ESSAYS ON FACETS OF THE EDUCATION SECTOR IN DEVELOPING COUNTRIES

Thesis Submitted to Jadavpur University

For the degree of

Doctor of Philosophy (Arts)

By

Sayoree Gooptu

Department of Economics

Jadavpur University

Kolkata- 700032, India

2023

Certified that the Thesis entitled

Essays on Facets of the Education Sector in Developing Countries submitted by me for the award of the Degree of Doctor of Philosophy in Arts at Jadavpur University is based upon my work carried out under the supervision of Dr. Vivekananda Mukherjee, Professor, Department of Economics, Jadavpur University and that neither this thesis nor any part of it has been submitted before for any degree or diploma anywhere/elsewhere.

…… ……

Countersigned by the Supervisor
Prof. Vivekananda Mukherjee
Dept. of Economics
Jadavpur University
Dated:

Candidate
Dated:

Dedicated to my maternal grandmother, my Didan,

for inspiring me to be a good human being.

Acknowledgements

My long journey of PhD would never have been possible without the support and encouragement of several people.

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Vivekananda Mukherjee for his constant guidance throughout my PhD work. His invaluable advice and feedback have enhanced my thought process and helped me improvise my research. I am grateful to him for providing me the freedom to explore new research ideas and keeping patience with my progress at every stage. His encouragement and support have helped me to get through the ups and downs and complete writing this thesis. Working under him for the last few years has not only enriched me academically, but has also taught me to cope with difficult situations.

I would also like to extend my thanks to Professor Achin Chakraborty, for providing constructive feedback on my thesis during the Research Advisory Committee meetings, which has helped me refine my work.

I am extremely grateful to the faculty members of the Department of Economics, Jadavpur University, who have taught me since my undergraduate days. They have played a pivotal role all throughout and encouraged me to pursue research. I must thank the Department for providing with excellent infrastructural facilities, which has facilitated my research. The departmental seminars, workshops and conferences have given me the forum for exchange of ideas and helped me grow academically. The non-teaching staff were also extremely kind throughout my journey as a student and scholar in the department.

I would like to thank the University Grants Commission for providing me with the fellowship under NET Junior Research Fellowship and NET Senior Research Fellowship during the course of my PhD research work.

The feedback and suggestions I have received from the conference and seminar participations at Indira Gandhi Institute of Development Research, Jawaharlal Nehru University, Presidency University, Azim Premji University, Indian Statistical Institute, New Delhi and BITS Pilani University, Hyderabad were of immense help. I would like to thank all the scholars and participants at these gatherings for the insightful discussions. The first chapter of this thesis has benefitted from extensive comments from the anonymous referees of the journal Review of Development Economics. The suggestions received from Professor Shubhashis Gangopadhyay and Professor Joydeep Bhattacharya have helped me provide a concrete shape to the first chapter of my thesis. The sound conceptual advice received from Mr. Purnendu Kishore Banerjee on data-cleaning and management has been extremely helpful. I am immensely thankful to Dr. Soham Sahoo for minutely explaining to me the nuances of empirical work and for being patient with all my queries. The suggestions received from Dr. Sanmitra Ghosh and Professor Dipankar Coondoo have also been extremely useful. The third chapter has benefitted from the helpful comments received from Dr. Bheemeshwar Reddy and Dr. Durgesh Chandra Pathak at the seminar presentation at BITS Pilani Hyderabad. I express my gratitude to Dr. Anirban Mukherjee, for the beneficial feedback received during my promotion interview for NET Senior Research Fellowship. I am also grateful to Professor Saumen Chattopadhyay for his comments during my MPhil viva-voce examination.

The discussions with esteemed researchers, who visited our department have been extremely beneficial. I am deeply indebted to late Professor Heinrich Ursprung whose extensive comments helped me in writing the first chapter. I also express my gratitude to Professor Arye Hillman, for motivating me with my academic pursuits.

I also wish to thank all my mentors who have inspired me to strive for better. In this regard, I would especially like to thank Dr. Sahana Roy Chowdhury, for always encouraging me to live up to my potentials.

I am thankful to my all my colleagues at Jadavpur University who have been a part of my journey as a research scholar, and provided a good peer environment. My seniors at the department were also very supportive and ready to provide advice at every juncture of my PhD.

Lastly, I would like to thank the people without whom, it would have been impossible. My parents, Kakali Gooptu and Soumitra Gooptu, have been the very first teachers in my life, without whom, I wouldn't have been where I am today. They have always been very supportive and appreciative of my work. I specially thank my mother for providing me with the comfort and support at home, to complete my research without any disruptions. My sister, Shirsha has always been my biggest strength, my confidante, who has been with me through everything and has kindled hope in me even during the toughest times. This journey would never have been complete without all three of them and their utmost trust in me. Their perpetual encouragement and support made it seem a little less arduous. I would like to make a special mention to my grandfather, Balarko Gupta, for his unflinching belief in my potentials and for being a constant source of assurance.

Sayoree Gooptu
Department of Economics
Jadavpur university
Kolkata

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

The present thesis delves into different issues related to education sectors in developing countries. The extensive impact of education across diverse spheres of a society is manifested in terms of increasing labour productivity, reduced inequality, lower poverty rates, better health outcomes, greater social mobility etc. Besides, the huge positive externality associated with education entails a spillover effect in the economy. As advocated by the proponents of the human capital theory, education serves as a driver of the growth and development of a nation as it enhances the skills and economic capabilities of individuals. Thus, investments in education are crucial for fostering the demographic dividend, which contributes to growth if employed productively. Despite being well versed in the benefits associated with education, most developing countries are often found grappling with multifaceted issues such as school dropouts, overeducation, where individuals work in jobs with lower educational requirements, mushrooming of private tuitions stemming from poor school teaching etc. The thesis focuses on each of these issues separately. First, it explores how a reallocation of the government education budget might cater to both school dropouts and overeducation. Second, it investigates the role of private tuition costs in determining the choice of schooling among Indian households at secondary and higher secondary levels. Third, it provides a detailed analysis of overeducation in different elementary occupations in India, where education requirements are typically low.

School dropouts constitute one of the most important issues plaguing education sectors in developing countries as they hinder the process of human capital accumulation, resulting in productivity loss. Despite the universality of such a fact, it is quite discouraging to evidence high rates of school dropouts in developing countries worldwide. According to the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2019) there existed 258.4 million out-of-school children in the world in 2018, among which the majority

belonged to upper-secondary grades. The scenario was the worst in Sub-Saharan Africa (97.5 million) and South Asia (93 million). In India, after the enactment of the Right to Education (RTE) Act by the government, which aimed at providing free elementary education for children in the 6-14 age group, though almost every child gets enrolled in school at the initial school going age, there still exists a large section of students dropping out before completion. Although the gross enrolment ratio at the primary level in India is as high as 103.4 percent, 1.5 per cent, 3 per cent and 12.6 per cent of the students have dropped out at primary, upper primary and secondary level in the year 2021-22[1].

At the other extreme, there exists an overeducation problem, which stems from an excess supply of educated individuals given the availability of job opportunities. According to the All India Survey on Higher Education (AISHE, 2018; 2021), there has been an increasing trend in the gross enrolment ratio in higher education over the last couple of years (23 per cent in 2013-14 to 27.3 per cent in 2020-21). This reveals the increasing preference for higher education among households[2]. However, this is accompanied by a high rate of unemployment among youth of 15 per cent in the 15-29 age group in India according to the Periodic Labour Force Survey (PLFS, 2021) for the year 2019-20[3]. Contingent on grim job prospects and an accentuating demand for higher education among households, individuals are often forced to work in occupations requiring lesser skills or education than what they acquire[4]. Overeducation is a particular and one of the most discussed types of skill mismatch in the labour market.

[1] Unified District Information on Education Plus (UDISE Plus, 2022).
[2] There has been an increase in higher education enrolment of around 16 percent in India between 2016 and 2021. The Compound annual Growth Rate (CAGR) for higher education in the same period is 3.8 percent. Also, between 2016 and 2021, there has been a 28 percent increase in the number of universities and 9.4 percent rise in the number of colleges respectively (AISHE, 2021). Enrolment in all higher education degrees (except MPhil) has increased over the same period.
[3] The PLFS 2019-20 data also shows that according to the principal status, the rates of unemployment are higher among the more educated. The corresponding rates for higher secondary and graduate individuals are 10.35 and 17.4 percent respectively. It is also seen that 18.6 percent and 4.4 percent of higher secondary and higher-educated individuals in the workforce in the 15-59 age group works as regular/casual workers in elementary occupations where education requirements are very low.
[4] In the job competition model proposed by Thurow (1975), individuals are keen on investing more in education as a defensive mechanism to ensure a higher position in the job queue.

According to McGuinness (2006), there is a potential welfare loss associated with overeducation since it involves wasteful expenditure on individuals without optimal utilization of the generated skills. Thus, simply promoting the expansion of higher education without generating productive employment results in workers landing in jobs with lesser education requirements, as overeducated workers. In this context, the skewed spending towards higher education in developing countries or the 'tertiary tilt[5]' discussed by Gruber and Cosack (2014) might be adding fuel to the overeducation scenario in a country.

However disparate the two phenomena – school dropout and overeducation might appear, they coexist in developing economies and require simultaneous attention for policy formulations. Either of these outcomes arises out of household decision-making in their child's education, which involves substantial investment. The government being one of the principal financiers of education in developing countries (UNESCO, 2016), plays a role in resolving discrepancies that might arise out of these households' education decisions. Such a complementarity between household and institutional investment provided by the government was first discussed by Majumdar (1983), which implies that a certain level of investment on the part of households is required in education parallelly to infrastructural investments like setting up schools and colleges, providing for classrooms, appointing teachers etc. In India, the combined expenditure of the Centre and states in education has increased consistently over the years from Rs 3.5 lakh crore to Rs 5.7 lakh crore between 2014-15 and 2020-21[6]. However, the percentage of such expenditure in the GDP still hovers around a meagre 3 per cent. The allocation (actual estimates) for higher education in the total education budget by the Ministry of education, Government of India, has increased from 32.43 per cent in 2011-12 to 41.73 per

[5] The 'tertiary tilt' is discussed by Gruber and Kosack (2014) and is evidenced with data from World Development Indicators provided by the World Bank. In South Asian countries, in 2015, government expenditure per student as percentage of GDP per capita at primary, secondary and tertiary level was 9.2, 10.5 and 29.4 respectively, against the corresponding figures of 20.0, 22.7, 27.7 in the OECD countries.

[6] Economic Survey(2021-22, 2022-23)

cent in 2021-22 (PRS, 2022; 2023). The corresponding compounded annual growth rate for the annual expenditure during the same period has been 6 percent. While the long-term implications of higher education on growth and development might justify an increasing proportion of spending on the same, it must not occur at the expense of school education quality. This holds especially in developing countries where education attainment and quality are very low (Su, 2004). In India, the Annual Status of Education Report (ASER) provide grim evidence on learning outcomes. According to ASER (2019) report for 2018, only 44.2 per cent and 22.7 per cent of children in the fifth grade of government schools can read a standard 2-level text and do divisions respectively[7]. Poor learning outcomes driven by inferior school quality are the primary reasons for children dropping out of school (Hanushek et al., 2006). However, it must also be remembered that just as important infrastructural investment is, for improving school quality (Glewwe et al, 2014), it depends to a large extent on the quality of teachers in schools who are products of the higher education system. Siphoning away funds from higher to school education raises the school quality and reduces higher education quality. The improvement in school quality incentivises households to complete their child's schooling. This lowers school dropouts. On the other hand, a fall in higher education quality lowers excess demand for higher education, resulting in lower overeducation rates. However, since the fall in the overall higher education quality affects the quality of teachers, it might end up perversely affecting school quality, causing school dropouts to rise. This shows that indeed these two problems are interlinked. In the first chapter of the thesis, we discuss the government's role in catering to both the school dropout and overeducation issues through a reallocation of the education budget, given the complementarities between household and institutional investment on one hand and school and higher education quality on the other.

[7] The National Achievement Survey (NAS, 2021), provides information on the learning achievements of students in classes 3,5,8 and 10 in India. It shows that majority of the states performed below the national average in the year 2021.

The significance of teachers in the education system is closely tied to the quality of education and student outcomes. However, formal education is often infiltrated with a plethora of issues related to poor teaching quality in developing countries. In this context, one can discuss the parallel education system popularly referred to as *private tuition* or *shadow education*, which has grown rapidly, catering to the existing deficiencies in formal education (Bray, 1999; Dang and Rogers, 2008; Ghosh and Bray, 2020) There are several factors which adds on to the demand for private tuitions such as high stakes examinations, student competition, peer pressure etc. Apart from all these, one of the most important reasons for opting for private tuition is poor teaching quality[8]. Given the existence of a private tuition market, households can opt for supplemental education for their children outside formal schooling. Household choice of their child's schooling depends on a host of factors. According to the NSS 75th Round on Education 2017-18 (NSS, 2019), among secondary and higher secondary students going to private unaided schools, 40 per cent reported the reason to be poor quality of nearby government schools, 17 per cent reported the reason to be the availability of special facilities[9] and around 17.5 per cent reported their preference for English medium of instruction. Despite such arguments in support of private schools, affordability is an important force driving the demand for their government counterparts. Given the differences in school-specific characteristics and teaching qualities between the two types of schooling, it follows that the presence of a private tuition market can mar the teaching quality gap between the two and enable households to select the more feasible option. Thus, private tuition can play an important role in affecting school choice. In the second chapter of the thesis, we explore the impact of private tuition costs on the choice of schooling, given the relative differences in

[8] The literature on teacher quality in India reveals a large extent of teacher absenteeism especially in government schools (Kremer et al., 2005; Muralidharan et al, 2017; Duflo et al, 2012a).

