yes	Yes	Yes	Yes
Observations	43419	5167	46399	5728

Notes: * significant at 10%; ** significant at 5%, *** significant at 1%. The table reports the marginal effect from probit regressions, using monthly per capita expenditure (MPCE) as a proxy for income. Absolute t-statistics calculated using standard errors clustered by state are reported in parentheses. The interaction term is between MPCE and credit availability, measured in columns 1-2 employing bank availability and in columns 3-4 employing provident fund membership. Bank availability variable is a dummy equal to 1 if in the child's state of residence 15,000 people or less are served by one bank. Provident fund membership is a dummy equal to 1 if at least one of the child's parents is a member of a provident fund. Lagging regressions are for children aged 12 or 15 currently attending school. A child is said to lag behind at school if he/she attends a school at a lower level than the level expected by his/her age. All regressions include fixed effect and are weighted. The additional control variables are the same as in tables 3.4 and 3.3.

TABLE 3.6
SCHOOL ATTENDANCE AND INCOME SHOCKS

Age Group	All	6-11	12-14	15-18	All	6-11	12-14	15-18
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Baby	-0.022	-0.021	-0.008	-0.055	-0.021	-0.019	-0.003	-0.068
	(2.90)***	(2.48)**	(0.81)	(3.54)***	(2.83)***	(2.44)**	(0.32)	(4.03)***
Household Member Left	-0.029	-0.012	-0.046	-0.052	-0.026	-0.009	-0.046	-0.043
	(5.20)***	(1.77)*	(3.90)***	(4.16)***	(4.22)***	(1.38)	(3.24)***	(3.12)***
Job Loss	**-0.08**	**-0.068**	**-0.072**	**-0.136**	**-0.094**	**-0.068**	**-0.101**	**-0.202**
	(8.91)* **	**(6.10)*** **	**(5.29)*** **	**(7.99)*** **	**(8.19)*** **	**(4.56)*** **	**(6.07)*** **	**(9.46)*** **
Bank Interaction Term					**0.011**	**0.009**	**0.015**	**0.02**
					(1.39)	**(0.78)**	**(1.53)**	**(1.97)****
Observations	42704	23388	10823	8493	36635	20411	9238	6986

Notes: * significant at 10%; ** significant at 5%, *** significant at 1%. The table reports the marginal effect from probit regressions. Absolute t-statistics calculated using standard errors clustered by state are reported in parentheses. The interaction term is between job loss and credit availability. Bank availability variable is a dummy equal to 1 if in the child's state of residence 15,000 people or less are served by each bank. Job loss variable is a dummy equal to 1 if at least one of the parents lost their job between the two visits. All regressions are for children who attended school on the first visit. All regressions are weighted.

TABLE 3.7
EDUCATIONAL ACHIEVEMENTS AND BANK AVAILABILITY

	Completed Primary School					School Years				
Sample	All	All	All	Rural	Urban	All	All	All	Rural	Urban
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Bank Availability	**0.006082**	**0.006378**	**0.003593**	**0.008597**	**0.002675**	**0.053962**	**0.056464**	**0.041361**	**0.067218**	**0.04584**
	(1.75)*	**(2.26)****	**(1.54)**	**(2.35)****	**(1.01)**	**(1.70)**	**(1.78)***	**(2.11)***	**(2.04)***	**(1.29)**
Individual Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Economic Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Political Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
School controls	NO	YES	YES	NO	NO	NO	YES	YES	NO	NO
Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time Trend	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO
Observations	241405	205617	205617	145183	96222	241405	205617	205617	145183	96222
R-squared						0.29	0.29	0.29	0.24	0.2

Notes: * significant at 10%; ** significant at 5%, *** significant at 1%. Columns 1-5 report the marginal effect from probit regressions, and columns 6-10 report OLS coefficients. Absolute t-statistics calculated using standard errors clustered by state are reported in parentheses. Bank availability is the number of years during school age years in which 15,000 people or less were served by one bank branch. Economic controls include real state domestic product per capita, rural and urban population, scheduled caste population, and scheduled tribe population. Individual controls include dummies for gender, rural residency, scheduled castes and scheduled tribes, and dummies for various religions. Political variables include a dummy for the leading political group in the Lower House as defined in the Data Appendix of chapter 4, political margin, and voter turnout. All regressions include state and cohort fixed effects and are weighted.

Chapter 4

Bicameralism and Public Education Provision: Theory and Evidence from India.

4.1 Introduction

Understanding the effect of political institutions on public policy is of key interest to political economy. One political institution which is widespread but has received little attention in the literature is bicameralism.[1]India is a federal country which inherited a bicameral legislative system from British rule. Some Indian states are still bicameral today, while others have changed to unicameralism, creating a unique time-state variation in the existence of Upper Houses. In addition teachers are granted special representation in the Upper Houses that no other profession enjoys. Employing both household data and state panel data this chapter brings together the unique institutional variation and teacher representation to study their combined effect on public provision of education.

Education is primarily a state matter in India and as a result there is significant variation in educational achievements across the Indian Federation, for example female literacy varies from 20 percent in Rajasthan to 86 percent in Kerala. Understanding how political institutions influence education is an important step in explaining this variation in educational outcomes.

[1] Around one third of the countries in the world today have bicameral legislatures, including almost all federal countries.

Traditionally, public finance has assumed that governments provide an optimal allocation of public goods. The literature analysing political influence on public goods provision was pioneered by Stiglitz (1974) who highlighted the fact that the median voter equilibrium does not necessarily exist. While the median voter approach is valuable in analysing certain aspects of public goods provision, its main disadvantage is that it is an 'institution-free' idea, and by its nature does not provide an insight into comparative institutions. Recently economists have devoted more attention to the effect of political institutions on political and economic outcomes. In their pioneering work, Persson and Tabellini (2004) show in cross-country regressions that a parliamentary system generates more public goods and higher taxes than a presidential system. Besley and Case (2003) provide a comprehensive review of the empirical evidence from the United States. Exploiting the broadly uniform institutional setting in the USA this literature generates scope for isolating the real effect of political institutions. However, this research is limited by the extent of institutional variation which has occurred within the USA.

Literature specifically concerning the role and economic impact of Upper Houses is quite limited even though Upper Houses are widespread. It has been argued that bicameralism generally reduces corruption while in federal nations the main justification for the existence of Upper Houses is to represent states' individual interests.[2] Ansolabehere etc. (2002) examines in a theoretical paper the effect of unequal voting weights in the federal Upper House and deduces that unequal weights do not necessarily imply a bias against small-state in public goods provision. Money and Tsebelis (1997) find a U-shaped relationship between the size of the government and the number of times laws are transferred between the two houses of the French parliament. Testa (2003) studies the effect of Upper Houses combined with lobbying groups on corruption employing cross-country regressions. She finds that when the two houses are of different political composition bicameralism increases corruption, while if the two houses are of similar political composition bicameralism reduces corruption.

The key limitation for empirical research of institutions in general and bicameralism in particular is that there is limited variation within countries, so identification is based upon cross-country regressions. In this framework it is difficult to account for unobserved variables. In addition it is difficult to interpret the results due to incomparability of definition of variables, sampling methods and institutions. Qualitative differences between Upper Houses are rarely accounted for even though Upper Houses

[2] For a full discussion of bicameralism see Money and Tsebelis (1997).

vary significantly worldwide, both in terms of powers and in methods of selecting members. At one extreme are Upper Houses that are constitutionally coequal with Lower Houses,[3]in the middle are Upper Houses which are coequal subject to restrictions, such as on monetary bills,[4]and at the other extreme are Upper Houses which have an essentially advisory capacity or can only delay the passage of laws.[5]Membership selection can be by direct election, indirect election, nomination, or heredity.[6]

India provides a unique opportunity to examine the effect of bicameralism on public goods provision. The constitution defines the composition of Upper Houses (known at the state level as Legislative Councils), and gives the states the power to abolish their Upper Houses, or to create new ones, resulting in a unique time-state variation in the existence of comparable institutions. Secondly, while teachers in India are granted a special privilege to choose their own representatives in the Upper Houses, primary school teachers are excluded, as they are not entitled to take part in these elections.[7]This implies that there may be different effects of having an Upper House on primary schools and higher-level schools. I exploit this fact in my empirical estimation. This special privilege of teachers does not allow me to distinguish between the effect of Upper Houses and teachers' political representation, but does allow me to focus my research on a well-defined subset of public goods: education.

Employing both household data and state panel data for the sixteen major Indian states over the period 1960-1992, I examine the effect of Upper Houses combined with teacher representation on education. The empirical analysis relies on the time and state variation in the existence of Upper Houses. The main findings of my empirical section are that the existence of an Upper House combined with teacher representation benefits middle and secondary schools at the expense of primary schools, with a fall in primary school completion rates of 3 percentage points, and an increase of 3 percentage points in middle and secondary school completion rates.

This chapter employs a theory based on a bargaining process between Upper Houses and governments. The government knows that if it makes a budget allocation which

[3]For example Upper Houses in the United States, Italy and Switzerland.

[4]For example, Upper Houses in India, Belgium and South Africa.

[5]Such as the French Senate, the British House of Lords, and the Irish Senate.

[6]Generally speaking, there is a correlation between the mode of selecting members and the powers given to the Upper House: the more democratic the selection process the more power the Upper House will enjoy.

[7]See Kingdon and Muzammil (2000, 2003) for a systematic study of teacher involvement in politics in Uttar Pradesh in India, including the historical reasons for their involvement and its effect on educational legislation and appointment of teachers.

is too different from the preferred allocation of the Upper House, amendments will be proposed. There is a cost associated with disagreements between the Houses, so the government is to some extent willing to compromise its preference in advance.

The theoretical model shows that groups of teachers who have political representation in the Upper House of the Indian state legislature are more successful in attracting resources than unrepresented groups of teachers. The empirical section shows that these extra resources are partially responsible for improvements in school outcomes.

Extensive literature has investigated the effect of school inputs including teachers' salaries on student outcomes and found mixed results. According to Hanushek (2002) only 9% of the estimates in the USA for the level of teachers' education and 14% of the estimates for teacher–pupil ratios show positive and statistically-significant effects on student performance. Specifically regarding India, Kingdon and Teal (2005) found that teachers' salaries improve students' achievements in private schools but not in public schools, where most teachers have a permanent job and are not worried about losing their position.

While the evidence is inconclusive, it does not say that money and resources do not matter but that simply spending more money on schools does not necessarily improve outcomes, especially if the resource allocation within some of the schools is inefficient. Moreover my empirical results do not suggest that the improvements in educational outcomes are solely explained by smaller teacher-student and school-student ratios. Other mechanisms should be considered, for example, teachers' representation could attract more quality teachers into teaching. Alternatively, representation could also serve to provide better facilities for both students and teachers, which improve school attendance.