[9] See NSS 75th Round on Education 2017-18 Report No. 585 (NSS, 2020). It shows that these special facilities include air-conditioned classrooms, state-of-the-art teaching aids, day boarding facility, transport facility, hostel facility, laboratory/library, school timing, extra-curricular activities, separate toilet facilities for boys and girls, co-educational structure.

teaching quality between private and government schools and other school-specific attributes driving school choice preferences. The empirical results are also explained with the help of a theoretical model which provides an insight into the teaching quality differences between private and government schools.

Given the linkage between the education sector and labour markets, it is also important to study the implications that education decisions might have on labour market outcomes. As discussed earlier, the prevalence of overeducation in the labour market is indeed a matter of efficiency concern given the huge expenditure that occurs on account of equipping individuals with unproductive education. Overeducation in the US labour market was first brought to notice by Freeman (1976) whereby, the excess supply of higher educated graduates resulted in a substantial lowering of wages, giving rise to a section of overeducated workers. The discussion on overeducation was carried forward by Duncan and Hoffman (1981), Verdugo and Verdugo (1989), Rumberger (1981), who empirically estimated the extent of overeducation and their impact on wages. However, the conventional methods used in the literature for calculating overeducation are subject to biases. For instance, in the commonly used empirical method/ realised matches method, the mean/modal level of education is treated as a benchmark. Any worker with excess education is considered to be overeducated. This might lead to biased estimates. Suppose that there is a high supply of educated workers in a particular region, compared to the demand. This might drive up the average years of education in that occupation, deeming workers in the same occupations in other regions, where absorption rate is high, to be undereducated. Thus, despite low returns to education in the former region with poor demand conditions, this method undermines the extent of overeducation. It fails to account for the demand side heterogeneities that might exist at regional levels. A more recent approach was introduced by Mehta et al. (2011) for calculating overeducation in four developing countries (Mexico, Philippines, Bangladesh and India) using returns to education,

which catered to both the demand and the supply side of the labour market. In this thesis, we investigate the existence of overeducation in elementary occupations in the Indian labour market using the returns to education approach. According to the official requirements given by the National Classification of Occupations (NCO 2004), jobs within elementary occupation require low skill with up to ten years of formal education requirement. However, the PLFS data confirms the presence of tertiary educated workers in these occupations. Given the heterogeneity in the nature of work, we explore whether all these workers are actually overeducated or not, within different disaggregated categories of elementary occupations. The chapter also calculates the overall overeducation in an occupation, that might consist of lower education levels getting insignificant returns over illiterate workers and compares with the conventional modal method of estimating overeducation.

In the next section, we review the relevant literature.

1.2 Literature Review

Ever since the advent of the human capital theory in the literature, (Schultz, 1961; 1963; Becker, 1964, Mincer, 1974), the concept of investment in individuals has gained traction. Given the immense impact that education has on growth and development (Lucas, 1988; Romer, 1990), the proponents of this theory have argued how human capital embodied in individuals generates productive outcomes necessary for both short and long term gains (Chattopadhyay, 2012). Besides, education has a huge social benefit, which is useful for the society in both current and future generations (McMahon, 2004). The discussion on investment in education was extended by Majumdar (1983), which additionally pointed out the complementarity between household and institutional investment, path dependence in educational investment and interdependence between school and higher education quality. In Majumdar's (1983) approach, education achievement depends not only on the child's ability,

but also the household investment and institutional investment for providing infrastructural facilities such as building up of schools and colleges alongside recruitment of teachers. Quality of schooling is a crucial factor determining quality of higher education, as reflected in the education production functions (Arcalean and Schiopu (2009); Su (2004)). This is true given that basic education forms the foundation for future achievement and returns. The reverse impact of higher education on school quality through the quality of teaching though non-trivial, gets less attention in the literature. There is indeed evidence of a huge impact that teachers have on the quality of schooling of a child. Azam and Kingdon (2014), Chetty et al. (2014), Slater et al. (2009) shows a positive significant impact of teachers on learning outcomes using the teacher value added method. Kingdon (2006) shows a positive impact of teacher's education on student performance. However, the school education quality functions presented in the literature rarely incorporate this additional component induced by higher education. For instance, in a theoretical framework, Gilpin and Kaganovich (2009) models the quality of teachers to be function of the overall higher education quality, which in turn also affects school quality.

Given the importance of education investments, it is essential to study household investment decisions in their child's education. There exists a vast literature that looks into the sequential decision making of households in education investments for their children. Galor and Zeira (1993) develops an overlapping generation model which consists of utility maximizing households who decide on their children's education level depending on their income distributions and credit constraints. However, the model does not differentiate on the basis of child's ability while taking decisions. In linking child labour with inequality of income, Ranjan (2001) discusses the household optimization of utility for two periods where households choose between sending their child to the unskilled labour market and sending him to school. While studying the effects of school quality on students' choices of pursuing higher

education, Castello-Climent and Hidalgo-Cabrillana (2012) also looks into household decision making in education in a dynamic general equilibrium model. However, in their study, given that primary education is fully funded by the government and higher education only requires additional household spending, it does not accommodate the complementarity between government and household investments as pointed out by Majumdar (1983). Using a sequential education decision making structure, Abbott et al. (2013) studies the equilibrium impact of government provision for college financial aids. However, it does not study the impact of a government budget reallocation on outcomes.

The importance of government redistribution of education funding is explored by several papers. Su (2004) shows that for an economy in the early stage of development, it is necessary for the government to allocate a major share of funds on basic education, for achieving both equity and efficiency. However, Su (2004) abstracts away from private investment in education. Hidalgo-Hidalgo and Iturbe-Ormaetxe (2005) explores the impact of reallocation of education budget between basic and higher education on equity and efficiency. They reveal through a theoretical model that a transfer of resources from higher to basic education positively impacts the average productivity across the population, as well as reduces the minimum income required to attain higher education. However, though this holds for a developed country, the paper questions the policy's effectiveness in a developing country with very low higher education enrolment. Arcalean and Schiopu (2009) develops a theoretical model which considers the interdependence between private and government investments and studies the effect of a reallocation of the government education budget on growth. The two important takeaways from their model are that a higher share of the education budget must be devoted towards basic education irrespective of the size of the budget and that a rise in tertiary enrolment does not necessarily lead to a higher growth. However, none of these papers studying the government's role in education outcomes incorporates the interdependence between school

and higher education outcome as discussed by Gilpin and Kaganovich (2009). Also, despite these studies focusing on broader perspectives like growth, efficiency and equity, the literature does not analyse the impact of a government budget reallocation on specific issues such as school dropouts and overeducation.

Schooling systems in developing countries are ridden with high rates of dropouts. Using panel data for Andhra Pradesh, a state in India, Nakajima et al. (2018) shows that the quality of education is one of the most crucial indicators of dropouts at the upper primary and upper secondary levels of education. Hirakawa and Taniguchi (2020) studies primary school dropouts in Cambodia and shows that school level factors such as teacher absence, interactions with teachers and test achievements significantly affects school dropouts. Siddhu (2011), Gibbs and Heaton (2013) explain different factors affecting dropouts during the transition from primary to secondary, which accounts for a major share of school dropouts in in developing countries. Acknowledging the positive effect of secondary education on economic growth, Mussida et al. (2018) analyses the impact of extending compulsory education to secondary level in ten developing countries across the world[10]. Their paper finds that increasing secondary school dropouts reduces the possibility of employment in non-elementary occupations.

Just as pervasive the problem of school dropouts characterised by low educational attainment is, so is the problem of overeducation, which concerns the excess supply of higher educated individuals forced to work in occupations with lesser education requirements. The broad definition of skill mismatch includes various sub-concepts such as skill shortages, skill gaps, vertical mismatches (such as overeducation and overskilling) and

[10] These developing countries are taken from three world macro regions – (i) Sub Saharan Africa, (ii) South East and Central Asia and (iii) Latin America and Caribbean.

horizontal mismatches[11]. Among them, overeducation is one of the most frequently used measures. Overeducation is not specific to developed countries only (Cultrera et al, 2022, Croce and Ghignoni, 2012, Leuven and Oosterbeek, 2011). It is seen in many developing countries as well (Battu and Bender, 2020; Quinn and Rubb, 2006). While Castello-Climent and Mukhopadhyay (2011) advocate a greater expenditure on higher education in developing countries, to facilitate growth, the existence of overeducation among tertiary graduates cannot be brushed aside. Jaume (2021) points out that an education expansion policy without appropriate job opportunities is likely to aggravate the incidence of overeducation. Also, higher educated workers are unable to utilise their productivity advantage they secure through education. The first chapter in the thesis fills the gap in the literature by looking at budget reallocation policies to resolve specific problems such as school dropouts and overeducation, both of which needs to be addressed in developing nations.

The choice of schooling is quite an important issue in the literature on education economics. As Lahoti and Mukhopadhyay (2019) points out, some of the main factors affecting parental decisions for their child's schooling involves infrastructure, affordability, medium of instruction, teaching quality, availability of special facilities etc. The general scepticism of households about government schools in developing countries has often been highlighted in the literature. Glick and Sahn (2006) shows that multigrade teaching in government schools is a primary deterrent against choice of government schools in Madagascar. Nishimura and Yamano (2013) reveals the ineffectiveness of the Free Primary Education Policy implemented in Kenya which involved free provision of education for

[11] Skill shortages refer to the situation when there exist vacancies which cannot be filled up due to lack of qualified candidates. Skill gaps refer to the case when workers lack the required skills to deliver the roles in the job. Vertical mismatches imply a vertical gap between the education and skill level of workers, whereas, horizontal mismatch implies an individual working in a job completely unrelated field of study. Under vertical mismatches, overskilling might appear to be a better estimate for skill mismatch compared to overeducation, as it accounts for actual skill levels of workers. However, it is extremely difficult to get data on skills, especially on those derived from experience, innate ability or from informal sources.

children in government schools. The policy had failed to meet its objective due to the poor quality of government schools and resulted in a mushrooming of private schools. In the Indian context, government schools are beset with a host of problems. Chaudhary et al. (2006), Kremer et al. (2005), Duflo et al (2012a), Kingdon, (1996) emphasizes on the acute teacher absenteeism present in government schools. However, despite all odds against government schools, being the cheaper option, they are still preferred by a large section of households (Harma, 2011; Woodhead et al., 2013).

The literature also reveals the surge in private tuitions across developed and developing countries. Bray (1999, 2013) provides a clear picture of the evolving private tuition markets in different countries across the world, especially Asia, Africa, Eastern Europe and Latin America. The intake of private tuitions is driven by diverse factors. For instance, Tan (2017) argues that the consumption of private tuitions is determined by parentocracy, where parents wish their children to secure a competitive edge over other students. Peer pressure might be another reason driving demand for private tuitions in schools (Sujatha, 2014). Bhorkar and Bray (2018) and Bhorkar (2023) also discusses the role of entrance examinations that lead to an uptake in the incidence of private tuitions. According to them, parents feel that mainstream education is inadequate in providing resources and learning for these dynamic examination systems. Alongside all these factors, a very important reason for intake of private tuitions is the poor quality of teaching in schools. According to Biswal (1999), the corruption in government schools characterised by shirking behaviour of teachers forces children to opt for private tuitions. This also follows from the literature on teacher absenteeism discussed earlier. As pointed out by Sujatha (2014), one of the reasons for attending private tuitions is that a large proportion of teachers do not teach well in class. The effect is more pronounced in government schools. Also, from the theoretical analyses in Kim and Lee (2010) and Dang and Rogers (2008), private tuitions are expected to fulfil the

education requirements which the supply constrained public education system fails to provide. From these existing analyses, it is evident that private tuition acts as a tool for bridging the gap between private and government school teaching. The literature also reveals that the incidence of private tuitions is higher for higher levels of schooling (Azam, 2015; Dang, 2007; Bray, 1999; Bray et al., 2013; Tansel and Bircan, 2006).

Another dimension which the literature on private tuition focuses on, is the impact private tuition has on academic achievement. Aslam and Atherton (2012) show that government school students in India and Pakistan benefit more from private tuition classes in comparison to private schools. The largest gain accrues to the poorest of students. This reveals how inadequate the government school teaching is in both the countries. Dang (2007) and Dongre and Tewary (2015) shows that private tuition has a positive significant impact on the performance of students in Vietnam and India respectively. However, the literature does not discuss the effect of private tuition market on school choice. Using data on rural Bihar, Banerji and Wadhwa (2015) provides evidence on poor quality of government schools inducing a shift to private schools and also the rising participation in private tuitions. However, the correlation between school choice and private tuitions has not been drawn. The third chapter of the thesis attempts to connect the two important literatures, one on school choice, the other on private tuitions. It looks into the effect of private tuition expenditure on school choice behaviour of households. It also provides a comparison between private and government school teaching quality which is largely unobserved in the literature.

Measurement of overeducation entails a comparison of the level of education acquired by a worker versus their occupational requirements. Conventionally, three methods are used to calculate overeducation and undereducation empirically– subjective method, which depends on the self-assessment of workers (Duncan and Hoffman (1981), Battu et al. (1999), Galasi (2008)), job evaluation method, which entails comparison of occupation

requirements done by professional job analysists (Rumberger, 1981, Hartog and Oosterbeek, 1988) and realised matches or empirical method, where the education level of workers are compared to the mean or modal level of education within an occupation (Verdugo and Verdugo, 1989; Kiker et al, 1997). However, each of these methods have several shortcomings. While the first method is subject to bias due to non-response or overreporting of educational requirements, the second method might be extremely costly, given the need to upgrade occupational requirements frequently, whereas the third method suffers from the inability to address the demand side of education, via the labour market. A comparatively newer method of estimating overeducation has been introduced by Mehta et al. (2011) in the context of four developing countries, which has also been adopted later by Roy Chowdhury et al. (2021) for measuring overeducation in the 'manufacturing' and 'mining and construction' sectors in India. This methodology involves calculation of the rate of return to education levels using the Mincerian equation (Mincer, 1974) and identifying the education level that provides significant positive returns. Any worker with excess education getting insignificant returns is deemed as overeducated. The empirical literature on overeducation shows that though the returns to surplus education for overeducated workers are positive, their wages are lower compared to matched workers (Sicherman, 1991; Darko and Abrokwa, 2020; Cohn and Khan, 1995; Verhaest and Omey, 2012). The literature on overeducation in India is relatively limited (Roy Chowdhury et al., 2021; Kukreja, 2018; Sharma and Sharma, 2017; Mukherjee and Paul, 2012; Sengupta, 2017). Besides, except Roy Chowdhury et al. (2021) and Mehta et al. (2011), none examines the presence of overeducation in elementary occupations. The fourth chapter of the thesis uses the returns to education method to estimate overeducation in disaggregated categories within elementary occupations in India, where educational requirements are low according to the official standards.

1.3 Outline of the chapters

The thesis consists of three core chapters. This section provides a concise description of these three chapters.

Chapter 2 explains the role of a government budget reallocation policy on two pressing issues in developing countries, namely school dropouts and overeducation. The theoretical model set up in the chapter consists of utility maximizing households, who decide for their child's education across two periods – school and higher education. We consider identical households differing along ability of their child. The model consists of education production functions which includes three important features -first, the path dependence between school and higher education investment, second, the complementarity between household and institutional investments and third, the interdependence between school and higher education outcomes. These three characteristics are instrumental in deriving the results of the model. The optimization exercise is solved by backward induction, with households first making choices about higher education and then deciding whether or not to complete schooling. At the equilibrium, the representative household chooses one of the three options for their child based on the endogenously determined ability thresholds– dropping out of school and joining the unskilled labour market, completing school and joining the semi-skilled labour market and opting for higher education. However, we consider that higher education does not always ensure a skilled employment. In such a situation, higher-educated individuals are absorbed in semi-skilled occupations as overeducated workers. Given the household choices of education derived in the model, we check for the implications of a government reallocation of education budget in favour of school education. The results show that such a policy resolves both the school dropout and overeducation problems under certain conditions, else ends up aggravating these issues. The chapter also shows that such a reallocation is ineffective in completely eliminating school dropouts.

Chapter 3 highlights the role of the private tuition markets, which operates parallel to the formal education system and supplements it. One of the primary sources of private tuition demand among households apart from the necessity to meet the competitive needs of students and peer pressure, is driven by the poor quality of teaching in schools. The analysis in this chapter is an attempt to connect the private tuition market with the choice of schooling at secondary and higher secondary levels. Given that the private tuition markets can assuage the deficiencies in formal school teaching, it could be feasible for the parent to send the child to the more affordable schooling option after taking into account the cost of private tuitions. As known from the school choice literature, the household choice of schooling depends on a host of factors other than teaching quality such as medium of instruction, availability of special facilities, discipline, infrastructure etc. While private schools fare better in these school specific characteristics, government schools remain the more affordable option. The empirical model constructed in this paper addresses how the cost of private tuition interacts with school specific characteristics and determines school choices of households. Due to unavailability of data on the cost of private tuitions, we take the share of private tuition expenditure of a child in the monthly per capita consumption expenditure of the household as a suitable estimate. The results show how private tuition crowds out private schooling as long as it is a reasonable option to households. The chapter also constructs a theoretical school choice model simultaneously with the empirical framework to infer about the relative differences in teaching quality between government and private school teaching quality.

Chapter 4 deals with the measurement of overeducation in elementary occupations in the Indian labour market. Following from the Periodic Labour Force Survey Data, we find around 3.8 percent of the tertiary educated in the working age group works in elementary occupations, for the year 2018-19. Given the substantial investment costs associated with higher education, this is indeed a matter of concern owing the low skilled

elementary occupations they are forced to work in. As against the conventional methods, which calculates overeducation using some predefined measures, we adopt the returns to education approach which first identifies the education level in an occupation providing positive significant returns and then calculates the proportion of workers with excess education. Since we are especially concerned with overeducation among higher-educated, we identify the elementary occupations where tertiary graduate workers do not get any positive significant returns and then calculate their proportion in the occupation. The analysis in this chapter starts with the measurement of overeducation in elementary occupations considered as a broad category and then proceeds into narrowly defined disaggregated categories within elementary occupations. The three occupation groups that we focus on in the study are 'Domestic and related helpers, cleaners and launderers', workers in 'Mining and construction' and 'Messengers, Porters, Door Keepers and Related Workers'. The main objective of this chapter is to test the hypothesis of whether all higher-educated workers in elementary occupations are overeducated for getting insufficient returns. Owing to the heterogenous nature of work, we focus on gender specific occupation categories in narrowly defined occupation categories, to get precise estimates. The chapter also briefly discusses the extent of overall overeducation in these occupations, which also includes workers with lower education levels, and compares with the estimates under the modal method of defining overeducation.

1.4 Results

Chapter 2 addresses the issues of school dropouts and overeducation in a unique theoretical model which accounts for complementarity between household and infrastructural investment in education production function on the one hand, and school and higher education, on the other. The results derived in the chapter show that a reallocation of the education budget in favor of school education can solve both problems in one go if two conditions are satisfied – the quality of school education is more responsive to school infrastructure compared to

teaching quality and the semi-skilled wages are sufficiently responsive to school education quality. The chapter, on the one hand, questions the efficacy of the widely practiced budget reallocation policy in favor of school education and on the other, shows that the reallocation cannot eliminate school dropout.

Chapter 3 argues that private tuition to some extent crowds out private schooling. *Ceteris paribus* it compensates for the teaching quality deficiency in government schools. As long as the cost of private tuition is below a threshold, a household may prefer to send its child to a government school and private tuition. With costlier private tuition, government school is substituted by private school. Using the National Sample Survey 75th Round, 2017-18 data on education in India, we estimate a negative significant impact of the share of private tuition expenditure in monthly per capita expenditure of a household on government school enrolment of its child, who is in the 13-18 age group. The result remains robust after controlling for district-level school-related variables. The chapter also uses a theoretical model for explaining its empirical finding. The theory argues that the empirical results derived in the paper provide hard-to-find evidence in support of the widely held view that on average the unaided private schools in India impart better teaching quality compared to the government schools. The results also help us to understand the likely impact of policies like the crackdown on private tuition as has happened in China recently on school choice and quality of education in a country.

Chapter 4 estimates overeducation in different disaggregated categories within elementary occupations in India. It uses the returns to education approach for estimating overeducation which takes into account both the demand and supply sides of the labour markets. In two occupations - 'Domestic and related helpers, cleaners and launderers', and 'Mining and construction', workers with higher education do not get any positive significant return, and thus confirms the presence of overeducation. On the contrary, higher-educated workers in 'Messengers, Porters, Door Keepers and Related Workers', despite falling under

the broad category of elementary occupations, gets significant positive returns. Thus, while the chapter confirms the existence of overeducation in some elementary occupations it points out in some others that despite low education requirements, workers with higher education get positive significant returns. This indicates, that though the official education requirement for a job might be low in an occupation, a worker with a higher than required education level might signal higher productivity to the employer and secure a higher return. In a way, it circumvents the existing biases associated with the predefined measures of overeducation. While calculating the overall rates of overeducation in an occupation, the chapter finds that among domestic workers and 'mining and construction' workers, those with any education are overeducated. For 'Messengers, Porters, Door Keepers and Related Workers', all workers with education between 'below primary' and 'higher secondary' are overeducated. In comparison to the returns to education approach, the estimates match for the domestic workers, whereas, leads to an underestimation in the latter two occupations using the modal method of calculating overeducation.

1.5 Plan of the Thesis

The rest of the thesis is organized as follows. Chapter 2 analyses the impact of a government reallocation of the education budget on school dropouts and overeducation. Chapter 3 investigates the role of private tuition cost on the choice of schooling at secondary and higher secondary level among students in India. Chapter 4 explores the extent of overeducation present in elementary occupations in India, which are typically characterised by very low education requirements going by the official standards.

CHAPTER 2*

SCHOOL DROPOUT AND OVEREDUCATION IN DEVELOPING ECONOMIES: FEASIBILITY OF A BUDGETARY SOLUTION.

***** This chapter is based on

Gooptu, S., & Mukherjee, V. (2022). School dropout and overeducation in developing economies: Feasibility of a budgetary solution. *Review of Development Economics*, 27(2), 825-852. https://doi.org/10.1111/rode.12964

2.1. Introduction

The United Nations Educational, Scientific and Cultural Organization (UNESCO, 2019) reports that there were 258.4 million out-of-school children in the World in 2018. The scenario was the worst in Sub-Saharan Africa (97.5 million) and South Asia (93 million)[12]. At the same time, according to International Labour Organization (ILO, 2020), around 258 million workers were overeducated across 114 countries (from all regions and income levels across the world) in the sense that their education level exceeds his/her job requirement[13]. In developing countries like India, Roy Chowdhury et al (2019) estimate that 66 percent of workers in the Indian manufacturing sector are overeducated. Kukreja (2018) reports the education mismatch rate in the Indian textile and clothing industry is 67.61 percent which is much above the world average. Due to school drop-outs, an economy faces a loss in labour productivity, national income, and the high social return associated with school education (Rumberger, 1987; Levin, 1972). Similarly, due to overeducation, an economy loses out because of the high opportunity costs associated with education expenditure.

The chapter addresses the dual problem of the high rate of school drop-out and the high incidence of overeducation, prevalent in developing countries, in a theoretical framework.

[12] According to the data from World Development Indicators by World Bank, out of the children in the primary school going age, 9 percent, 20 percent, 3 percent are out of school in South Asian countries, Sub Saharan countries and Latin American and Carribean (excluding high income) countries in the year 2020. The corresponding percentage for least developed countries (by United Nations classification) is 17 percent. The United Nations (2020) estimates show that the rate of school completion at primary level is only 34 percent for the poorest 20 percent of households in low-income countries. For more evidence on school dropouts in developing countries, see Mussida et al (2018), Gibbs and Heaton (2013), Hirakawa and Taniguchi (2020) and Woldehanna et al. (2021)

[13]See McGuinness (2006), McGuinness et al. (2017), Leuven and Oosterbeek (2011) for definitions of overeducation. Conventionally it is measured as vertical mismatch in skill level of a worker if his/her educational qualification exceeds the mean or modal level of education at the job or the required education level of the job. Mehta et al. (2011) uses the 'returns to skill' approach in measuring overeducation in developing countries. Duncan and Hoffman (1981) and Verdugo and Verdugo (1989) show that even though overeducated workers receive positive significant returns, they receive less than that of the appropriately matched workers. The literature also provides recent evidence on overeducation in developed (Cultrera et al. (2022), Kracke et al. (2017)) and developing countries (Castro et al. (2022), Battu and Bender (2020), Morsy and Mukasa (2020)). According to ILO (2020), Madagascar has a large share of overeducated workers (23 percent in 2015) among low -income countries. Among low-middle income countries, Vietnam, Mongolia, Micronesia, Tajikistan and Georgia have relatively higher rates of overeducation in 2018 (20,23,25, 22 and 24 percent respectively).

The government being the most important financier of education in developing countries (UNESCO, 2016), we study the role it can play in solving these problems. Assuming that the government's objective is a reduction of both, the chapter examines the feasibility of solving these problems through the reallocation of its education budget, which is known for its bias in favor of higher education[14]. Of late, although developing countries tried to achieve universal enrolment in primary schools for fulfilling the UN's Millennium Development Goals, from the data reported above it appears that the question of school drop-out and overeducation have not received sufficient attention. The chapter finds the conditions under which the reallocation instrument would achieve the above objective.

The chapter models a single-child household's decision about the schooling and higher education of its child. The model considers an educational production function that incorporates the following features of a typical education process (Majumdar (1983)): (i) household and institutional investments are complementary to each other in determining educational outcome; (ii) path dependence in choice of educational investment at the higher level and (iii) interdependence between school and higher education outcomes. This approach acknowledges that educational achievement depends not only on the ability of the child but also on the investment in education made by the household and institutional investment provided by the government which comes in the form of the building of schools and colleges with all the relevant facilities and recruitment of teachers[15]. The school outcome has crucial

[14] Refer to Footnote 5.