The rest of the chapter is organised as follows: section 2 discusses the theoretical framework; section 3 describes the relevant institutional background; section 4 reviews the econometric strategy and discusses the results; section 5 concludes.

4.2 Theoretical Framework

The aim of this section is to offer foundations for the empirical tests and a structured scheme for interpreting the empirical results. I present a theoretical model which illustrates how differences between the two legislative houses in their preferences over educational inputs are translated into policy decisions even in the absence of a power of veto. I compare the policy taken in states with Upper Houses and states without

Upper Houses. In doing so I generate a number of predictions about the influence of bicameralism on educational provision as well as providing a useful organisational device for discussing my empirical assessment.

The model explores only the official influence channel, whereby teachers electing members of the parliament can affect legislation and budget decisions. The legislative rules are modelled according to the legislative process in India but could easily be adopted to fit other countries. The empirical strategy focuses on identifying the effect on educational inputs, which are decided by monetary bills. Accordingly, the theoretical model focuses only on monetary bills. However it is straightforward to extend the model to non-monetary bills, on which the Upper House can have an even greater effect.[8]In addition, I do not rule out some alternative mechanisms of influence.

The model is based on a bargaining process between the government and the Upper House. I do not model the bargaining process between the coalition parties and take the government structure as given. The model focuses on the allocation of educational inputs between primary schools and higher levels of schools. I do not explicitly model the objective functions of each agent but since secondary school teachers and middle school teachers attached to secondary schools have representatives in the Upper House and not in the Lower House, I assume that the preferences of the Upper House are more in favour of these teachers.

Consider a model with two agents $i \in \{1, 2\}$ with $i = 1$ being the government and $i = 2$ being the Upper House. With a given educational budget each agent has a preferred allocation between primary schools and higher schools. Let $x = 1$ be the preferred point of the government, where educational achievements in all schools are equally cared for. Let $x = 0$ be the preferred point of the Upper House, where only middle and secondary school achievements are of concern to it.[9]Thus, in the absence of an Upper House the policy implemented is $x = 1$. The policy offer available to each agent is denoted by $x_i \in [0, 1]$. The utilities from a chosen policy x are $(x, 1 - x)$, representing the proportionate utilities to the government and the Upper House respectively. These given utilities correspond to the fact that each agent achieves higher utility as the chosen allocation is closer to its preferred point. Thus, the bargaining process focuses on choosing a point along the continuum between these two preferred allocations.

[8]For example, Kingdon and Muzammil (2003) find that in Uttar Pradesh teachers have affected non-monetary bills, such as laws regarding supervision of teachers.

[9]These assumptions are just for simplicity. All that really matters is that the Upper House has some bias in favour of middle and secondary schools.

The political process has three stages. In the first stage the government makes an offer x_1. In the second stage the Upper House can either accept the offer or propose an amendment x_2. If the Upper House accepts the offer the bargaining ends and the utilities are $(x_1, 1-x_1)$. Otherwise a third stage takes place where the Lower House votes on the amendment, and either accepts or rejects it. In the case of acceptance of the amendment, the resulting utilities are $(\delta_1 x_2, \delta_2(1-x_2))$, whereas if the government rejects the amendment the allocation of the budget is passed in its original form despite dissent by the Upper House, and the utilities are $(\delta_1^2 x_1, \delta_2^2(1-x_1))$. $\delta_i \in [0,1]$ are discount factors representing the cost to the government and to the Upper House of disagreement, either due to reputational damage or as a result of delaying the budget legislation.

It is easy to show that the limiting case where the Upper House does not suffer at all from a disagreement, i.e. where $\delta_2 = 1$, is not interesting as it always results in a disagreement. Since the government anticipates that the Upper House will not accept any policy except for its own preferred policy, the government does not make any effort to compromise. Intuitively, if the Upper House does not lose anything from suggesting amendments, it will always amend the budget to represent its preferred choice, and therefore the price of making it indifferent is too high. In the same way, it is also easy to show that if the government does not suffer at all from a disagreement, i.e. $\delta_1 = 1$, then the government always follows its preferred policy. This is because the government can simply reject any amendment and receive its original offer without cost. Hence, I assume that δ_1 and δ_2 are strictly less than 1.

I solve the model backwards. In the last stage of the bargaining process the government will logically accept any amendment x_2 if $\delta_1 x_2 \geqslant \delta_1^2 x_1$, therefore in the second stage if the Upper House chooses to make an amendment, its optimal value will be $x_2 = \delta_1 x_1$, in which case the utilities are $(\delta_1^2 x_1, \delta_2(1-\delta_1 x_1))$. Hence in the first stage the government knows that the Upper House is willing to accept any offer x_1 if $1 - x_1 \geqslant \delta_2(1-\delta_1 x_1)$. Thus if the government seeks a compromise, its optimal offer is $x_1 = \frac{1-\delta_2}{1-\delta_2\delta_1}$. The government prefers this compromise to simply ignoring the Upper House and forcing through its preferred value of 1, if $\frac{1-\delta_2}{1-\delta_2\delta_1} \geqslant \delta_1^2$. In other words agreement is reached if the benefit from an agreement is non-negative.[10]

Therefore the optimal offer $x^* = \begin{cases} 1 & \text{in the absence of an Upper House} \\ & \text{or if } \frac{1-\delta_2}{1-\delta_2\delta_1} < \delta_1^2 \text{ ;} \\ \frac{1-\delta_2}{1-\delta_2\delta_1} & \text{otherwise} \end{cases}$

[10] If $\delta_1 = \delta_2 = \delta$ the condition is fulfilled for $\delta \leq 0.75$.

Proposition 3 ***Upper House Effect***

For a sufficiently low and positive discount factor the existence of an Upper House generates a smaller x^* compared to policy in the absence of an Upper House.

Proof. In the presence of an Upper House with $\delta_2 > 0$[11]and $\frac{1-\delta_2}{1-\delta_2\delta_1} \geqslant \delta_1^2$, $x^* = \frac{1-\delta_2}{1-\delta_2\delta_1}$, which must be smaller than 1 because $\delta_1 < 1$. ■

This Proposition states that policies taken in states with Upper Houses are further away from governments' ideal points than policies taken in states without Upper Houses. Intuitively, the reasons for the result are clear. Where costs are associated with a parliamentary conflict the government is willing to compromise its preferences even in the absence of a power of veto.

To conclude the model generates an important implication for the empirical strategy. It shows that Upper Houses affect the legislation even in the absence of a power of veto. It also shows that if the Upper House favours secondary and middle schools, then its existence will serve to improve educational outcomes of middle and secondary schools at the expense of primary schools. In other words, I expect selective teacher representation in the Upper House to have a negative effect on primary schools and a positive effect on higher level schools. It is this prediction of the model which I will bring to the data. In particular, I will exploit the variation in the existence of Upper Houses, and will use different measures of educational policy and educational outcomes, e.g. number of school teachers, number of schools, and school completion rates, to compare policy taken in states with teacher representation in the Upper Houses with policy taken in states where the government does not need to be involved in political bargaining with an Upper House.

4.3 Institutional Background and Data

India provides a unique opportunity to examine the effect of bicameralism due to its state-time variation in the existence of comparable Upper Houses. In addition the composition of Upper Houses provides special representation to a very narrow group of people and hence reveals more concrete information about the direction of preference gaps between the Upper House and the Lower House. Moreover the exclusion

[11] At the extreme case when $\delta_2 = 0$, $x^* = 1$. In other words, if the Upper House discounts disagreement by so much that it would never dare make an amendment, then the government does not make any concession and chooses its preferred policies.

of primary school teachers generates contradicting predictions for different levels of schools, which I explore empirically.

4.3.1 Upper Houses

The introduction of an Upper House in India in 1921 was a product of British rule.[12]Upon independence a great discussion arose around the necessity of the bicameral system. Nevertheless, the new constitution of India established the Upper House as a federal chamber, representing the interests of the states in federal legislation. Article 168 of the constitution created a similar bicameral structure at the state level for various states of India, and article 169 of the constitution granted the states' Lower Houses the power to abolish Upper Houses or create one where none exists. Originally, the constitution of India provided for having bicameralism in six states.[13]By 1960 after several reorganisations and separations of states, eight states had an Upper House. West Bengal, Punjab, Andhra-Pradesh, and Tamil-Nadu subsequently abolished their Upper Houses in late 1969, early 1970, 1985, and 1986, respectively.[14]No state without an Upper House chose to create one.

In structuring the states' Upper Houses India followed the Irish model of granting representation to 'wiser groups'. In particular, India followed Ireland in granting graduates from higher education a special right to elect representatives to the Upper House, and in an exceptional step granted the same privilege to teachers. Teachers were the only profession singled out for this special privilege. Article 171 of the constitution lays down the composition of the Upper House. According to the constitution, local bodies elect one third of the Upper House members. The Lower House elects another third. The governor appoints a further sixth of the members. University graduates elect one twelfth of the members, and teachers elect the final twelfth. Each member of the Upper House is elected for six years and every two years a third of the members are replaced. Therefore elections take place every other year.

A person is entitled to vote on teachers' representatives in the Upper House provided that he is a resident of the state and has at some time within the previous six years been

[12]Tripathi (2002).

[13]Bihar, Bombay, Madras (later renamed Tamil-Nadu), Punjab, Uttar Pradesh and West Bengal. The basic rule was that states with Lower Houses with more than 160 members were also granted Upper Houses.

[14]For a discussion of the reasons for abolition see section 4.4.1. Graph 1 shows the variation over time.

engaged in teaching in educational institutions within the state,[15]which are deemed to be not lower in level than secondary school. For this purpose the government specified with the election commission all educational institutions that are deemed to be not lower than secondary school. In practice, the statement about secondary school teachers has been interpreted to include also middle school teachers if they are teaching in a middle school attached to a secondary school[16]. As a result, it is mainly primary school teachers who are excluded from the political process.

Upper Houses affect public goods provision mainly through monetary legislation. The constitution makes a distinction between monetary bills and non-monetary bills. Since the Upper House is not an elected body in the regular sense it is granted less powers over legislation than the Lower House, and these powers are significantly reduced in respect of monetary bills. Article 198 of the constitution "Special Procedure in Respect of Money Bills" restricts the introduction of monetary bills to the Lower House. After the bill has been passed by the Lower House it is transferred to the Upper House for its recommendations, and the Upper House has a period of fourteen days to return the bill to the Lower House with its recommendations. The Lower House may either accept or reject any or all of the recommendations of the Upper House. If the Lower House accepts any of the recommendations of the Upper House, the monetary bill is deemed to have been passed by both Houses. If the Lower House does not accept any of the recommendations of the Upper House, the monetary bill is deemed to have been passed by both Houses in the form in which it was passed by the Lower House without any of the amendments recommended by the Upper House. Hence the influence of the Upper House is generated by two characteristics of the legislative process: first, the amendment rights granted to the Upper House, and second, the inability of the government to change its initial budget once it has been passed over to the Upper House. Accordingly the source of power of the Upper House is the government's fear of reputational damage as a result of disagreement between the Houses. This fear causes the government to choose an allocation of public goods which is not too far from the preferred allocation of the Upper House.[17]

[15] The candidates are not required to be teachers.

[16] Kingdon and Muzammil (2003).

[17] An alternative source of political power, which I do not model, is that elements of the government coalition could side with the Upper House.

4.3.2 Data

I test the effect of bicameralism on the provision of education using both micro-level regressions and state-level regressions. Accordingly, I employ both household data and a state-panel dataset, combined with data on the existence of Upper Houses. The household data is the National Sample Survey (NSS) employment and unemployment round 55 conducted in 1999. After restricting the sample to the sixteen major states of India the data has around 150,000 observations for rural areas and 100,000 for urban areas. The state panel data was employed by Besley and Burgess (2000) and covers the sixteen major states of India over the period 1960-1992.[18]

For the micro-level regressions I employ the NSS data, retaining individual age and educational outcomes and all invariable individual characteristics such as gender, social group, and religion, and then merging the dataset with the state-panel data. In the merged data set individuals have their own characteristics but also all the relevant state information at the year of starting each school level, i.e. when the child was six years old for primary school, eleven years old for middle school and fourteen years old for secondary school. The NSS data does not include information about state of birth, but does include information about state of residency and migration. In cases of migration I take the previous state of residency as a proxy for state of birth.

I employ two different measures of Upper House existence. The first is a dummy variable equal to 1 if on the year starting school there was an Upper House in the state. As a second measure I define a variable that measures how many years out of the potential maximum years of studies at each school level the individual lived in a state which had an Upper House. For example, for an individual born in 1960 in West Bengal this number in primary school regressions is 4, since he started school in 1966, and the Upper House was cancelled at the end of 1969.[19]

In the second part of my analysis I employ state-level regressions to study the effect of bicameralism on educational inputs and on state-level educational achievements. Owing to concerns over biases in the enrolment data I construct school completion rates based on the NSS data. For each cohort I construct variables measuring the percentage of people completing primary school, middle school and secondary school, out of the people who were at the age of starting each of these schools, and then merge

[18] The Data Appendix describes the main data sources employed and various calculations and adjustments that have been made for the econometric analysis.

[19] For full details see the Data Appendix.

these cohort completion rates with the state panel data.[20]

Table 4.1 provides a statistical description of the data. Approximately 40% of the observations have an Upper House, with 4 states having an Upper House for all of the specified period and 4 states abolishing their Upper Houses at some point midway. The remaining 8 states did not have an Upper House at any time during the specified period. The data also reveals substantial variation in the political composition of various states, but in most of the states the Congress party is the dominant party. Educational inputs are relatively low. For example, the average number of teachers per 1000 children in the relevant age group is 16.42, 13.85 and 17.51 in primary, middle, and secondary schools respectively. The average number of schools per 1000 children in the relevant age group is 6.21, 1.82 and 1.03 for primary, middle, and secondary schooling respectively. Finally, the average real expenditure per child in the relevant age group is 48.9 Rs, 41.7 Rs and 87.4 Rs in primary, middle, and secondary schools respectively. The enrolment data, as pointed out in the past by Drèze (1998), is clearly biased upwardly with an average of 98% of boys reportedly being enrolled in primary school and a maximum enrolment level supposedly exceeding 100%.[21] School completion rates represent more accurately the true educational achievements in India, with 56%, 39% and 23% of the children completing primary, middle, and secondary schools respectively.

4.4 Empirical Estimation

4.4.1 Identification Strategy

The theoretical section generates a testable prediction regarding the effect of bicameralism and teacher representation on the provision of education in India. Bicameralism is expected to shift resources away from the allocation which would be preferred by the government in the absence of an Upper House. Thus political representation of middle and secondary school teachers in the Upper House is expected to shift resources toward the preferred allocation of these teachers.[22] In terms of the theory, the existence

[20] The primary school completion rate of people born in year x is merged with state data for the year x+6, the middle school completion rate of people born in year x is merged with state data for the year x+11, and the secondary school completion rate of people born in year x is merged with state data for the year x+14.

[21] A potential problem arises if there is a correlation between the bias in the enrolment data and the Upper House, for example if teachers over-report student enrolment as an alternative source of power in the absence of an Upper House.

[22] Even though university teachers enjoy similar political power I limit my analysis to school teachers because higher education comes under the responsibility of the federal government rather than the

of Upper Houses generates a smaller value x^*, which represents a closer allocation to the ideal public goods allocation of the Upper House. Accordingly, the empirical section employs various measures of educational achievements and educational inputs in different types of schools as the dependent variables.

The comparison in table 4.2 between states with and without an Upper House in the early Sixties reveals significant differences between the two. In general, states with a second chamber have more schools and teachers per 1000 children, and higher secondary school completion rates. However, the allocation of Upper Houses is not random. In fact according to the constitution only states with Lower Houses exceeding 160 members were assigned a bicameral parliament. Table 4.2 shows that states with Upper Houses also have significantly greater population sizes and higher GDP per capita. Therefore differences in educational inputs and educational outputs across those states cannot necessarily be attributed to the Upper House.

The source of variation comes from the fact that the Indian constitution allows states to abolish their Upper Houses. As a result there are three types of states: states which had Upper Houses during the entire period; states which never had Upper Houses; and finally states which abolished their Upper Houses. In what follows I refer to these three categories as 'bicameral states', 'unicameral states' and 'transformed states', respectively.

In the micro-level regressions it is possible to refer to a change in the parliamentary regime as generating a treatment. The combination of date of birth and state of residency determines who is exposed to the treatment. Young cohorts in transformed states form the treatment group. The estimation compares the difference in the probability of school completion between young cohorts and older cohorts in transformed states against untransformed states.[23]

In the state-level regressions I examine the change in educational achievements and educational inputs of untransformed states and transformed states before and after each transformed state abolished its Upper House. Over time we expect an increase in educational achievements and educational inputs in both types of states. However, the difference in differences between the rates of improvements can be interpreted as the effect of Upper House abolition. For example, a smaller improvement in the number of middle school teachers in transformed states may indicate that the rate of increase in number of middle school teachers has been slowed down as a result of abolition of

state governments.

[23] Untransformed states comprising unicameral states and bicameral states.

the Upper House.

A non-trivial assumption of the differences in differences method is that in the absence of abolition the patterns of education in both types of states should be the same. However, if the increase in the number of teachers is positively correlated with their initial numbers, and states which abolished their Upper Houses had lower educational achievements in the early Sixties, then we would observe the same pattern even without any effect from the Upper House.

Therefore an interesting feature of bicameralism in India is that an implication of this identification assumption can be tested. States which had the same parliamentary regime during the entire period, i.e. bicameral states and unicameral states, should not be affected by the abolition of Upper Houses elsewhere. Therefore education in these two categories would be expected to evolve comparably across the period. In the following section I test this identification assumption.

Another key assumption is that the abolition of Upper Houses is exogenous to educational achievements. Unfortunately, the reasons for the abolition of the Upper Houses are not completely clear as there was no discussion recorded of the abolition decision in the parliaments of Punjab or Andhra-Pradesh. From the discussions in West Bengal and Tamil-Nadu it appears that the main flaw in those Upper Houses was that political parties used them as a default job granted to candidates who did not succeed in elections to the Lower House. A second argument was that the states' Upper Houses had not served the purpose for which they had been created. Teachers' political privilege was mentioned among the reasons to abolish the Upper Houses but was not the primary reason.[24]Discussions in the Seventies between the federal government and state governments regarding the removal of teacher representation show that there is no clear correlation between states' opinions regarding teacher representation and their decisions regarding the abolition of their Upper Houses. For example, while Tamil-Nadu recommended to the federal government to retain teacher representation and even suggested to extend teacher representation to include primary school teachers, it abolished its Upper House. Meanwhile Bihar and Karnataka recommended the removal of teacher representation but chose to retain their Upper Houses. By contrast, Maharashtra recommended to keep teacher representation and retained its Upper House, and Andra Pradesh recommended the removal of teacher representation and cancelled its Upper House. So although econometric issues concerning non-random abolition of Upper Houses are still present it seems reasonable to assume that there

[24]Kingdon and Muzammil (2003).

is no reverse causality whereby educational achievements or educational inputs caused the abolition of Upper Houses.

A second concern is that there might be some underlying omitted factor, such as political change, driving both the abolition of Upper Houses and the allocation of educational inputs. In that case we would expect to find a correlation between bicameralism and various other government expenditures and policies, especially on the total education budget. However, this idea is not supported by the data.[25]Moreover, since fixed effects answer only the concerns over fixed unobserved state characteristics, each regression includes additional controls. In particular, the state domestic product and population separated into rural and urban areas account for the ability to invest in education and for the expected returns from education.[26]In addition, I use political variables to account for changes in preferences of the Lower House, which might be linked to other policies affecting education.[27]Finally I check the robustness of the results to inclusion of a state-specific linear time trend, to capture state-specific trends in attitudes toward education.

A general concern with the differences in differences method is that standard errors could be serially correlated. As Bertrand, Duflo and Mullainathan (2003) show, this would usually result in upwardly biased t-statistics, and in over-rejection of the null hypothesis. I employ the common solution and cluster the standard errors.[28]

4.4.2 Micro-Level Regressions

I begin by considering the effect of Upper Houses on educational outcomes. The theory section predicts an improvement in middle and secondary school achievements at the expense of primary schools.[29]To test this, I perform two types of regressions: individual

[25]I do not find an effect of bicameralism on government expenditures and taxes. The only correlation which I do find is between bicameralism and land reforms.

[26]Other controls such as roads, banks, and GINI coefficients for infrastructure, credit constraints, and inequality are also tried. While the results are robust the sample size is significantly reduced due to missing variables.

[27]The political variables include a dummy for election year and a dummy if the leading political group in the Lower House is one of four main groups of parties: Janata parties, Hindu parties, hard-left parties, and regional parties, as classified in Besley and Burgess (2002), where the base group is Congress parties.

[28]Concerns over serial correlation require clustering at the state level. However I present more conservative results clustering at the district level.

[29]On the other hand Kingdon and Muzammil (2003) claim that teachers' political power could have an adverse effect on teachers' efforts, due to lack of appropriate supervision and accountability. Therefore, while the net effect of teachers' representation in the Upper House on primary schools is expected to be negative the effect is ambiguous for middle and secondary schools.

data regressions and state panel data regressions.