[15] Although in OECD countries the main source of funding in primary and secondary education is public (OECD (2021)), Nordbloom (2003) includes parental investments in other goods associated with education in the human capital production function such as computers, books and other complements to formal education. In South Asian countries like India, China and Japan, apart from these, a substantial amount is spent for private tuition of the child. Apart from the opportunity cost of not sending the school dropped-out child to un-skilled labor market, the 75th Round National Sample Survey on education in India (NSS, 2019) finds that the average annual expenditure on education in the year 2017-18 varies from INR 1030 (15 USD) at pre-primary level to INR 7000 (102 USD) at higher secondary level in government schools. For private aided institutes, it varies from INR 13223 (193 USD) at pre- primary level to INR 16445 (240 USD) at higher secondary level.

implications for higher education: the set of opportunities available at higher education depends on the irreversible investment made at the school level. The quality of school education, on the other hand, depends on the higher education outcome since schoolteachers are appointed from the higher education system, which is emphasized in papers like Chetty et al (2014), Aaronson et al. (2007), Azam and Kingdon (2014)[16]. While deciding about the allocation of its budget at different stages of education, the government considers these features of education production function and internalizes it. In the theoretical literature papers like Arcalean and Schiopu (2009) although considered production functions that capture the complementarity between household and institutional investments in determining the educational outcome, the interdependence between school and higher education outcomes through teachers' quality has not been addressed. On the contrary, papers like Gilpin and Kaganovich (2009)[17] acknowledge the impact of teaching quality via aggregate higher education quality, on quality of schooling. but does not incorporate the complementarity between household and institutional investment. The current chapter addresses both these features. The existing literature on school education while focuses on school-specific strategies, like improvement of school infrastructure, reduction of teacher-absenteeism, and improvement of the governance structure of the school through an increase in per-pupil expenditure (Angrist and Lavy, 1999; Duflo et al., 2012a; Duflo et al., 2012b; Björkman, 2004; Glewwe et al., 2014; Baron, 2022; Jackson and Mackevicius, 2021; Lee and Polachek, 2017), this chapter emphasizes on the quality of school

[16] Clotfelter et al (2007), Kingdon (2006) argue that students' outcomes are positively affected by teachers' credentials and possession of a master's degree, respectively. Glewwe et al (2014) in their review of the related literature in developing country context find that the most consistent determinants of learning at schools are "having teachers with greater knowledge of the subjects they teach, having a longer school day, and providing tutoring. However, Hanushek and Rikvin (2006) and Hanushek (1997) point out that this effect can be weak.

[17]Gilpin and Kaganovich (2009) in an overlapping generation model study the trade-off between quantity and quality of teachers in school education in presence of better outside opportunities for college graduates. The feature of the framework of the present chapter that the quality of higher education appears as an externality in a household's investment decision is similar to their paper. However, since the model, developed in the context of US, assumes universal basic education, the quality of school education is not a function of the household's investment in school education as is the case in most of the developing countries in the world.

teachers, which depends on higher education quality, and matters for the success of such strategies.

In our model, the otherwise identical households decide on educational investment according to the perceived ability level of their child. Similar to papers like Galor and Zeira (1993), Baland and Robinson (2000), Ranjan (2001) the households are assumed to be credit-constrained. At the benchmark equilibrium, every household sends its child to higher education after completion of his school education. Low-ability children drop out before school completion. Some high-ability children, after their higher education, find one of the limited numbers of jobs available in the skilled labor market. The rest are absorbed into the semi-skilled labor market as overeducated workers. Starting from the benchmark situation, we derive the conditions under which a reallocation of the government's budget in favor of school education would promote the completion of school education and the reduction of enrolment in higher education. The reallocation improves the school infrastructure and worsens the higher-education infrastructure. Consequently, in the first round, the higher-education quality falls with an uncertain effect on school-education quality. The uncertainty is derived from two opposing effects working on schooling-outcome: a direct positive effect from improved infrastructure and an indirect negative effect from the fall in teaching-quality since the schoolteachers are college graduates. In the second round, the school-input impacts the quality of higher education with a feed-back effect on school education itself and so on. If the direct effect dominates the indirect effect, school education quality improves, and households are incentivized to spend more towards the completion of their child's schooling. The school dropout rate falls. If the improved schooling quality sufficiently raises the semi-skilled wage, more households reverse their decision of sending them to college after school completion. This simultaneously solves the problem of overeducation and overspending in higher education. If the indirect effect on school-education is dominant, schooling-outcome

deteriorates, and the reallocation renders counterproductive results. Therefore, the chapter questions the efficacy of the well-received policy of biasing the education budget towards school education. It also shows that such a reallocation cannot completely wipe out the school drop-out problem in developing countries.

The scope of the present chapter is closely related to papers like Galor and Zeira (1993), Hidalgo-Hidalgo and Iturbe-Ormaetxe (2012), and Arcalean and Schiopu (2009) that use theoretical models to study the effect of reallocating the education budget across education levels.[18] The main difference between the model in this chapter and Galor and Zeira is that in Galor and Zeira the main source of heterogeneity is wealth and the inability of the poor to borrow at a lower rate of interest, while in the current chapter the ability of the students and their earning potentials are different. The households are assumed identical in terms of wealth but they are credit constrained. Another difference is the structure of the labor market; the model in the chapter has a semi-skilled labor market. While Hidalgo-Hidalgo looks into the effect of the budget-reallocation on equity and efficiency, Arcalean and Schiopu focus on growth. In a recent paper, Abbott et al. (2013) use a sequential decision-making structure similar to ours, but they do not study the reallocation problem. The present chapter differs from the scope of all these papers as it deals with the problems of school drop-out and over-education. The model highlights the positive external effect of higher education on school education through improved teaching quality, which unless sufficiently weak, impairs a budget reallocation towards schooling in reducing school-dropouts and overeducation. Our treatment shows that the problems of school dropouts and overeducation are related and acknowledging these interdependencies is important for policy formulation. For example, although papers like Castello-Climent and Mukhopadhyay (2011) advocate higher expenditure on higher

[18] See also Su (2004), Blankenau (2005) for other models relating hierarchical education structure and budget reallocation to growth, efficiency and equity.

education[19], the present chapter shows such a policy can be counterproductive in developing countries. In our context, a solution to the problem of overeducation provides a stimulus to growth as well, as it saves scarce resources by avoiding expenditure on higher education. So, this chapter ends up addressing the problem of 'tertiary tilt' that exists in education-budget of developing countries and finds out the conditions under which a budget-reallocation helps.

The next section of the chapter describes the model. Section 3 generates the equilibrium, followed by the results in Section 4. The section following concludes, discusses the limitations of the model, and offers a sketch of possible extensions.

2.2.The Model

Consider an economy consisting of only single-child households of measure 1. The households are identical to each other except that the ability level of the child, denoted by $a \geq 0$, differs from one household to another and follows a uniform distribution over $[0, A]$. It is also assumed that the child's ability is perfectly known to his parents. A household lives for two periods. In period 0 it is born, and it dies at the end of period 1. The household may invest in its child's education at two different points in its lifetime: at $t = 0$ it may spend on the child's school education; if the investment is sufficient, at the end of $t = 0$, the child completes school education and then, at $t = 1$, the household may spend on his higher education. The model abstracts away from household bargaining and assumes that parents make all the decisions on behalf of their child regarding the child's education. There exist three different kinds of labor markets: one that employs unskilled labor, one that employs semi-skilled labor, and the other that employs skilled labor. At $t = 0$, if the child does not complete school education, he is sent to an unskilled labor market where he surely gets a job[20]. At $t = 1$, if the child is not sent to

[19] Hidalgo-Hidalgo and Iturbe-Ormaetxe (2012) points out the caveats against reallocating education budget towards basic education in developing countries for addressing both equity and efficiency.

[20] The number of students not completing schooling may also subsume those students who never attended school. However, we assume universal enrolment inspired by the achievement of countries like India, where the

higher education, after completion of schooling, he is sent to a semi-skilled labor market where also he gets a job for sure. We assume, every child enrolled in higher education, completes it and becomes eligible for joining the skilled labor market. However, it is common knowledge that a job is not guaranteed in the skilled-labor market. Only $\delta \in (0,1)$ proportion of higher education degree holders get a skilled job; the rest are compelled to work in the semi-skilled market. A child with a higher education degree, if works in a semi-skilled labor market, is considered as overeducated. Let a_H be the threshold ability level, the students with ability above which pursue higher education. Since a follows a uniform distribution over $[0, A]$, the number of students attending higher education is $\left(1 - \frac{a_H}{A}\right)$. Out of them $\delta\left(1 - \frac{a_H}{A}\right)$ secure a skilled job, the remaining $(1-\delta)(1-\frac{a_H}{A})$ enters the semi-skilled market as overeducated workers. Also, let a_S be the ability threshold, the children with an ability level above which complete schooling. Thus, the number of school dropouts is given by $\frac{a_S}{A}$. The ability-level thresholds a_H and a_S are endogenously determined in the model and are influenced by the government's budget-reallocation decision.

The complementarity between household and infrastructural investment in the production of educational quality both at the school and higher education stage is the focal point of the chapter[21]. We assume that the infrastructural investment is provided by the government[22]. Following Bearse et al. (2005), we capture the complementarity between the two

successful implementation of 'Right to Education' act ensured that all the children are enrolled in schools. It had led to a rise in enrolment rates at primary levels of schooling in India. The UDISE Plus (2022) Report shows that the gross enrolment at primary level is 103.4 percent. However, despite reduction in the proportion of never enrolled children, the problem of dropouts persists. The UDISE Plus (2022) Report also shows that the dropout rate at secondary level (class 9-10) is 12.6 percent.

[21] The degree of complementarity between the two, however, is quite a debatable issue and the literature has not yet been able to reach a consensus on this.

[22] The optimal ownership of infrastructural investment in education is debated. However, the magnitude of positive externality associated with it, makes the government natural choice for provision of such an investment.

types of investment through the use of the Constant Elasticity of Substitution (CES) production function at both stages of education.

The higher education production function is written as

$$q_H = aZ(Q_H)\, q_S[\alpha_1(c_1^H)^{\rho_1} + \beta_1(c_1^G)^{\rho_1}]^{\frac{1}{\rho_1}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.1)$$

where q_H is the quality of higher education achieved by a child given the quality of his school education q_S. In equation (2.1), c_1^H and c_1^G are the respective household and government investments per student in higher education and $\alpha_1 > 0$, $\beta_1 > 0$ are the weights of these investments in the determination of higher education quality. The degree of complementarity between the two types of investments is represented by $\rho_1 \in (-\infty, 1)$. Equation (2.1) captures the fact that the quality of higher education achieved by a child is affected by the innate ability a of the child which is required for the absorption of the available higher education quality; all other things remaining the same, higher a implies higher q_H. Also, it is important to notice from equation (2.1) that the quality of school education of a student has an impact on the quality of higher education he can achieve; higher q_S implies higher q_H. While the quality of schooling is measured by the marks or grades obtained by the student in the final school-level examination, the quality of higher education is measured by the higher education degree achieved. q_H is also a function of the overall higher education quality $Q_H = \frac{1}{A}\int_{a_H}^{A} q_H(\cdot)da$, $Z'(Q_H) > 0, Z''(Q_H) < 0$.

Like the higher education production function, the school education production function is written as:

$$q_S = aT(Q_H)[\alpha_0(c_0^H)^{\rho_0} + \beta_0(c_0^G)^{\rho_0}]^{\frac{1}{\rho_0}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.2)$$

where c_0^H and c_0^G represent the respected levels of household and public investment per child for completion of school education; $\rho_0 \in (-\infty, 1)$ represents the degree of complementarity

between the two. The parameters $\alpha_0 > 0$ and $\beta_0 > 0$ are the weights of these two types of investments in the determination of the quality of school education. Equation (2.2) shows that the quality of school education positively depends on the teaching quality T at schools. Since the teachers are appointed from the higher education system, similar to Z, T is also assumed to be a function of Q_H with $T'(Q_H) > 0, T''(Q_H) < 0$. However, in both (2.1) and (2.2) the households consider Q_H as given, and they are too small to influence its value. As q_S depends on Q_H, notice that equation (2.1) represents how children with different ability levels absorb the direct and indirect impact of available higher education quality to determine their own higher education quality.

Let $\bar{c}_0^H$ be the threshold level of investment that is necessary for the successful completion of a child's schooling. If $c_0^H \leq \bar{c}_0^H$, the household is short of investing enough to complete the child's schooling. If $c_0^H > \bar{c}_0^H$, the child completes school. For simplicity, we normalize the threshold investment level $\bar{c}_0^H = 0$.

As mentioned before, three different types of labor markets exist in the economy: unskilled, semi-skilled, and skilled. The minimum wage in the economy $\underline{w}$ is received in the unskilled labor market independent of ability or the level/quality of education. The wage rate at the other labor markets, however, are functions of educational level/quality and includes skill premium. A child can join the semi-skilled labor market only after completion of school education. The wage rate in the semi-skilled labor market w_I is a function of the quality of school education. We assume, $w_I(q_s) > \underline{w}$ for all $a > 0$ and $q_s > 0$. Similarly, a child can join the skilled labor market only after completion of higher education. The wage rate at the skilled labor market w_F is a function of the quality of higher education. We assume, $w_F(q_H) > w_I(q_s)$ for all $a > 0$, $q_s > 0$ and $q_H > 0$. However, while an educated child can join the semi-skilled labor market for sure, joining the skilled labor market is assumed to be restrictive. Given that

child with higher education can secure a job in the high skilled labor market only with probability $\delta \in (0,1)$, the expected labor-market return from higher education is given by $[w_I(q_S) + \delta(w_F(q_H) - w_I(q_S))]$. The term $(\delta(w_F(q_H) - w_I(q_S)))$ represents the wage premium that the household expects to receive on sending the child for higher education.