Employing the NSS data I run the following linear probability regression for individual i from cohort c and state s:

$$h_{scij} = \alpha + S + C + \beta_j UH_{scij} + \lambda_j X_{scij} + \epsilon_{scij} \tag{4.1}$$

where h_{scij} is a dummy educational outcome variable, UH_{scij} is the variable of interest being either a dummy for the existence of an Upper House[30]or a variable measuring how many years each individual was exposed to an Upper House, X_{scij} is a vector of exogenous variables determining educational outcomes,[31] S is state fixed effect and C is cohort fixed effect. The regressions are run separately for 3 different levels of education, indicated by the subscript j. For primary schools $j = p$, for middle schools $j = m$ and for secondary schools $j = h$.

I expect differences in culture and attitudes toward education to be captured in the state fixed effects. In other words, the Upper House coefficient estimates the differences in differences of educational achievements between old individuals and young individuals in transformed states, against states which did not change the structure of parliament during the period.

As the theory predicts that bicameralism shifts the allocation of public goods in the direction of represented teachers, I expect to find contradicting effects on primary schools and higher level schools. For primary schools I expect bicameralism to be negatively associated with educational achievements ($\beta_p < 0$), whereas in the case of middle and secondary schools I expect bicameralism to be positively associated with educational achievements ($\beta_m, \beta_h > 0$).

Table 4.3 reports my key results on the effect of Upper House abolition on school completion employing a linear probability model. All the regressions include state and cohort fixed effects, child characteristics and state characteristics from the year the child was at the age of starting school j. The standard errors are clustered at the sampling unit.[32]Columns 1 and 2 show that in line with the theory, the effect of Upper Houses is to reduce the probability of completing primary school. Columns 3-6 apply the same regressions to middle and secondary schools. In line with the theory the

[30]Alternatively, we can interpret this variable as an interaction term between being old and living in a transformed state.

[31]Controls include individual controls such as gender, social group, rural residency, and other state controls from the year when each individual started school j.

[32]The district in most cases.

coefficients are now positive and significant. The even columns report the results using Upper House years as the variable of interest instead of a dummy variable. The two different measures of Upper House existence generate consistent results. For example, the coefficient in column 3 is three times as large as that in column 4 and the coefficient in column 5 is twice as large as that in column 6, corresponding to three years of middle school and two years of secondary school. Therefore from now on I report only results with an Upper House dummy. These results show that having an Upper House with teacher representation increases the probability of completing secondary and middle schools but reduces the probability of completing primary school, indicating that the political structure results in a real effect on outcomes.

If Upper Houses lead to higher drop-out rates at the primary school level then fewer people start middle and secondary schools. As a result the effect of Upper Houses on middle and secondary schools in columns 3-6 is under-estimated. The right-hand side of the table presents estimations conditional on completing the previous educational level. For example, the regressions on middle school completion ask what the determinants of middle school completion are for people who completed primary school. The downside of the conditional results is that if Upper Houses lead to higher drop-out rates at the primary school level then the better achievements in middle and secondary schools may reflect the selection of better students attending middle and secondary school. In fact, the conditional results are larger and more significant than the non-conditional results.

I next perform a robustness check to compare transformed states with bicameral states and unicameral states separately. Specifically, I run regression (4.1) twice more, the first time limited to individuals born in transformed states and bicameral states, and the second time limited to individuals born in transformed states and unicameral states. The Upper House coefficient estimates the differences in differences between bicameral states or unicameral states and transformed states before and after abolition.

Table 4.4 reports the results of this robustness check. The upper part of the table shows the difference in differences between transformed states and bicameral states. All definitions are the same as before, with all regressions weighted and standard errors clustered at the sampling unit. In line with previous results, the coefficients for primary schools are negative and significant and the results for secondary schools are positive and significant, but the results for middle school limited to a sub-sample of states are no longer significant. The middle part of the table reports similar results for the subset of unicameral states and transformed states.

The identification rests on the assumption that in the absence of abolition education

would have evolved similarly in both types of states. However, if improvement in educational achievements is negatively correlated with the initial level of education and transformed states had higher educational achievements in the Sixties, then we would observe the same pattern even without any effect from the Upper House.

To test the identification assumption I therefore run regression (4.1) one final time for individuals born in untransformed states (bicameral states or unicameral states). To estimate the differences in differences between bicameral states and unicameral states before and after abolition I replace the Upper House variable with a dummy variable equal to 1 if the state is a bicameral state and the year is after 1970, being the year when two of the transformed states abolished their Upper Houses.[33] If the identification assumption is valid then the difference in differences between transformed states and untransformed states is expected to be significant, but the difference in differences between bicameral states and unicameral states is expected to be insignificant.

The lower part of table 4.4 shows the differences in differences between bicameral states and unicameral states before and after 1970. In line with the identification assumption all of the coefficients are now insignificant. These results suggest that the differences in differences estimations are not caused by an inappropriate identification assumption.

I perform an additional robustness check including a state-specific time trend to account for a state-specific trend toward education. The results are robust to this specification and have similar coefficients. The results are presented in table 4.5. As the table shows the results are robust to specific time-trend specification and have similar coefficients.

To conclude, my results are consistent with the theory, and demonstrate that political structure affects educational outcomes. Upper Houses with special representation granted to middle school teachers and secondary school teachers improve educational achievements in those schools, to the disadvantage of primary schools which are excluded from the political process.

4.4.3 State-Level Regressions

For the state panel data regressions I run the following regression for state s at time t:

[33] West Bengal and Punjab abolished their Upper Houses at the end of 1969 and the beginning of 1970, respectively.

$$H_{stj} = \alpha + S + T + \beta_j UH_{stj} + \lambda_j X_{stj} + \epsilon_{stj} \quad (4.2)$$

where H_{stj} is the completion rate of school j, UH_{stj} is the key explanatory variable of bicameralism, X_{stj} is a vector of exogenous variables affecting educational achievements,[34] S is a state fixed effect, and T is a time fixed effect.

Table 4.6 reports the results for school completion rates and is provided as a test of whether bicameralism also has a substantial effect at the aggregated state level. Column 1 reveals a significant negative effect of Upper Houses on primary school completion rates. Specifically, an Upper House reduces primary school completion rates by between 3.4 and 3.9 percentage points. Column 3 reveals the opposite results for secondary schooling: Upper Houses increase the percentage of people completing secondary school by 2.9 percentage points. Column 4 reports the pooled regressions for all levels of school and indicates that Upper Houses create a bias against primary schools. Columns 5 and 6 provide the conditional results, with the dependent variables being the percentage of people completing middle school out of people who completed primary school and the percentage of people completing secondary school out of people who completed middle school. The conditional results are larger and more significant, but as explained earlier might reflect some selection bias. The lower part of table 4.6 reports the control group regressions and as before the coefficients are not significant, providing additional support that the differences in differences estimation is not driven by an unsuitable identification assumption. Taken together, the results demonstrate that teachers' political power in the Upper House affects educational achievements not only at the individual level but also at the state level.

Following the clear effect of political structure on educational achievements, I now try to identify the educational inputs which account for the differences. In particular, I ask whether bicameralism affects teachers' employment, the number of schools, and educational expenditures. I run equation (4.2) again, but this time with the number of teachers, number of schools and real expenditure per child as the dependent variables. All other definitions are unchanged except that some of the control variables are no longer relevant.

Table 4.7 reports the results. The even columns report results for number of schools per 1000 children and the odd columns report results for the number of teachers per

[34] including state domestic product, rural and urban population, scheduled caste population, scheduled tribe population, and political variables.

1000 children.[35]The results are somewhat weaker, but remain consistent with previous results. Column 2 reveals a negative effect of bicameralism on the number of primary schools. The existence of an Upper House reduces the number of primary schools per 1000 children by 0.4, which is around 6% of the average level in the sample. Columns 3 and 4 demonstrate the opposite effect for middle school inputs. Bicameralism increases the number of middle school teachers per 1000 children by 1.91, which accounts for more than 13% of the average level in the sample. The number of middle schools is similarly increased and accounts for just under 20% of the average level in the sample. The results for secondary school teachers are less significant but in the same direction. Section B of the table reports results from a robustness check and compares between bicameral states and transformed states. Section C tests the identification assumption and shows that the differences in differences between bicameral states and unicameral states before and after 1970 are not significant.

An interesting feature of the results is that I do not find any effect on total educational expenditure. In other words, my results show that Upper Houses have a significant effect on educational inputs consistent with the allocation of resources being shifted from schools with no political representation toward schools with political representation, but not a change of total resources.

4.5 Conclusion

Economists are far from clear about the role institutions play in the field of public policy. In particular, very little is known about bicameralism and public goods provision. Two unique features of the Indian legislature allow me to assess the effect of bicameralism on the provision of education. First, India chose to retain the bicameral system after independence both at the federal level and at the state level, but left the decision on whether to have a state Upper House for each individual state to make, resulting in a time-state variation in the existence of bicameralism. Second, teachers were granted a special representation in the Upper House in the belief that they are nobler than regular people and would provide a better judgment on political issues without grabbing rents for themselves. This allows me to generate predictions on a specific branch of public goods: education.

The chapter provides evidence that contrary to their intended purpose Upper

[35]Since the expenditure results are insignificant I do not report them.

Houses have served to shift the balance between different educational institutions. In fact, teachers have used their political power to transfer educational resources away from primary schools and toward higher levels of education. While this chapter cannot distinguish between teachers' sincere concern about education and rent-seeking motives, it does provide evidence that political power has a significant effect on resource allocation and real outcomes even in the absence of power of veto. In that sense these results are valuable to the debate regarding the existence of Upper Houses.

Moreover, while India is a unique example of granting political power to teachers, the lesson is applicable to many other countries, which also have an Upper House with members selected by elitist criteria. For example, the British House of Lords has been the subject of much recent debate. The controversy centres on its composition and whether hereditary peers and appointments are the preferred methods for selecting its members. This chapter highlights the potential risks of granting political power to selected groups. However it does not consider potential positive contributions such as experience and professional knowledge.

Taken together these results also allow for a better understanding of the significant variation in educational outcomes across different Indian states. For example, they show that Upper Houses reduce the primary school completion rate by 3 percentage points which accounts for more than 15% of the state variation in primary school completion.

4.6 Data Appendix

This appendix describes the main data sources employed and the various calculations and adjustments which have been made for the econometric analysis. As I mention in the main text the data used in this chapter comes principally from two sources. The first is the 1999 employment and unemployment round of the National Sample Survey (NSS). This is a household-level dataset, which contains among other things information on the age and education of each individual in the household and their last state of residency if they came from another state. In all of my calculations I take the last state of residency as a proxy for the state of birth. The second data source is Indian state panel data, collected by Besley and Burgess (2000) from a wide variety of original sources. This data covers the main sixteen Indian states for the period 1960-1992. Haryana split from the state of Punjab and enters the sample in 1965.

Upper House Variables: Based on Kingdon and Muzammil (2003) I construct a

dummy variable UH_j for each state and year with an Upper House. Using the National Sample Survey I construct an alternative variable UH_j for each individual older than 16,[36]which measures the number of schooling years out of the maximum potential years of school type j during which an Upper House existed. So for individuals who were born in a state which still has an Upper House this number in primary school regressions is equal to 5, whereas for individuals living in a state which never had an Upper House this number is equal to 0. Otherwise, if z defines the first calendar year after abolishing the Upper House, b defines individual birth year, m_j defines the potential number of years in school type j,[37]and g_j defines the age of starting school type j, then $UH_j = z-(b+g_j)$ subject to a maximum value of m_j and minimum value of 0.

Educational Variables: Data on educational variables is obtained from annual publications by the Department of Education in India: Education in India and Selected Educational Statistics. Number of schools per child and number of teachers per child are calculated employing the corresponding population information from the census.

If there are composite schools such as middle schools with primary classes, or secondary schools with middle classes, these have been included in the number of primary and middle schools respectively. For example, if in a town there are two primary schools and one middle school with primary classes, the number of primary schools in the town has been given as three and that of middle schools as one even though there are only three educational institutions. The same is the case with secondary or higher secondary schools.

The expenditure data is adjusted by a price index taken from Besley and Burgess (2002). This price index is a weighted average of the consumer price index for agricultural labourers (CPIAL) and the consumer price index for industrial labourers. Ozler, Datt and Ravallion (1996) have further corrected these price indices to take account of inter-state differences in the cost of living and have also adjusted CPIAL to take account of rising firewood prices.

I also use the NSS data to construct a measure of the percentage of people in each cohort graduating from primary, middle and secondary schools out of the total cohort.

Population Variables: The population estimates including scheduled caste populations and scheduled tribe populations are constructed using the Census data for the years 1961, 1971, 1981, and 1991. Between any of these years each sector of the

[36]The age of completing lower secondary school.

[37]5 for primary school, 3 for middle school and 2 for lower secondary school.

population is assumed to grow at a constant rate of growth.

Political Variables: Data on the number of seats won by each political party at each of the state elections is obtained from Butler, Lahiri and Roy (1991) "India Decides: Elections 1952-1991". Political margin is calculated as the gap in seats between the leading party and its nearest rival divided by the total number of seats. Following Besley and Burgess (2002) I construct five broad political groups: i) Congress Parties comprising the Indian National Congress, Indian National Congress Urs and Indian National Congress Socialist Parties; ii) Janata Parties comprising the Janata, Janata Dal and Lok Dal Parties; iii) Hard Left Parties comprising the Communist Party of India and Communist Party of India Marxist Parties; iv) Hindu Parties comprising the Bharatiya Janata Party; v) Regional parties comprising parties unique to each state.

Economic and Infrastructure Variables: The primary source of state income is an annual publication by the Department of Statistics. The data is adjusted by the price index. The primary source of bank branch data is from the Reserve Bank of India (2000), collected by Burgess and Pande (2003). Data on poverty is constructed by Ozler, Datt and Ravallion (1996) based on various rounds of the National Sample Survey.

REFERENCES

Agarwal, R. D. (1982), "Law of Education and Educational Institutions in India", Law Book Company, Allahabad.

Ansolabehere, Stephen, James M. Snyder and Michael M. Ting, (2002), "Bargaining in Bicameral Legislatures: When and Why Does Malapportionment Matter?", Department of Economics, MIT, Mimeo.

Beladi, S. H. (1988), "Theory and Practice of Parliamentary Procedure in India", Tripathi Private Ltd., Bombay.

Bertrand, Marianne, Ester Duflo and Sendhil Mullainathan, (2003), "How Much Should We Trust Differences-In-Differences Estimates?", Department of Economics, MIT, Mimeo.

Besley, Timothy and Robin Burgess, (2000), "Land Reform, Poverty and Growth: Evidence from India", *Quarterly Journal of Economics*, 115 (2), 389-430.

Besley, Timothy and Robin Burgess, (2002), "The Political Economy of Government Responsiveness: Theory and Evidence from India", *Quarterly Journal of Economics*, 117(4), 1415-1452.

Besley, Timothy and Anne Case, (2003), "Political Institutions and Policy Choices: Evidence from the United States", *Journal of Economic Literature*, 41(1), 7-73.

Besley, Timothy and Maitreesh Ghatak, (2004), "Public Goods and Economic Development", Working Paper, London School of Economics.

Burgess, Robin and Rohini Pande, (2005), "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment", *American Economic Review*, 95(3), 780-795.

Butler, David, Ashok K. Lahiri and Roy Prannoy, (1991), "India Decides :Elections 1952-1991", Aroon Purie, New Delhi.

Chawla, D.D. (1967), "Election: Law and Practice", Rahri Brothers, Delhi.

Drèze, Jean and Geeta G. Kingdon, (1998), "Biases in Education Statistics", The Hindu, 6th March, 1998, New Delhi.

Druckman, James N. and Michael F. Thies, (2002), "The Impact of Bicameralism on Government Formation and Duration", *American Journal of Political Science*, 46(4), 760-771.

Hanushek, Eric A. (2002), "Publicly Provided Education", in Handbook of Public Economics, Vol. 4, edited by A.J. Auerbach and M. Feldstein, Amsterdam, North Holland.

Heller, William B. (1997), "Bicameralism and Budget Deficits: The Effect of Parliamentary Structure on Government Spending", *Legislative Studies Quarterly*, 22(4), 485-516.

Kaul, M. N. and S. L. Shakdher, (1978), "Practice and Procedure of Parliament", Metropolitan Book Co., New Delhi.

Kingdon, Geeta G. and Mohd. Muzammil, (2000), "A Political Economy of Education in India: The Case of Uttar Pradesh", Department of Economics, Oxford, Mimeo.

Kingdon, Geeta G. and Mohd. Muzammil, (2003), "The Political Economy of Education in India: Teacher Politics in Uttar Pradesh", Oxford University Press, New Delhi.

Kingdon, Geeta G. and Francis Teal, (2005), "Does performance related pay for teachers improve student performance? Some evidence from India", Department of Economics, University of Oxford, Global Poverty Research Group, Working paper 14.

Money, Jeannette and George Tsebelisand, (1995), "Bicameralism Negotiations: The Navette System in France", *British Journal of Political Science*, 25(1), 101-129.

Money, Jeannette and George Tsebelisand, (1997), "Bicameralism", Cambridge University Press, New York.

Mughan, Anthony and Samuel C. Patterson, (2001), "Fundamentals of Institutional Design: The Functions and Powers of Parliamentary Second Chambers", *Journal of Legislative Studies*, 7(1), 39-60.

Mukherjea, A. R. (1983), "Parliamentary Procedure in India", Oxford University Press, Calcutta.

Ozler, Berk, Gaurav Datt and Martin Ravallion, (1996), "A Data Base on Poverty and Growth in India", World Bank, Mimeo.

Persson, Torsten, Gerard Roland and Guido Tabellini, (1997), "Separation of Powers and Political Accountability", *Quarterly Journal of Economics*, 112(4), 1163-1202.

Persson, Torsten, Gerard Roland and Guido Tabellini, (2000), "Comparative Politics and Public Finance", *Journal of Political Economy*, 108(6), 1121-1161.

Persson, Torsten and Guido Tabellini, (2004), "Constitutional Rules and Fiscal Policy Outcomes", *American Economic Review,* 94, 25-46.

Reserve Bank of India, (2000), "Directory of Commercial Bank Offices in India", Reserve Bank of India, Mumbai.

Rudolph, S. H. and L. I. Rudolph, (1972), "Education and Politics in India: Studies in Organization, Society, and Policy", Harvard University Press, Cambridge, MA.

Shakdher, S. L. (1992), "The Law and Practice of Election in India", National Publication House, New Delhi.

Stiglitz, Joseph E. (1974), "The Demand for Education in Public and Private School Systems", *Journal of Public Economics*, 3, 349-386.

Testa, Cecilia, (2003), "Government Corruption and Legislative Procedures: Is One Chamber Better than Two?", STICERD Development Economics Discussion Paper Series No. 41.

Tripathi, R. C. (2002), "Emergence of Second Chamber in India", National Publication House, New Delhi.

[redacted]