Assumption 2.1: $w_F(q_H, a) = \underline{w} + \gamma q_H$ and $w_I(q_S, a) = \underline{w} + \mu q_S$ where $\gamma > 0,\ \mu > 0,$ $\gamma q_H > \mu q_S$ and $\gamma \frac{\partial q_H}{\partial q_S} > \mu$. As $a \to 0,\ w_F(0) = w_I(0) \to \underline{w}$.

The assumption about the linearity of wage functions simplifies the analysis. The inequality $\gamma q_H > \mu q_S$ ensures that the wage premium from higher education is positive. We also assume $\gamma \frac{\partial q_H}{\partial q_S} > \mu$ i.e. the school education has a greater impact on the skilled wage than the semi-skilled wage. However, a child with an ability close to zero commands a wage close to the $\underline{w}$, independent of the segment of the labor market he works in. The endogenization of wages would determine γ and μ through the interplay of market demand and market supply at respective skill levels. But, for simplification, we abstract away from this analysis.

At period $t = 0$, the household endowment and consumption are denoted by y_0 and x_0 respectively, and the same at $t = 1$ are denoted by y_1 and x_1 respectively. The chapter assumes the presence of credit constraints that prevents a household from borrowing against its future earnings, and also, from lending or saving at the initial period. Thus, whatever income is earned, is consumed in the period itself. Therefore, the lifetime utility function of the household is written as:

$$v = x_0 + \theta x_1, \qquad (2.3)$$

where $\theta \in [0,1]$ is the discount factor attached to period 2 consumption.

The modelling technique here follows papers like Baland and Robinson (2000) and Ranjan (2001). The linearity of the utility function is a simplifying assumption. The results of the chapter go through a non-linear specification as well.

The representative household's career-choice problem for its child involves maximization of v by choice of $\{x_0, x_1\}$ subject to the budget constraint, which turns out as an equivalent choice of c_0^H at $t = 0$ and c_1^H at $t = 1$.

With the specific features of the model described above, the decision tree of the household is given in figure 2.1 below.

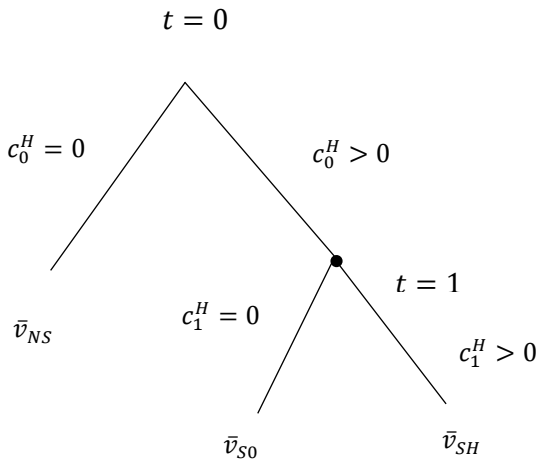


Figure 2.1: The Decision Tree of Representative Household

In Figure 2.1, the payoffs of the household from each possible choice about the career path of its child are shown. At period $t = 0$, the household has two options: either to choose $c_0^H = 0$ or to choose $c_0^H > 0$. If it chooses $c_0^H = 0$, the child drops out from school and works in the unskilled labor market, the payoff of the household for which is given by $\bar{v}_{NS}$. However, if the household chooses $c_0^H > 0$ at $t = 0$, it arrives at the decision node at $t = 1$ where there is the

option of choosing either $c_1^H = 0$ or $c_1^H > 0$ for the child. If $c_1^H = 0$ is chosen, the child does not go for higher education and enters the semi-skilled labor market. The payoff of the household if this option is chosen is given by $\bar{v}_{S0}$. On the other hand, if the household chooses $c_1^H > 0$, the child is sent for higher education, on completion of which he enters the skilled labor market with probability δ. The payoff of the household if this action is chosen is given by $\bar{v}_{SH}$. The following analysis derives the values of $\bar{v}_{NS}$, $\bar{v}_{S0}$ and $\bar{v}_{SH}$ and the equilibrium decision path of a household.

2.3.The Equilibrium

2.3.1 The households' problem

We solve a household's decision-making problem by application of backward induction method starting at $t = 1$.

<u>Decision Making at $t = 1$</u>

Option 1: *Given $c_0^H > 0$, the household sends its child to higher education and decides about $c_1^H > 0$.*

Given its period 0 choice of c_0^H (and therefore, x_0), while maximizing v by choosing x_1 (and therefore, c_1^H) the household faces uncertainty in securing a skilled job. Since there is a dearth of skilled jobs, with probability $(1 - \delta)$ the child may end up in the semi-skilled labor market. If he joins the skilled labor market and earns $w_F(q_H, a)$. Then, by assumption 2.1, the budget constraint of the household is:

$x_1{}^h = y_1 + \underline{w} + \gamma q_H - c_1^H$.

However, if he finds his job in the semi-skilled labor market, the budget constraint of the household is:

$x_1{}^{l} = y_1 + \underline{w} + \mu q_S - c_1^H$.

Therefore, using (2.3), the household's maximization problem can be written as:

maximization of its expected utility

$$v = x_0 + \theta[y_1 + \underline{w} + \mu q_S + \delta[\gamma q_H - \mu q_S] - c_1^H] \quad (2.4)$$

subject to $c_1^H > 0$. Since q_H is function of c_1^H from the higher education production function defined in (2.1), the first-order condition for the choice of $c_1^H(q_S, a, c_1^G, \delta, \gamma, Q_H) > 0$ is:

$$\delta\gamma.\frac{\partial q_H}{\partial c_1^H}(q_S, a, c_1^G, c_1^H, Q_H) = 1. \quad (2.5)$$

At the equilibrium $c_1^H(q_S, a, c_1^G, \delta, \gamma, Q_H)$ balances the expected marginal benefit of investment in higher education with its marginal cost given by 1. From (2.1), $c_1^H(q_S, a, c_1^G, \delta, \gamma, Q_H)$ also determines the quality of higher education received by the child.

Observation 2.1: (i) $\frac{\partial c_1^H}{\partial a} > 0, \frac{\partial c_1^H}{\partial q_S} > 0, \frac{\partial c_1^H}{\partial c_1^G} > 0, \frac{\partial c_1^H}{\partial \delta} > 0, \frac{\partial c_1^H}{\partial \gamma} > 0, \frac{\partial c_1^H}{\partial Q_H} > 0$.

(ii) $\frac{\partial q_H}{\partial a} > 0, \frac{\partial q_H}{\partial q_S} > 0 \frac{\partial q_H}{\partial c_1^G} > 0, \frac{\partial q_H}{\partial \delta} > 0, \frac{\partial q_H}{\partial \gamma} > 0, \quad \frac{\partial q_H}{\partial Q_H} > 0.$

Proof: See the Appendix 2.A. □

Observation 2.1 specifies the factors that induce higher investment in a child's education c_1^H, which leads to a better higher education degree, q_H. All the parameters that increase the expected return from investment in its child's higher education favour such an investment. It rises either with the higher probability of landing a skilled job, δ, or with the increased sensitivity of the skilled wage rate to higher education quality, γ. Any exogenous improvement in the quality of school education, q_S will have a similar effect. More is spent on a child with a higher ability a, as he is expected to do better in higher education. Because of the complementarity of c_1^H and c_1^G in the higher education production function, any increase in c_1^G

improves q_H. An improvement in the higher education quality, Q_H, improves the quality of teachers at higher education institutes, improving both c_1^H and q_H.

If a household decides to send its child to higher education, its payoff, which is the expected value of its indirect utility is:

$$v_{SH} = x_0 + \theta(y_1 + \underline{w} + (1-\delta)\mu q_S + \delta\gamma q_H(q_S, a, c_1^H(\cdot), c_1^G, Q_H) - c_1^H(\cdot)). \quad (2.6)$$

Option 2: *Given* $c_0^H > 0$, *the household chooses* $c_1^H = 0$.

In this case, in period 1, as the household decides against sending the child for higher education after schooling, its payoff is:

$$v_{S0} = x_0 + \theta(y_1 + \underline{w} + \mu q_S). \quad (2.7)$$

The preferred option is decided by comparison of v_{SH} and v_{S0}.

Let us define, $\varphi = v_{SH} - v_{S0}$.

Substituting from equations (2.6) and (2.7) it follows:

$$\varphi = \theta(\delta(\gamma q_H(q_S, a, c_1^H(\cdot), c_1^G, Q_H) - \mu q_S) - c_1^H(\cdot)). \quad (2.8)$$

Lemma 2.1: (i) $\frac{\partial \varphi}{\partial a} = \varphi'(a) > 0$ for all a in $[0, A]$;

(ii) $\frac{\partial \varphi}{\partial c_1^G} > 0$; $\frac{\partial \varphi}{\partial q_S} > 0$; $\frac{\partial \varphi}{\partial \delta} > 0$; $\frac{\partial \varphi}{\partial \gamma} > 0$; $\frac{\partial \varphi}{\partial \mu} < 0$; $\frac{\partial \varphi}{\partial Q_H} > 0$; $\frac{\partial \varphi}{\partial \theta} > 0$ *if* $\delta(\gamma q_H(.) - \mu q_S) > c_1^H(\cdot)$.

Proof: See the Appendix 2.A.

Lemma 2.1(i) states that after completion of school education at t =1 the amount of loss a household incurs in terms of payoff by not sending its child to higher education inflates with the ability of the child. The loss is higher if a child with higher ability is not sent for higher education after completion of his school education. The second part of the lemma states that

given $a > 0$, how the loss inflates or deflates with respect to change in c_1^G, q_S, δ, γ, μ, and Q_H. Also, it increases with θ if the value of the premium from attending higher education, exceeds the investment cost of higher education.

Observation 2.2: (i) If $\varphi(A) > 0$, there exists a value of $a = \bar{a}(q_S, c_1^G, \delta, \gamma, \mu, \theta, Q_H) \in [0, A]$ that satisfies $\varphi(\bar{a}) = 0$ and at $t = 1$, a household opts for the higher education of its child if and only if $a \geq \bar{a}$.

(ii) If $\varphi(A) \leq 0$, no household opts for the higher education of their child.

Proof: See the Appendix 2.A.

Since the objective of the chapter is to analyse overeducation, we avoid the case of $\varphi(A) \leq 0$.

Assumption 2.2: $\varphi(A) > 0$.

Assumption 2.2 ensures that all the children with their ability $a \geq \bar{a}$, if completes school education, are enrolled in higher education. Notice that if a child drops out of school at period $t = 0$, the question of opting for higher education does not arise at $t = 1$.

It follows that the average quality of higher education in the economy is determined as:

$$Q_H = \frac{1}{A}\int_{\bar{a}}^{A} q_H\left(Q_{H,} q_S, .\right) da. \tag{2.9}$$

Now that a household's decision regarding its child's higher education, if he completes school education, is resolved, we go one-step backward to the discussion related to the household's decision regarding the child's schooling at $t = 0$.

<u>Decision Making at $t = 0$</u>

Here, a household faces two options: either to send the child for completion of school education by investing $c_0^H > 0$ or to choose $c_0^H = 0$ by sending the child to the unskilled labor market where he earns $\underline{w}$. If the latter option is chosen, the household earns v_{NS}. While deciding about

its choice of $c_0^H > 0$, the household realizes that the action would take it to the decision node at $t = 1$, where, observation 2.2 suggests that a child with ability $a \geq \bar{a}$ is sent for higher education and v_{SH} is obtained, and a child with $a < \bar{a}$ is not sent for higher education and v_{S0} is obtained. Therefore, the household takes account of the ability level of its child along with v_{SH} and v_{S0} while evaluating its options at $t = 0$. We discuss the options separately below.

Option 1: $c_0^H = 0$.

Since the household does not invest in the completion of school education, here the payoff of the household is

$$v_{NS} = x_0 + \theta\, y_1. \qquad (2.10)$$

Now the choice of the household is limited in deciding only the value of x_0 that maximizes v_{NS} as in (2.10). From its budget constraint $x_0 = y_0 + \underline{w}$, it follows:

$$\bar{v}_{NS} = y_0 + \underline{w} + \theta y_1. \qquad (2.11)$$

Notice that $\bar{v}_{NS}$ is not a function of the ability of the child. The household compares $\bar{v}_{NS}$ with the payoff, it would have received if it had supported the child in the completion of his school education.

Option 2: $c_0^H > 0$.

Case 1: $a \geq \bar{a}$.