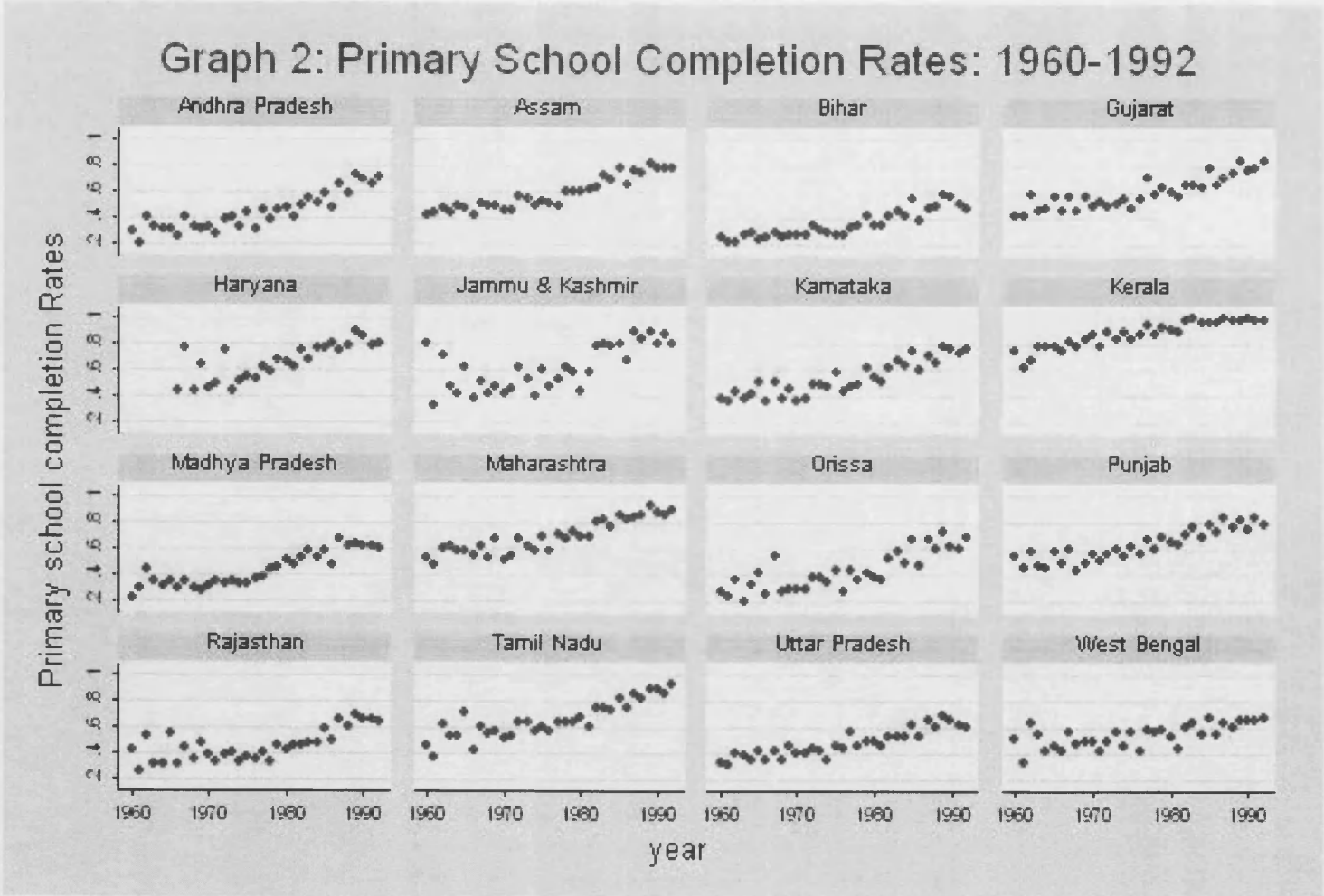
IA: 1960-1992

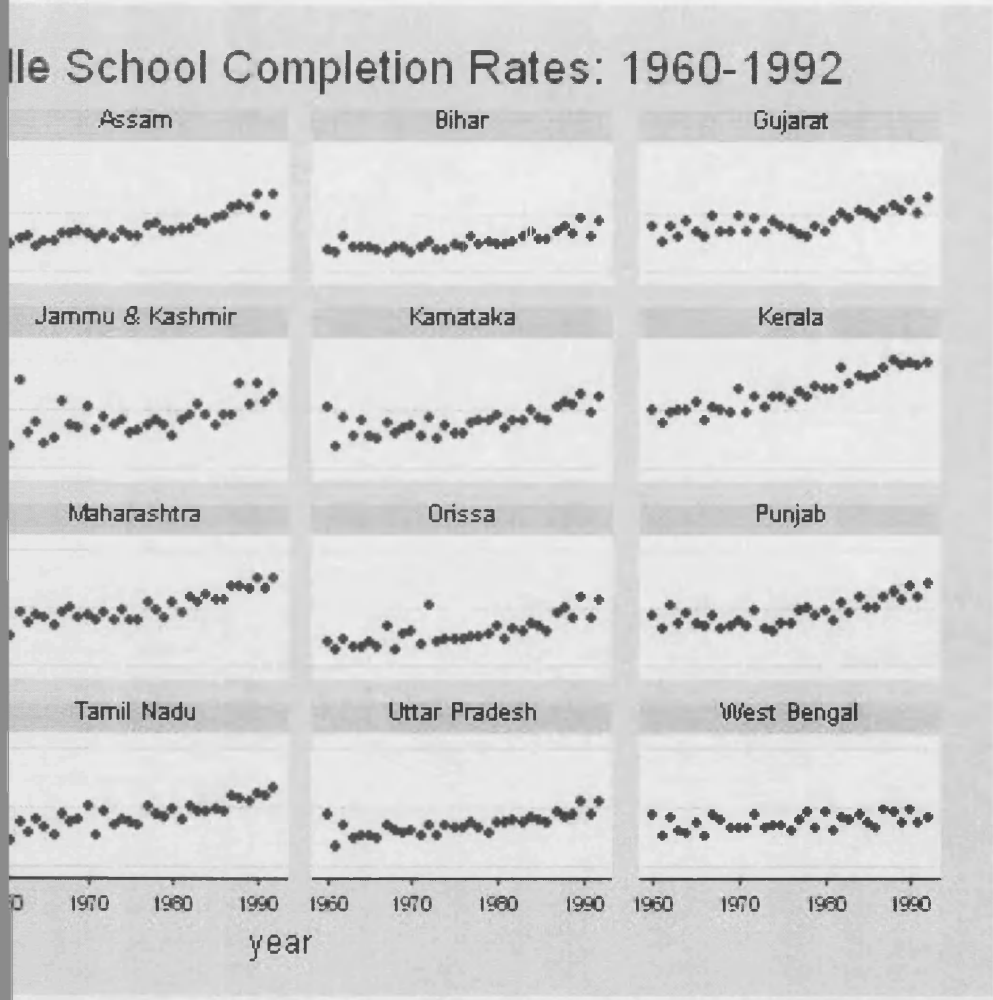
1966	1967	1968	1969	1970	1971
1	1	1	1	1	1
0	0	0	0	0	0
1	1	1	1	1	1
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
1	1	1	1	1	1
0	0	0	0	0	0
0	0	0	0	0	0
1	1	1	1	1	1
0	0	0	0	0	0
1	1	1	1	0	0
0	0	0	0	0	0
1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	1	0	0

1978	1979	1980	1981	1982	1983
1	1	1	1	1	1
0	0	0	0	0	0
1	1	1	1	1	1
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
1	1	1	1	1	1
0	0	0	0	0	0
0	0	0	0	0	0
1	1	1	1	1	1
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
1	1	1	1	1	1
1	1	1	1	1	1
0	0	0	0	0	0

	1984	1985	1986	1987	1988	1989	1990	1991	1992
Pradesh	1	1	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0
ı*	0	0	0	0	0	0	0	0	0
&Kashmir**	0	0	0	0	0	0	0	0	0
ka	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0
Pradesh	0	0	0	0	0	0	0	0	0
shtra	1	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
an	0	0	0	0	0	0	0	0	0
adu	1	1	1	0	0	0	0	0	0
adesh	1	1	1	1	1	1	1	1	1
engal	0	0	0	0	0	0	0	0	0

ed on Kingdon (2003) *Haryana split from the state of Punjab in 1965. ** Kashmir has its own constitution, with an Upper House which
t powers and different structure to other Upper Houses in Indian states, therefore I treat Kashmir as a unicameral state.


Graph 2: Primary School Completion Rates: 1960-1992
Andhra Pradesh
Assam
Bihar
Gujarat
Haryana
Jammu & Kashmir
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Orissa
Punjab
Rajasthan
Tamil Nadu
Uttar Pradesh
West Bengal
Primary school completion Rates
.2 .4 .6 .8 1
1960 1970 1980 1990
year


lle School Completion Rates: 1960-1992
Assam
Bihar
Gujarat
Jammu & Kashmir
Karnataka
Kerala
Maharashtra
Orissa
Punjab
Tamil Nadu
Uttar Pradesh
West Bengal
1960
1970
1980
1990
year

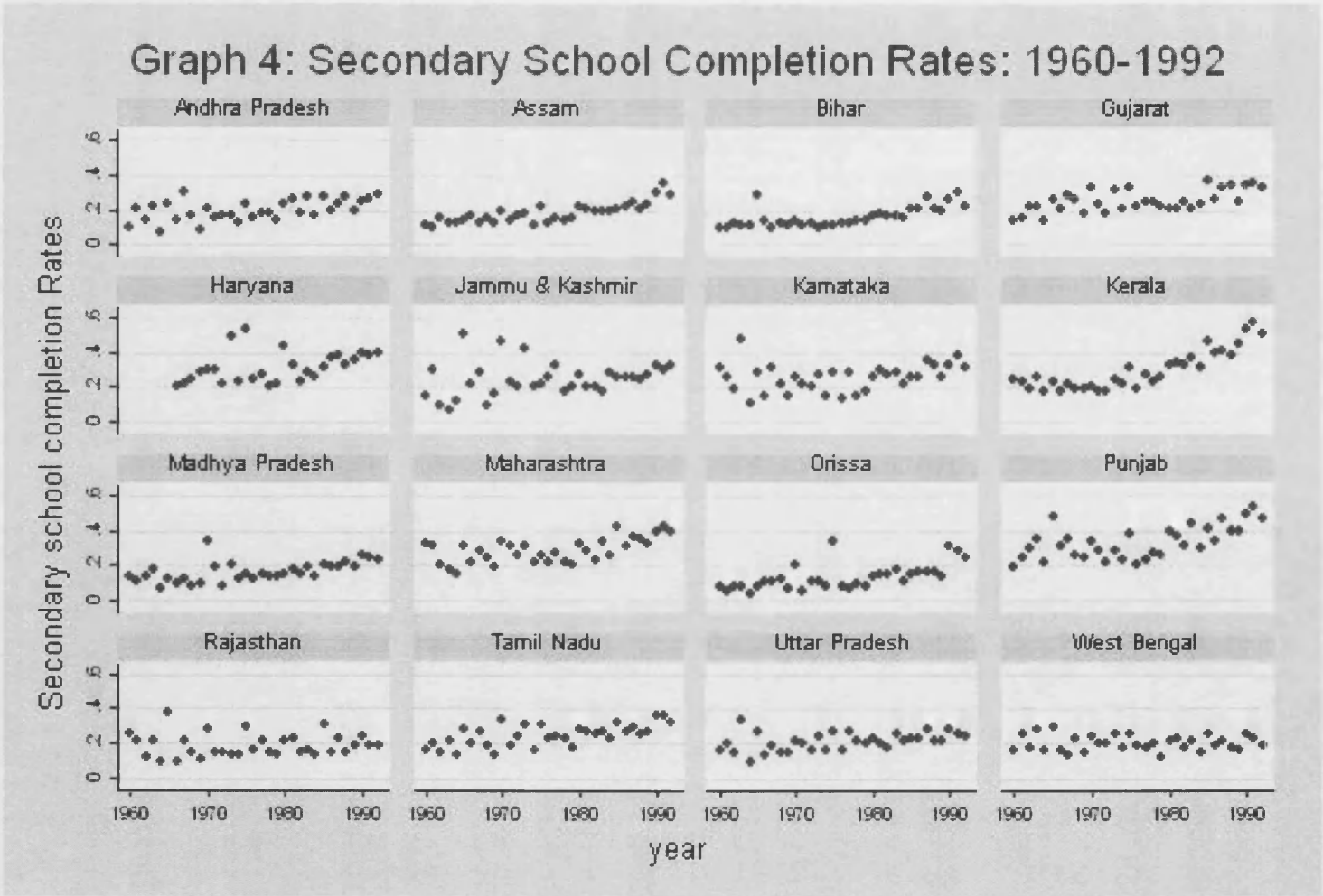

Graph 4: Secondary School Completion Rates: 1960-1992
Andhra Pradesh
Assam
Bihar
Gujarat
Haryana
Jammu & Kashmir
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Orissa
Punjab
Rajasthan
Tamil Nadu
Uttar Pradesh
West Bengal
Secondary school completion Rates
0
.2
.4
.6
1960
1970
1980
1990
year

TABLE 4.1
DATA DESCRIPTION

Variable	Mean	SD	SD Within States	Number of Observations
	(1)	(2)	(3)	(4)
Upper House Existence	0.39	0.49	0.21	491
	Enrolment Percentage			
Boys' Primary Schools (6-11)	97.98	16.77	13.03	380
Girls' Primary Schools (6-11)	67.20	24.89	14.84	380
Boys' Middle Schools (11-14)	52.76	16.53	11.33	395
Girls' Middle Schools (11-14)	27.68	19.11	11.14	395
Boys' Secondary Schools (14-17)	26.40	7.04	4.90	395
Girls' Secondary Schools (14-17)	11.67	7.96	4.08	393
	School Completion Rates			
Primary Schools	0.56	0.17	0.12	491
Middle Schools	0.39	0.14	0.10	491
Secondary Schools	0.23	0.09	0.07	491
	Real Expenditure per Child (Rupees)			
Primary Schools	48.90	25.55	20.34	328
Middle Schools	41.65	34.43	26.87	328
Secondary Schools	87.43	46.31	35.43	328
	Number of Schools per 1000 Children			
Primary Schools	6.21	4.00	1.13	427
Middle Schools	1.82	1.37	0.62	424
Secondary Schools	1.03	0.64	0.12	425
	Number of Teachers per 1000 Children			
Primary Schools	16.42	12.26	4.46	425
Middle Schools	13.85	8.67	4.46	425
Secondary Schools	17.51	12.17	7.35	426
	Controls			
Real State Domestic Product Per Capita (Rs)	1069.67	545.67	261.45	452
Rural Population ('000)	29774	20114	5818	491
Urban Population ('000)	8204	6095	2925	491
Scheduled Caste Population %	14.56	5.98	1.27	491
Scheduled Tribe Population %	7.32	7.51	0.997	491
Gini Coefficient %	29.62	3.79	2.10	491
Highways (Km.)	1724.33	744.11	281.66	352
Banks Per 1000 People	0.04	0.03	0.03	491
	Political Variables			
Congress %	0.50	0.25	0.21	489
Hard Left %	0.07	0.13	0.06	489
Janata %	0.14	0.24	0.19	489
Hindu %	0.04	0.08	0.06	489
Political Margin %	0.42	0.20	0.18	486
Turnout %	61.32	10.21	5.63	489

Notes: The data is for the sixteen main states over the period 1960-1991. Haryana split from the state of Punjab in 1965. From this date on, I include separate observations for Punjab and Haryana. I therefore have a total of 491 possible observations. Deviations from this are accounted for by missing data. See the Data Appendix for details on the construction and sources of the variables.

TABLE 4.2
DESCRIPTIVE STATISTICS BY PARLIAMENTARY REGIME (1960-1964)

Parliamentary System	Bicameral	Unicameral	Difference (1)-(2)
	(1)	(2)	(3)
Primary School Teachers	15.80	7.80	7.99***
	(1.07)	(0.69)	(1.27)
Middle School Teachers	10.45	6.05	4.4***
	(0.74)	(0.65)	(0.98)
Secondary School Teachers	11.68	5.45	6.23***
	(0.65)	(0.57)	(0.86)
Primary Schools	6.78	3.58	3.21***
	(0.40)	(0.34)	(0.52)
Middle Schools	1.40	0.79	0.61***
	(0.14)	(0.10)	(0.17)
Secondary Schools	0.70	0.31	0.39***
	(0.03)	(0.02)	(0.04)
Primary School Completion Rate	0.42	0.45	-0.03
	(0.02)	(0.03)	(0.03)
Middle School Completion Rate	0.32	0.30	0.03
	(0.02)	(0.02)	(0.03)
Secondary School Completion Rate	0.20	0.15	0.051***
	(0.01)	(0.01)	(0.02)
Population ('000)	38062	18238	19824***
	(2770)	(1430)	(3117)
Rural Population Proportion	0.79	0.85	-0.06***
	(0.01)	(0.01)	(0.02)
Real State Domestic Product Per Capita (Rupees)	886.43	779.17	107.26***
	(33.07)	(15.30)	(36.44)

Notes: * T-test is significant at 10%; ** T-test is significant at 5% ***T-test is significant at 1%. Standard deviations are in parentheses. Column 1 is for all bicameral states in 1960 including those who chose to abolish the Upper House in later years. Column 3 reports the differences between bicameral states and unicameral states. The data is for the sixteen main states over the period 1960-1964, excluding Haryana. See the Data Appendix for details on the construction and sources of the variables. Number of schools and number of teachers are per 1000 children.

TABLE 4.3
BICAMERALISM AND SCHOOLING: INDIVIDUAL DATA (All States)

	Unconditional Regressions						Conditional Regressions			
	Primary School		Middle School		Secondary School		Middle School		Secondary School	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Upper House	**-0.025**	**-0.007**	**0.021**	**0.007**	**0.024**	**0.012**	**0.032**	**0.009**	**0.037**	**0.021**
	(2.30)**	**(2.80)*****	**(2.65)*****	**(2.16)****	**(2.92)*****	**(2.71)*****	**(3.23)*****	**(2.45)****	**(2.50)****	**(2.59)*****
Economic Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Political Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Cohort and State FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	253394	253394	222332	222332	199164	199164	120917	120917	81937	81937
R-squared	0.25	0.25	0.25	0.25	0.2	0.2	0.25	0.25	0.21	0.21

Notes: * significant at 10%; ** significant at 5%, *** significant at 1%. The table reports the coefficients from linear probability regressions. Absolute t-statistics calculated using standard errors clustered by sampling units (districts) are reported in parentheses. See the Data Appendix for details on the construction and sources of the variables. The data is from the National Sample Survey 1999 for all individuals born between 1955 and 1984. Upper House variable in the even columns measures the number of years out of the maximum potential schooling years during which an Upper House existed. In odd columns, Upper House is a dummy equal to 1 if an Upper House existed when individual started school. The regressions in columns 7-8 are conditional on primary school completion and the regressions in columns 9-10 are conditional on middle school completion. Economic controls include real state domestic product per capita, rural and urban population, scheduled caste population and scheduled tribe population. Results are robust to the inclusion of Gini coefficient and banks per capita. Individual controls include dummies for gender, rural residency, scheduled castes and scheduled tribes, and dummies for various religions. Political variables include political margin, voter turnout, and a dummy for the leading political group in the Lower House as defined in the Data Appendix. All regressions include state and cohort fixed effects and are weighted.

TABLE 4.4
BICAMERALISM AND SCHOOLING: ROBUSTNESS CHECK

	Primary School	Middle School	Secondary School
	(1)	(2)	(3)
A. Comparison of Interest: Bicameral States and Transformed States			
Upper House	**-0.031**	**0.005**	**0.026**
	(2.60)* **	**(0.55)**	**(2.85)*** **
Observations	163286	144593	129506
R-squared	0.26	0.25	0.20
B. Comparison of Interest: Unicameral States and Transformed States			
Upper House	**-0.036**	**0.002**	**0.023**
	(3.15)* **	**(0.20)**	**(2.37)** **
Observations	155309	138187	124838
R-squared	0.26	0.25	0.20
C. Control Group: Bicameral States and Unicameral States			
Upper House × After 1970	**0.006**	**0.01**	**0.011**
	(0.58)	**(1.08)**	**(1.53)**
Observations	184501	163454	145984
R-squared	0.28	0.26	0.2

Notes: * significant at 10%; ** significant at 5%, *** significant at 1%. The table reports the differences in differences from linear probability regressions. Absolute t-statistics calculated using standard errors clustered by sampling units are reported in parentheses. See the Data Appendix for details on the construction and sources of the variables. The data is from the National Sample Survey 1999 for all individuals born between 1955 and 1984. Upper House variable is a dummy equal to 1 if an Upper House existed when individual started school. Economic controls include real state domestic product per capita, rural and urban population, scheduled caste population, and scheduled tribe population. Results are robust to the inclusion of Gini coefficient and banks per capita. Individual controls include dummies for gender, rural residency, scheduled castes and scheduled tribes, and dummies for various religions. Political variables include dummy for the leading political group in the Lower House as defined in the Data Appendix, political margin, and voter turnout. All regressions include state and cohort fixed effects and are weighted.

TABLE 4.5
BICAMERALISM AND SCHOOLING: TIME TREND

	Primary School	Middle School	Secondary School
	(1)	(2)	(3)
A. Comparison of Interest: All States			
Upper House	-0.024	0.028	0.021
	(2.33)**	(1.93)*	(1.94)*
Observations	253394	222332	199164
R-squared	0.27	0.25	0.2
B. Comparison of Interest: Bicameral States and Transformed States			
Upper House	-0.028	0.002	0.021
	(2.58)**	(0.17)	(2.00)**
Observations	163286	144593	129506
R-squared	0.26	0.25	0.2
C. Comparison of Interest: Unicameral States and Transformed States			
Upper House	-0.025	0.001	0.02
	(2.26)**	(0.08)	(1.87)*
Observations	155309	138187	124838
R-squared	0.26	0.25	0.2

Notes: * significant at 10%; ** significant at 5%, *** significant at 1%. The table reports the differences in differences from linear probability regressions. Absolute t-statistics calculated using standard errors clustered by sampling units are reported in parentheses. See the Data Appendix for details on the construction and sources of the variables. The data is from the National Sample Survey 1999 for all individuals born between 1955 and 1984. Upper House variable is a dummy equal to 1 if an Upper House existed when individual started school. Economic controls include real state domestic product per capita, rural and urban population, scheduled caste population, and scheduled tribe population. Results are robust to the inclusion of Gini coefficient and banks per capita. Individual controls include dummies for gender, rural residency, scheduled castes and scheduled tribes, and dummies for various religions. Political variables include dummy for the leading political group in the Lower House as defined in the Data Appendix, political margin, and voter turnout. All regressions include state and cohort fixed effects and state specific time trend and are weighted.