The household chooses x_0 to maximize v_{SH} as in (2.6) subject to the budget constraint: $x_0 + c_0^H = y_0$. After substitution of x_0 in equation (2.6), v_{SH} is rewritten as:

$$v_{SH} = y_0 - c_0^H + \theta(y_1 + \underline{w} + (1-\delta)\mu q_S + \delta\gamma q_H(q_S, a, c_1^H(\cdot), c_1^G, Q_H) - c_1^H(\cdot)),$$

which is a non-monotonic function of c_0^H. As the household maximizes v_{SH} by choosing $c_0^H > 0$, since equation (2.5) holds, the choice satisfies the first-order condition:

$$\theta \frac{\partial q_S}{\partial c_0^H}\left(\mu(1-\delta)+\delta\gamma.\frac{\partial q_H}{\partial q_S}\right)=1. \qquad (2.12)$$

Equations (2.12) is solved for $c_0^H(c_0^G, a, c_1^G, \delta, \gamma, \mu, Q_H, \theta) > 0$. The choice of c_0^H balances the marginal benefit from the higher quality of school education in terms of higher expected returns from skilled and semi-skilled labor markets with the marginal cost that arises due to the sacrifice of period 0 consumption.

By substituting $c_0^H(\cdot)$ in equation (2.2), the equilibrium value of q_S is derived as

$$\bar{q}_S = aT(Q_H)[\alpha_0(c_0^H(\cdot))^{\rho_0} + \beta_0(c_0^G)^{\rho_0}]^{\frac{1}{\rho_0}}. \qquad (2.13)$$

Observation 2.3:

(i) $\frac{\partial c_0^H}{\partial a} > 0, \frac{\partial c_0^H}{\partial c_1^G} > 0, \frac{\partial c_0^H}{\partial c_0^G} > 0, \frac{\partial c_0^H}{\partial \gamma} > 0, \frac{\partial c_0^H}{\partial \mu} > 0, \frac{\partial c_0^H}{\partial \delta} > 0, \frac{\partial c_0^H}{\partial Q_H} > 0, \frac{\partial c_0^H}{\partial \theta} > 0.$

(ii) $\frac{\partial \bar{q}_S}{\partial a} > 0,\ \frac{\partial \bar{q}_S}{\partial c_1^G} > 0, \frac{\partial \bar{q}_S}{\partial c_0^G} > 0, \frac{\partial \bar{q}_S}{\partial \gamma} > 0, \frac{\partial \bar{q}_S}{\partial Q_H} > 0, \frac{\partial \bar{q}_S}{\partial \mu} > 0, \frac{\partial \bar{q}_S}{\partial \delta} > 0, \frac{\partial \bar{q}_S}{\partial \theta} > 0.$

Proof: See the Appendix 2.A.

While the first part of Observation 2.3 characterizes the parametric changes in c_0^H, if the household decides to complete the child's school education, the second part characterizes the corresponding changes in his quality of school education, q_S. Notice that all parameters in the model unambiguously raise the marginal benefit of investing in school education for the household. It follows from the complementarity of investments in education production function that the household investment, c_0^H rises in tune with public investment in school education, c_0^G. Like Observation 2.1, here also, the higher the ability of the child, a, the greater the incentive of the parents to invest in his school education. On the other hand, responsiveness of skilled γ and semi-skilled wages μ to quality of school education and probability of getting a skilled job, δ also leads to a higher c_0^H and q_S. A higher level of public investment in higher

education, c_1^G, raises the marginal benefit of a household from investment in the child's school education and thus, c_0^H rises. Also, a rise in average higher education quality, Q_H , since it improves the teaching quality at schools, induces a household to spend more on its child's school education, i.e. higher c_0^H . In addition, for higher values of θ, when the household values the future more, it chooses a higher c_0^H with an expectation of higher returns in the future. This results in a higher q_S.

Substituting $c_0^H(\cdot)$ and $\bar{q}_S(\cdot)$ in v_{SH} we obtain its reduced form value as:

$$\bar{v}_{SH} = y_0 - c_0^H(\cdot) + \theta(y_1 + \underline{w} + (1-\delta)\mu\bar{q}_S(\cdot) + \delta\gamma q_H(\bar{q}_S(\cdot), a, c_1^H(\cdot), c_1^G, Q_H) - c_1^H(\cdot)). \quad (2.14)$$

Comparing equations $\bar{v}_{SH}$ and $\bar{v}_{NS}$ from equations (2.14) and (2.11) respectively we get:

$$\bar{v}_{SH} - \bar{v}_{NS} = \psi(a) - \underline{w}(1-\theta) \quad (2.15)$$

where, $\psi(a) = \theta[(1-\delta)\mu\bar{q}_S(\cdot) + \delta\gamma q_H(\cdot) - c_1^H(\cdot)] - c_0^H(\cdot)$.

Lemma 2.2: (i) $\frac{\partial\psi}{\partial a} = \psi'(a) > 0$ for all a in $[\bar{a}, A]$;

(ii) $\frac{\partial\psi}{\partial c_0^G} > 0$, $\frac{\partial\psi}{\partial c_1^G} > 0$ and $\frac{\partial\psi}{\partial Q_H} > 0$.

Proof: See the Appendix 2.A.

Lemma 2.2 implies that for a household with the ability of its child within the range $[\bar{a}, A]$ the amount of loss it incurs by not completing the child's school education inflates with the ability of the child. The higher-ability child receives a better quality of higher education degree and potentially can earn more in the skilled labor market which is sacrificed by not sending the child to school. For a child with ability a, the loss increases with a rise either in higher education or in school education investment by the government. It also rises with the increase in the average higher education quality.

Observation 2.4: (i) If $\psi(\bar{a}) \geq \underline{w}(1-\theta)$, all the children with their ability level $a \geq \bar{a}$ complete both school and higher education;

(ii) If $\psi(\bar{a}) < \underline{w}(1-\theta) < \psi(A)$, there exists a threshold of ability level $\hat{a} \in (\bar{a}, A]$ such that all the 2children with $a \geq \hat{a}$ complete both school and higher education; the children with $\bar{a} \leq a < \hat{a}$ join the unskilled labor market instead of completing schooling.

(iii) If $\underline{w}(1-\theta) \geq \psi(A)$, no children with their ability level $a \geq \bar{a}$ complete schooling.

Proof: See the Appendix 2.A.

Observation 2.4 shows that at the equilibrium of the model whether every child with $a \geq \bar{a}$ completes higher education or not, depends on the household's return from the option of not investing for the completion of the child's school education. Every child with $a \geq \bar{a}$ completes higher education if and only if the discounted wage at the unskilled labor market $\underline{w}(1-\theta)$ is sufficiently low. But with a higher value of $\underline{w}(1-\theta)$ even in the range $a \geq \bar{a}$, some children (below the threshold $\hat{a}$) will drop out of school. With a sufficiently high value of $\underline{w}(1-\theta)$ no one completes school education.

Case 2: $a < \bar{a}$.

The household chooses x_0 to maximize v_{S0} as in (2.7) subject to the budget constraint: $x_0 + c_0^H = y_0$. After substitution of $x_0 = y_0 - c_0^H$ in equation (2.7), v_{S0} is rewritten as:

$$v_{S0} = y_0 - c_0^H + \theta(y_1 + \underline{w} + \mu q_S)$$

which is a non-monotonic function of c_0^H. As the household maximizes v_{S0} by choosing $c_0^H > 0$, the choice satisfies the first-order condition:

$$\theta\mu\frac{\partial q_S}{\partial c_0^H}=1. \tag{2.16}$$

Equations (2.16) is solved for $c_0^H(Q_H,a,c_0^G,\mu,\theta)>0$. Since the child is not sent for higher education, the choice of c_0^H balances the marginal benefit from a higher quality of school education in terms of higher expected return from the semi-skilled labor market with the marginal cost that arises due to the sacrifice of period 0 consumption.

By substituting $c_0^H(\cdot)$ in equation (2.2), the equilibrium value of q_S is derived as

$$\tilde{q}_S=aT(Q_H)[\alpha_0(c_0^H(a,c_0^G,\mu,Q_H,\theta))^{\rho_0}+\beta_0(c_0^G)^{\rho_0}]^{\frac{1}{\rho_0}}\ . \tag{2.17}$$

Observation 2.5: (i) $\frac{\partial c_0^H}{\partial a}>0,\frac{\partial c_0^H}{\partial c_0^G}>0,\frac{\partial c_0^H}{\partial \mu}>0,\frac{\partial c_0^H}{\partial Q_H}>0,\frac{\partial c_0^H}{\partial \theta}>0;$

(ii) $\frac{\partial \tilde{q}_S}{\partial c_0^G}>0,\frac{\partial \tilde{q}_S}{\partial a}>0,\ \frac{\partial \tilde{q}_S}{\partial \mu}>0,\frac{\partial \tilde{q}_S}{\partial Q_H}>0,\frac{\partial \tilde{q}_S}{\partial \theta}>0$.

Proof: Similar to the proof Observation 2.3, follows from equations (2.16) and (2.17) above.

□

Unlike case 1 where $a\geq\bar{a}$, here, since the household does not send its child to higher education, the public investment in higher education, c_1^G , does not have any direct impact on c_0^H. A rise in the average higher education quality, Q_H, leads to a better overall teaching quality thus driving up household investment in school, c_0^H and the schooling quality, q_S. The household investment also gets positively affected by the child's ability, a, public investment in school education, c_0^G and responsiveness of semi-skilled wage to quality of school education, μ. Consequently, q_S of the child also improves. A rise in the discount factor, θ, also raises c_0^H and q_S.

A reduced form value of indirect utility, which is also the payoff of the household, in this case, is obtained by plugging $c_0^H(\cdot)$ and $\tilde{q}_S(\cdot)$ into equation (2.7) as:

$$\bar{v}_{S0} = y_0 - c_0^H(\cdot) + \theta(y_1 + \underline{w} + \mu\tilde{q}_S(\cdot)). \quad (2.18)$$

Comparing $\bar{v}_{S0}$ and $\bar{v}_{NS}$ from equation (2.18) and (2.11), we obtain:

$$\bar{v}_{S0} - \bar{v}_{NS} = \tau(a) - \underline{w}(1-\theta), \quad (2.19)$$

where $\tau(a) = \theta\mu\tilde{q}_S(\cdot) - c_0^H(\cdot)$.

Lemma 2.3: $\frac{\partial\tau}{\partial a} = \tau'(a) > 0$ for all values of a in $[0, \bar{a})$;

(ii) $\frac{\partial\tau}{\partial c_0^G} = \mu\frac{\partial\tilde{q}_S}{\partial c_0^G} > 0$, $\frac{\partial\tau}{\partial Q_H} = \mu\frac{\partial\tilde{q}_S}{\partial Q_H} > 0$ and $\frac{\partial\tau}{\partial c_1^G} = 0$.

Proof: See the Appendix 2.A.

Since a higher ability child in the range $[0, \bar{a})$ even if he does not join higher education earns a higher wage in the semi-skilled labor market, Lemma 2.3 suggests from equation (2.19) that the amount of loss incurred by a household from not sending the child to school inflates with the higher ability of the child. Given the ability of a child, the loss increases with the rise in public expenditure in school and average higher education quality.

Observation 2.6: (i) If $\tau(\bar{a}) > \underline{w}(1-\theta) \geq \tau(0)$, there exists an ability level $\tilde{a} \in [0, \bar{a})$ such that all the children with their ability level $a \geq \tilde{a}$ are sent for completion of schooling. The children with $a < \tilde{a}$ do not complete school.

(ii) If $\tau(\bar{a}) < \underline{w}(1-\theta)$, no children with their ability in $[0, \bar{a})$ completes school.

Proof: See the Appendix 2.A.

Observation 2.6 identifies the situations when a child who is not sent for higher education, is also not completing schooling.

How do $\psi(a)$ and $\tau(a)$ compare with each other?

Lemma 2.4: (i) $\psi(\bar{a}) = \tau(\bar{a})$;

(ii) $\psi'(\bar{a}) > \tau'(\bar{a})$.

Proof: See the Appendix 2.A.

From equations (2.15) and (2.19) it follows that $\psi(\bar{a}) - \tau(\bar{a}) = \bar{v}_{SH} - \bar{v}_{S0}$. But from lemma 2.1(i) we know that given the choice of c_0^H, at $\bar{a}$, $\bar{v}_{SH} - \bar{v}_{S0} = 0$. This explains the first part of lemma 2.4. Since the quality of higher education and the expected return from it rises with the ability of the child, for a child having an ability level marginally higher than $\bar{a}$, the opportunity cost of not sending him to higher education against sending him to the unskilled labor market is more than the opportunity cost of not sending him to the semi-skilled labor market. This intuitively explains the second part of the lemma.

Observation 2.2, 2.4, 2.6, and lemma 2.4 discussed above allow us to identify three possible equilibrium situations.

Observation 2.7: (i) If $\psi(A) > \psi(\bar{a}) = \tau(\bar{a}) > \underline{w}(1-\theta) \geq \tau(0)$, children with ability level $a < \tilde{a}$ do not complete schooling, children with ability within the range $[\tilde{a}, \bar{a})$ complete schooling but do not proceed with higher education, while children with ability level $a \geq \bar{a}$ continue with higher education.

(ii) If $\psi(A) > \underline{w}(1-\theta) > \psi(\bar{a}) = \tau(\bar{a}) > \tau(0)$, all the children with $a \geq \hat{a}$ complete both school and higher education; the rest joins the unskilled labor market.

(iii) If $\underline{w}(1-\theta) > \psi(A) > \psi(\bar{a}) = \tau(\bar{a}) > \tau(0)$, all the children in the range $[0, A]$ go to the unskilled labor market.

Proof: Follows from Observations 2.4 and 2.6.

Since there always exists some children who participate in higher education (by Assumption 2.2), we can ignore case (iii). While in case (i) the equilibrium higher education quality is $Q_H^0 =$

$\frac{1}{A}\int_{\bar{a}}^{A} q_H(Q_H^0, q_S(Q_H^0),.)da$, in (ii) the equilibrium average higher education quality is $Q_H^0 = \frac{1}{A}\int_{\hat{a}}^{A} q_H(Q_H^0, q_S(Q_H^0),.)da$.[23]

Lemma 2.5: $\frac{\partial Q_H^0}{\partial c_0^G} > 0$; $\frac{\partial Q_H^0}{\partial c_1^G} > 0$.

Proof: See the Appendix 2.A.

A rise in investment in higher education raises the higher education quality of the existing students. At the same time, it raises enrolment in higher education at the extensive margin. Both constitute the first-order effects of a rise in higher education expenditure by the government. There is a positive multiplier effect of the rise in higher education quality which operates through the improvement of teaching-quality in schools. An improvement in the quality of schooling positively affects the average quality of higher education and thus, the teaching quality at schools again. Similarly, a rise in government investment in schools raises school quality and provides an impetus to the quality of higher education as it has a better quality of students to start with. Also, there is an increase in participation in higher education. This raises the overall higher education quality with a feedback effect on teaching quality at schools. The subsequent rounds follow as in the previous case.

[23]For case (ii), the condition for existence of Q_H^0 is derived from: $Q_H = \frac{1}{A}\int_{\hat{a}}^{A} q_H(Q_H, q_S(Q_H),.)da$, where $q_H(Q_H, q_S(Q_H),.)$ is the value of q_H from equation (2.1) after solving the period 2 problem of the household's decision making. Given the assumptions of the model it is likely to be the case that the RHS of the above equation is a continuous, concave function with non-negative intercept. Therefore, the condition for existence of a unique non-zero Q_H^0 is given by: $\int_{\hat{a}}^{A} \frac{\partial q_H}{\partial Q_H} + \frac{\partial q_H}{\partial c_1^H}\frac{\partial c_1^H}{\partial Q_H} + \frac{\partial q_H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial Q_H} + \frac{\partial q_H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial c_0^H}\frac{\partial c_0^H}{\partial Q_H} + \frac{\partial q_H}{\partial c_1^H}\frac{\partial c_1^H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial Q_H} + \frac{\partial q_H}{\partial c_1^H}\frac{\partial c_1^H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial c_0^H}\frac{\partial c_0^H}{\partial Q_H} da - q_H(\hat{a})\frac{\partial \hat{a}}{\partial Q_H} < A$.

Observation 2.7 (ii) discussed above allows us to identify a benchmark equilibrium situation in the economy described below which we use to derive the results of the model.

2.3.1.1. The Benchmark Situation: $[\psi(A) > \underline{w}(1-\theta) > \psi(\bar{a}) = \tau(\bar{a}) > \tau(0)]$

We define the benchmark situation in the economy as the one where all the children who complete their school education also join higher education. Figure 2.2 describes the benchmark situation.

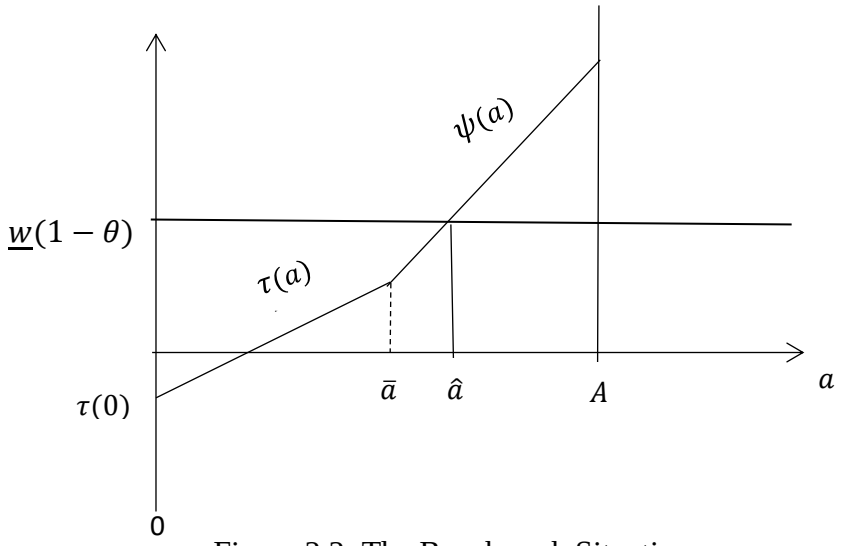


Figure 2.2: The Benchmark Situation

Observation 2.7(ii), by observations 2.2 and 2.4 imply that in such a situation all the children with their ability in the range $[\hat{a}, A]$ complete both school and higher education; the children with their ability in the range $[0, \hat{a})$ do not complete schooling and drop out. As defined earlier, the proportion of children in the economy completing school education and the number of students entering higher education are given by $E_S^0 = E_H^0 = 1 - \frac{\hat{a}}{A}$. The transition rate is 100%.

The proportion of children in the economy who drops out of school and joins the unskilled labor market is given by $L_u^0 = \frac{\hat{a}}{A}$.

The proportion of children who joins the semi-skilled labor market is given by $L_{SS}^0 = (1-\delta)(1-\frac{\hat{a}}{A})$. All the households, whose child works in the semi-skilled labor market after completing higher education, overspend on their child's education.

The quality of higher education is given by $Q_H^0 = \frac{1}{A}\int_{\hat{a}}^{A} q_H(Q_H^0, q_S(Q_H^0),.)da$.

2.3.2 The government's budget constraint

Suppose the government has an outlay of M for allocation on school education and higher education.

$$c_0^G + \left(1-\frac{\hat{a}}{A}\right)c_1^G = M. \tag{2.20}$$

Given M, from equation (2.20) it follows:

$$dc_1^G = -[\left(\frac{A}{A-\hat{a}}\right) - \frac{c_1^G}{A-\hat{a}}\frac{d\hat{a}}{dc_0^G}]dc_0^G. \tag{2.21}$$

From $\psi(\hat{a}) = \underline{w}(1-\theta)$, $\frac{d\hat{a}}{dc_0^G} < 0$ using Lemma 2.2, and therefore, in equation (2.21), $\left[\left(\frac{A}{A-\hat{a}}\right) - \frac{c_1^G}{A-\hat{a}}\frac{d\hat{a}}{dc_0^G}\right] > 0$.

2.4 The Results

Proposition 2.1: *Starting from the benchmark situation where no one joins the semi-skilled labor market after completion of school education, by reallocation of the education budget of the government from higher education to school education, if the conditions* $\frac{\partial \tilde{q}_S}{\partial c_0^G} + \frac{\partial \tilde{q}_S}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_0^G} - \frac{\partial \tilde{q}_S}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_1^G}\left(\left(\frac{A}{A-\hat{a}}\right) - \frac{c_1^G}{A-\hat{a}}\frac{d\hat{a}}{dc_0^G}\right) > 0$ *and* $w_I\left(\tilde{q}_S(c_0^{G\prime}, c_1^{G\prime}, \hat{a})\right) > k_1$ *hold, it is possible to implement an equilibrium in which (i) a proportion of children join the semi-skilled labor market after completion of school education; (ii) the school-dropout rate falls; (iii) enrolment in higher*

education falls; and (iv) the proportion of household over-spending in higher education falls. Otherwise, the benchmark equilibrium continues to hold.

For an equilibrium of the model to exist where the outcome (i) of proposition 1 holds (let us call it equilibrium profile 1), observation 2.7(i) implies that the following condition must hold: $\psi(A) > \psi(\bar{a}) = \tau(\bar{a}) > \underline{w}(1-\theta) \geq \tau(0)$.

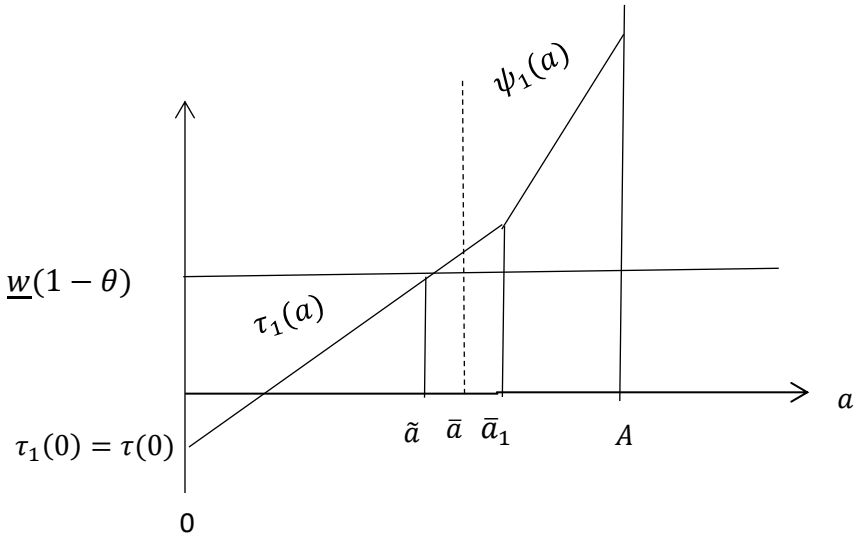


Figure 2.3: The equilibrium profile 1

In proposition 2.1, government reallocates its fixed educational budget from higher education to school education such that $(c_0^{G\prime}, c_1^{G\prime})$ is the new allocation. Consequently, by implications of lemma 2.2 and 2.3, while the positions of the $\tau(a)$ and $\psi(a)$ curve change to $\tau_1(a)$ and $\psi_1(a)$ respectively in figure 2.3 compared to figure 2.2, the position of $\underline{w}(1-\theta)$ remains unchanged. Then, it must be the case that the value of $\bar{a}$ in Figure 2.3 (call it $\bar{a}_1$) is greater than the value of $\bar{a}$ in Figure 2.2. Also, $\hat{a}$ in figure 2.2 (which indicates the threshold for determining higher education enrolment and school completion) must also be lower than $\bar{a}_1$ and greater than $\tilde{a}$. Notice that as $a \to 0$, $\tau_1(0) = \tau(0) = -c_0^H\left(c_0^{G\prime}, c_1^{G\prime}, 0\right)$ remains unchanged in the new situation.

Since the ability is distributed uniformly in $[0, A]$ and from observation 2.7(i) the households with the ability of their child lying in the range $[0, \tilde{a})$ do not invest in the completion of their child's school education, the school drop-out rate is given by $L_u^1 = \frac{\tilde{a}}{A}$. The school completion rate is given by $E_S^1 = \frac{A-\tilde{a}}{A} = 1 - \frac{\tilde{a}}{A}$.

Also, by observation 2.7(i), the enrolment ratio in higher education at equilibrium profile 1 is calculated as $E_H^1 = 1 - \frac{\bar{a}_1}{A}$.

Since $\tilde{a} < \bar{a}_1$ in figure 2.3, clearly, $E_H^1 < E_S^1$. The proportion of children who after completing school education joins the semi-skilled labor market is $E_S^1 - E_H^1 = \frac{\bar{a}_1 - \tilde{a}}{A} > 0$.

The proportion of children who join the semi-skilled labor market after completion of higher education is given by $L_{SS}^1 = (1-\delta)(1 - \frac{\bar{a}_1}{A})$. The households to which these children belong overspend on their child's education.

The quality of higher education is determined by the equation $Q_H^1 = \frac{1}{A}\int_{\bar{a}_1}^{A} q_H da$.

Notice that equilibrium profile 1 satisfies its features (ii), (iii), and (iv) described in proposition 2.1 since $L_u^0 - L_u^1 = \frac{\hat{a}-\tilde{a}}{A} > 0$, $E_H^0 - E_H^1 = \frac{\bar{a}_1 - \hat{a}}{A} > 0$ and $L_{SS}^0 - L_{SS}^1 = (1-\delta)\left(\frac{\bar{a}_1 - \hat{a}}{A}\right) = (1-\delta)(E_H^0 - E_H^1) > 0$.

For deriving the conditions under which Proposition 2.1 holds, we juxtapose Figure 2.2 representing the benchmark equilibrium, and Figure 2.3 representing equilibrium profile 1 in Figure 2.4 below.