TABLE 4.6
SCHOOL COMPLETION RATES: STATE-LEVEL

	Primary School	Middle School	Secondary School	All Schools	Middle School	Secondary School
	(1)	(2)	(3)	(4)	(5)	(6)
A. Comparison of Interest: Bicameral States and Transformed States						
Upper House	**-0.034**	**0.018**	**0.029**	0.023	**0.031**	**0.029**
	(2.75)**	**(1.37)**	**(2.32)****	(1.38)	**(1.67)**	**(2.27)***
Upper House × Primary School				**-0.068**		
				(2.65)**		
Observations	244	244	244	732	244	244
R-squared	0.96	0.92	0.83	0.77	0.64	0.85
P-value of test UH+UH×Primary School=0				0.026		
B. Comparison of Interest: Unicameral States and Transformed States						
Upper House	**-0.039**	**0.009**	**0.03**	0.02	**0.024**	**0.023**
	(2.42)**	**(0.75)**	**(3.19)*****	(1.11)	**(2.35)****	**(2.37)****
Upper House × Primary School				**-0.051**		
				(1.14)		
Observations	353	353	353	1059	353	353
R-squared	0.94	0.88	0.79	0.75	0.39	0.81
P-value of test UH+UH×Primary School=0				0.301		

TABLE 4.6 CONTINUE

	Primary School	Middle School	Secondary School	All Schools	Middle School	Secondary School
	(1)	(2)	(3)	(4)	(5)	(6)
C. Control States: Bicameral States and Unicameral States						
Upper House × After 1970	**0.024**	**-0.005**	**-0.026**	-0.009	**-0.012**	**-0.02**
	(1.33)	**(0.24)**	**(1.43)**	(0.48)	**(0.64)**	**(1.16)**
Upper House × After 1970 × Primary School				**0.02**		
				(0.92)		
Observations	368	368	368	1104	368	368
R-squared	0.94	0.90	0.76	0.75	0.41	0.80
P-value of test UH×After 1970+UH×After 1970×Primary School=0				0.434		

Notes: * significant at 10%; ** significant at 5% ***significant at 1%. The table reports the differences in differences from linear probability regressions. Absolute t-statistics are calculated using clustered standard errors. School completion rate is calculated using NSS round 55. Controls are as in table 4.4. Upper House is a dummy variable for states and years with an Upper House. Upper House interaction with primary school is a dummy for primary school observations in states and years with an Upper House. All regressions include state and time fixed effects and are weighted. Columns 5-6 are conditional on primary school completion and on middle school completion, respectively.

TABLE 4.7
BICAMERALISM AND EDUCATIONAL INPUTS

	Primary School		Middle School		Secondary School	
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Teachers	Schools	Teachers	Schools	Teachers	Schools
A. Comparison of Interest: Transformed and Untransformed States						
Upper House	**-0.355**	**-0.405**	**1.912**	**0.341**	**2.383**	**0.019**
	(0.42)	**(2.87)****	**(3.30)*****	**(2.93)****	**(1.15)**	**(0.33)**
Observations	411	413	411	410	412	411
R-squared	0.97	0.98	0.98	0.96	0.97	0.97
B. Robustness Check: Bicameral States and Transformed States						
Upper House	**-0.526**	**-0.581**	**1.54**	**0.251**	**2.629**	**-0.061**
	(0.52)	**(1.90)***	**(3.50)*****	**(1.42)**	**(2.36)***	**(0.43)**
Observations	219	219	219	220	218	218
R-squared	0.98	0.98	0.94	0.99	0.97	0.97
C. Control States: Bicameral States and Unicameral States						
Upper House × After 1970	**-0.52**	**0.095**	**2.165**	**-1.162**	**-0.126**	**0.089**
	(0.16)	**(0.07)**	**(0.73)**	**(1.13)**	**(0.70)**	**(0.88)**
Observations	318	318	319	320	319	318
R-squared	0.97	0.98	0.97	0.98	0.97	0.96

Notes: * significant at 10%; ** significant at 5% ***significant at 1%. The table reports the differences in differences from linear probability regressions. Absolute t-statistics are calculated using clustered standard errors. See the Data Appendix for details on the construction and sources of the variables. The dependent variable in the odd columns is teachers per 1000 children and in the even columns is schools per 1000 children. The data is for the sixteen main states including Jammu and Kashmir for the period 1960-1991. For robustness purposes I checked the results also excluding Jammu and Kashmir. Haryana split from the state of Punjab in 1965. From this date on, I include separate observations for Punjab and Haryana. I therefore have a total of 491 potential observations, of which 248 would be applicable to section B, and 367 to section C, but missing data reduces the sample sizes to those noted. Political parties' variables include dummy for the leading political groups in the Lower House, voter turnout, and a dummy for election year. Political margin is defined as the gap in percentage of seats in the Lower House between the leading party and its nearest rival. Upper House is a dummy variable for states and years with an Upper House. All regressions include state and time fixed effects.

Chapter 5

Conclusion

This thesis analyses inefficiencies in India's education system. It finds evidence that market failures and political considerations affect investment in education by households and government.

Chapter 2 assesses the influence of local education on individuals. The evidence strongly supports the existence of human capital externalities in India, implying that living in a region of higher average educational outcomes is highly beneficial to individuals. In line with theoretical predictions, the results turn out to be much more pronounced when they refer to a smaller local area, so the effect of local educational achievements is seen to diminish with size. The results also reveal that externalities are larger in non-primary industries, where education plays a more significant role in production than in primary industries.

I have employed the differences in differences method and paid special attention to characteristics which are potentially correlated with my measures of local educational achievements. Although I have not found a reliable instrument variable to account for endogeneity, if endogeneity and unobservable variables were the only driving force behind the results I should have found no difference between the primary and non-primary sectors, and it is difficult to think of omitted variables that could be correlated with expenditures of people working in non-primary industries but not correlated with expenditures of people working in primary industries.

The chapter presents a theoretical model to explain the influence of externalities on wages. The model demonstrates the differences between the empirical predictions of general equilibrium models and the predictions of externalities models. In doing so the chapter offers a contribution to the empirical literature assessing human capital

externalities. As the general equilibrium effect would cause skilled labour to be disadvantaged by an increase in total education, the size of the externality is likely to be even higher than the observed effect.

Most existing research on externalities studies the influence of education in countries where literacy rates and primary school completion rates are close to 100%, and the only significant variation is in high school and college education. By focusing on India I take advantage of substantial variation at low levels of education, and the magnitude of my results is considerably greater than those found in previous papers focusing on the USA.

From my results it seems that households in India do not fully internalise the social benefits of acquiring education. This implies that the government should be more active in reducing the direct and indirect costs of schooling in India. However, in generating policy recommendations it is important to consider a more complete cost-benefit analysis. For example it is essential to establish the level of education which is most important in generating social benefits. Future research is also required to measure the effect of average education on non-market outcomes, such as voting rates, crime and fertility decisions.

Chapter 3 analyses the effect of credit constraints on educational outcomes. The findings demonstrate a substantial effect of credit constraints on school attendance and on wealth inequality in educational outcomes. It also shows that temporary income shocks affect education negatively but that widely-available credit can mitigate their negative consequences. Finally the chapter reveals that credit constraints harm school completion rates.

These findings have important policy implications for India as well as for other developing countries. The first implication is that it is possible to promote educational outcomes by improving other markets in India such as credit markets. Deciding whether bank expansion should be implemented requires, of course, a full cost-benefit analysis and a comparison with other alternatives. But these results show that when examining the cost-benefits of credit market intervention the broader social affects of credit should be taken into account. Extending the research to study the effect of credit on other outcomes such as health and nutrition should be extremely interesting in this context. The second implication is that income inequality could have a long-term effect on development and on the persistence of inequality when credit markets are constrained. Thus an important mechanism to promote social mobility and development is to provide access to credit and education.

Existing research focusing on developed countries finds it difficult to identify any evidence of credit constraints affecting education, despite the fact that the nature of investment in education makes it difficult to obtain loans to finance education even in developed countries. One reason is the existing subsidies and loans provided in those countries. A second reason is that most research focuses on college education, and at this level it is very difficult to isolate unobserved ability from income. Focusing on India allows me to employ bank availability as a relatively exogenous source of credit availability. It also allows me to concentrate on low levels of education where ability is less likely to bias the results.

However, this chapter has a few limitations. Bank availability is likely to be relatively exogenous but is also likely to be correlated with unobserved state variables. While the chapter offers a valuable improvement compared to the existing literature and does its best to account for other state features, it remains possible that these results are affected by unobserved variables.

Chapter 4 examines how a special representation of middle and secondary school teachers in the Indian state Upper Houses influences the provision of education in India. The main finding is that Upper Houses are associated with an increase in middle and secondary school completion rates and a corresponding drop in primary school completion rates. This is driven to some extent by a shift in resource allocation, where teacher representation increases teachers' employment in represented schools and reduces employment in unrepresented schools.

These results reveal that political representation plays an important role in determining government actions. It is evident that teachers have used their power to affect resource allocation, but it remains unclear whether this was driven by rent-seeking motives or purely toward achievements in their schools. In addition, the chapter does not allow complete separation between the effect of teacher representation and the existence of an Upper House, although this concern is mitigated by showing that there are no effects on other public goods or other aspects of the economy. In spite of these limitations, the conclusion that political power matters is itself important, and gives support to a range of theoretical models and political arrangements.

The findings reported here are also important for political economy research because they empirically examine the role of a widespread political institution which is not fully understood. The effect of political institutions, and Upper Houses in particular, is usually difficult to address convincingly amid other reasons due to endogeneity and unobserved variables. This chapter employs the differences in differences method and

takes advantage of India's unique time-state variation in the existence of Upper Houses to overcome potential endogeneity and biases from omitted variables. The findings also provide some explanation for the substantial state-level variation in educational achievements in India. Furthermore, they offer a partial explanation for the weakness in the primary school sector compared with higher levels of education in India.

While India is a unique example of granting political power to teachers, I believe the lesson to be applicable to other countries, which also have Upper Houses with members selected through elitist criteria. The chapter highlights the potential risks of granting political power to selected groups in general, a lesson which is applicable to a wider set of countries.

Future research can implement the outlined methodology in other contexts. In particular it would be interesting to examine the effect of Upper Houses on corruption and on the length of the legislation process. In the context of teacher representation it could also be interesting to study the effect on absence of teachers and legislation regarding education.

Overall, the thesis provides a clear picture of under-investment in education in India. It demonstrates that private investment in education is unlikely to be optimal. High exposure to risk, especially in rural areas, without an appropriate mechanism to mitigate this risk, means that households sacrifice future income in favour of current child labour. Large externalities imply that the socially optimum level of education is higher than the existing one. Finally, government investment is affected by political considerations and is unlikely to be optimal. It seems that at least in some states secondary and middle schools receive relatively too many resources compared with primary schools.

To achieve the Millennium Development Goal of universal primary education, more attention should be paid to household and government incentives. In particular, it is possible that by improving other markets such as credit markets and labour markets, educational achievements would also improve. From the evidence presented above it seems that households do not fully internalise the benefits from education, so more effort should be made by the government to reduce the direct and indirect costs of schooling borne by households. Some attention should also be devoted to institutional design, including government and teacher power.