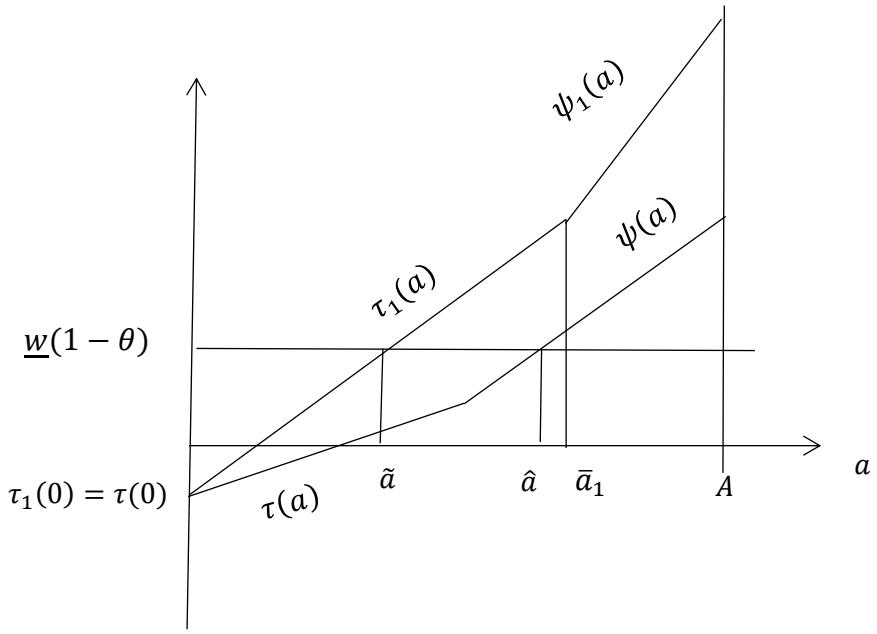


Figure 2.4: Juxtaposition of benchmark

Figure 2.4 shows that due to a rise in c_0^G and a corresponding fall in c_1^G, compared to the benchmark equilibrium both $\tau(a)$ and $\psi(a)$ shift in an upward direction. Also, the sufficient condition under which proposition 2.1 holds, is: $\tau_1(\hat{a}) > \underline{w}(1-\theta)$. The derivation of conditions in Proposition 1 is relegated to the appendix 2.A.

The intuition behind proposition 2.1 is simple. The reallocation of the government budget in favor of school education serves the purpose of inducing children to join the semi-skilled labor market after completion of school, if the net effect of budget reallocation improves the schooling quality of the child belonging to a household, who was indifferent between sending its child to the semi-skilled labor market and sending him for higher education, to such a level that it expects to receive sufficiently high wage by working at the semi-skilled labor market. The wage must cover not only the returns from joining the unskilled labor market, $\frac{w}{\theta}$, also the

investment made by the household for school completion $\frac{c_0^H(\,c_0^{G\prime},c_1^{G\prime},\hat{a})}{\theta}$. The term $\left(\frac{\partial\tilde{q}_S}{\partial c_0^G}+\frac{\partial\tilde{q}_S}{\partial Q_H}\,\frac{\partial Q_H^0}{\partial c_0^G}\right)$ represents the direct and indirect effects of increasing budget allocation in school education on schooling quality, both of which are positive. The direct effect is generated as the school infrastructure improves. The indirect effect is generated as better schooling quality helps to improve higher education quality, which has a feedback effect on schooling quality as the quality of teachers improves. The term $\frac{\partial\tilde{q}_S}{\partial Q_H}\,\frac{\partial Q_H^0}{\partial c_1^G}\left(\left(\frac{A}{A-\hat{a}}\right)-\frac{c_1^G}{A-\hat{a}}\,\frac{d\hat{a}}{dc_0^G}\right)$ represents the negative effect of the reduction of budget in higher education on schooling quality that works through a fall in teaching quality. For the budget reallocation to be effective it is necessary that the net of the above two effects is positive. Otherwise, the household does not even consider the completion of school education of its child as an option and the budget reallocation policy fails. The reallocation policy reduces both the school-dropout rate and the enrolment in higher education. The reduction in enrolment in higher education leads to a fall in the number of households that overspend on higher education.

Proposition 2.2: *A budgetary reallocation from higher education to school education cannot eliminate the problem of school-dropout.*

The intuition behind the result is the following. A household completes the school education of its child only if by doing so it enjoys a higher utility compared to not completing his school education. Now consider the case of the households with the ability of their child close to zero. A child with the minimum ability, owing to assumption 2.1 of the model, even if completes his school education and participates in the semi-skilled/high-skilled labor market, earns $\underline{w}$, the wage in the unskilled labor market. But since the completion of school education costs $c_0^H > 0$, the household enjoys a life-time utility of $[y_0 - c_0^H(.) + \theta(y_1 + \underline{w})]$ by continuing with the school-education of the child. In contrast, if the household allows the child to drop out of

school, he participates in the un-skilled labor market and earns $\underline{w}$. In such a case, the household's life time utility is $[y_0 + \underline{w} + \theta y_1]$, which clearly exceeds $[y_0 - c_0^H(.) + \theta(y_1 + \underline{w})]$ by $[\underline{w}(1-\theta) + c_0^H(.)]$ independent of the allocation of the government budget between school and higher education. Therefore, the budget reallocation policy although can reduce the problem of school drop-out as suggested by proposition 2.1, it would never be able to eliminate it as suggested by proposition 2.2.

2.5 Conclusions and Discussion

The chapter derives conditions satisfying which a government can achieve a reduced rate of school dropout and solve the problem of overeducated labor in the semi-skilled labor market by reallocation of its education budget between school and higher education. The theoretical model developed in the chapter takes account of the complementarities that exist between household and institutional investment at the school and the higher education, on the one hand, and between school and higher education quality, on the other, which is unique in the literature.

The results derived in the chapter have important implications for educational policy-making in developing countries. First, for a conventional policy of budget-reallocation towards schooling to work successfully in solving the problems of school dropout and overeducation, it is important to satisfy two conditions: (1) schooling-quality responds more to better school infrastructure than to teaching quality; and (2) the semi-skilled wage rate responds sufficiently to schooling-quality. While biasing the expenditure in favor of school education the policy makers often ignore the fact that schoolteachers are recruited from college graduates. Therefore, a fall in higher education quality may adversely impact teaching quality in schools, which may not be completely recouped by the rising investment in school infrastructure. In developing countries, the quality of education being poor, a marginal improvement of higher

education quality may have a large impact on improving the school education quality, rendering budgetary reallocation policy in favor of school education ineffective. Our results also suggest that the elimination of dropouts is not feasible through a budget reallocation policy. The results underline the importance of responsiveness of semi-skilled wage rate to schooling-quality. In this way, it implicitly supports the importance of vocational training in the school curriculum. Therefore, our results theoretically justify policies undertaken in some developing countries for strengthening their vocational training programs. For example, the National Education Policy (NEP, 2020) formulated by the Government of India has designed an ambitious plan of exposing 50 percent of school and higher secondary students to vocational training by 2025.

Like all theoretical models, the present model also has certain limitations. First, the model assumes identical income households with the uniformly distributed ability of their child. At the equilibrium, the households self-select themselves according to the ability of their child. However, in reality, the households differ also in their income level. It can be checked how the distribution of income and the distribution of ability interact with each other in the decision about the education of their child. Second, the model divides the entire education life of a child into two broad categories: school education and higher education. However, more divisions of both are possible. For example, school education could be divided between regular school and vocational training and higher education between undergraduate and postgraduate without changing the basic intuition of the model.

The present chapter can be extended in many directions. First, presently the model links educational investment decisions to the supply side of the labor market. Once the demand side is brought in, one can analyze the effect of demand-side shocks like the changing profile of Foreign Direct Investment on the education sector of an economy. Second, one can take up empirical studies to ascertain the effect of higher education quality on the school teaching

quality, the weak sensitivity of which is the necessary condition for the success of budget reallocation policy in solving the dual problem of school-dropout and overeducation in this chapter. These remain on our future research agenda.

Appendix 2.A

Proof of Observation 2.1: (i) From Equation (2.5):

$$\frac{\partial c_1^H}{\partial a} = -\frac{\frac{\partial^2 q_H}{\partial c_1^H \partial a}}{\frac{\partial^2 q_H}{\partial c_1^{H^2}}} > 0. \quad (2.A.1)$$

Since $\frac{\partial^2 q_H}{\partial c_1^{H^2}} < 0$ and $\frac{\partial^2 q_H}{\partial c_1^H \partial a} > 0$ from equation (2.1), the RHS of (2.A.1) is positive. The statement follows. The proof of the rest of the statements is similar.

(ii) From equation (2.1):

$$\frac{\partial q_H}{\partial a} = \frac{q_H}{a} + a q_S [\alpha_1 (c_1^H(.))^{\rho_1} + \beta_1 (c_1^G)^{\rho_1}]^{\frac{1}{\rho_1}-1} \alpha_1 (c_1^H(.))^{\rho_1 - 1} \frac{\partial c_1^H}{\partial a},$$

which is positive since $\frac{\partial c_1^H}{\partial a} > 0$ from above. The proof of the rest of the statements is similar.

□

Proof of Lemma 2.1: (i) From equation (2.8), we have,

$$\frac{\partial \varphi}{\partial a} = \theta \delta \gamma \frac{\partial q_H}{\partial a} + \theta \frac{\partial c_1^H}{\partial a} \left(\delta \gamma \frac{\partial q_H}{\partial c_1^H} - 1 \right). . \quad (2.A.2)$$

By application of equation (2.5) in (2.A.2):

$$\frac{\partial \varphi}{\partial a} = \theta \delta \gamma \frac{\partial q_H}{\partial a} > 0. \quad (2.A.3)$$

Since $\frac{\partial q_H}{\partial a} > 0$ the RHS of (2.A.3) is positive.

(ii) From equation (2.8),

$$\frac{\partial \varphi}{\partial c_1^G} = \theta [\delta \gamma \frac{\partial q_H}{\partial c_1^G} + \frac{\partial c_1^H}{\partial c_1^G} \left(\delta \gamma \frac{\partial q_H}{\partial c_1^H} - 1 \right)]. \quad (2.A.4)$$

By application of equation (2.5) in (2.A.4):

$\frac{\partial \varphi}{\partial c_1^G} = \theta\delta\gamma \frac{\partial q_H}{\partial c_1^G} > 0.$

Since $\frac{\partial q_H}{\partial c_1^G} > 0$ by Equation (2.1), the RHS of the above equation is positive.

The proof of rest of Lemma 2.1 follows similarly. □

Proof of Observation 2.2: (i) As $a \to 0$, since $q_H \to 0$ from equation (2.1) and $q_S \to 0$ from equation (2.2), from equation (2.5) it follows that $\varphi(a) \to (-\theta c_1^H(\cdot)) < 0$. Since $\varphi(a)$ is a continuous monotonically increasing function over $[0, A]$ from lemma 2.1, if $\varphi(A) > 0$, the statement of the first part of the observation follows from the definition of φ.

(ii) If $\varphi(A) < 0$, for all values of a in $[0, A]$, $\varphi(a) < 0$. Therefore, the statement of the second part of the observation follows from the definition of φ. □

Proof of Observation 2.3: (i) From (2.12):

$$\frac{\partial c_0^H}{\partial a} = -\frac{\frac{\partial^2 q_S}{\partial c_0^H \partial a}\left(\mu(1-\delta)+\delta\gamma.\frac{\partial q_H}{\partial q_S}\right)+\frac{\partial q_S}{\partial c_0^H}\delta\gamma.\frac{\partial^2 q_H}{\partial q_S \partial a}+\frac{\partial q_S}{\partial c_0^H}\delta\gamma.\frac{\partial^2 q_H}{\partial q_S \partial c_1^H}\frac{\partial c_1^H}{\partial a}}{\frac{\partial^2 q_S}{\partial {c_0^H}^2}\left(\mu(1-\delta)+\delta\gamma.\frac{\partial q_H}{\partial q_S}\right)}. \quad (2.A.5)$$

Since from equation (2.2) $\frac{\partial^2 q_S}{\partial {c_0^H}^2} < 0, \frac{\partial^2 q_S}{\partial c_0^H \partial a} > 0,$ from equation (2.1) $\frac{\partial q_H}{\partial q_S} > 0,$ and from Observation 2.1 (i) $\frac{\partial c_1^H}{\partial a} > 0$, the RHS of equation (2.A.5) is positive, the proof of the rest of first part of the observation is similar

(ii) From Equation (2.13),

$\frac{\partial \bar{q}_S}{\partial a} = \frac{\bar{q}_S}{a} + aT(Q_H)[\alpha_0(c_0^H(.))^{\rho_0} + \beta_0(c_0^G)^{\rho_0}]^{\frac{1}{\rho_0}-1}\alpha_0(c_0^H(.))^{\rho_0-1}\frac{\partial c_0^H}{\partial a}$ which is positive by application of Observation 2.3(i). The proof of the rest of the second part of the observation is similar. □

Proof of Lemma 2.2: (i) From the definition of $\psi(a)$,

$$\psi'(a) = \frac{\partial \bar{v}_{SH}}{\partial a} = \theta\gamma\delta \frac{\partial q_H}{\partial a} + \theta\left[\gamma\delta \frac{\partial q_H}{\partial q_S} + \mu(1-\delta)\right]\left(\frac{\partial \bar{q}_S}{\partial a}\right) + \frac{\partial c_0^H}{\partial a}\left(\theta \frac{\partial q_S}{\partial c_0^H}\left(\mu(1-\delta) + \delta\gamma.\frac{\partial q_H}{\partial q_S}\right) - 1\right) + \theta\left(\frac{\partial c_1^H}{\partial a} + \frac{\partial c_1^H}{\partial q_S}.\frac{\partial \bar{q}_S}{\partial c_0^H}.\frac{\partial c_0^H}{\partial a} + \frac{\partial c_1^H}{\partial q_S}.\frac{\partial \bar{q}_S}{\partial a}\right)\left(\delta\gamma.\frac{\partial q_H}{\partial c_1^H} - 1\right). \quad (2.A.6)$$

By application of equations (2.5) and (2.12) in (2.A.6):

$$\psi'(a) = \frac{\partial \bar{v}_{SH}}{\partial a} = \theta\gamma\delta \frac{\partial q_H}{\partial a} + \theta\left[\gamma\delta.\frac{\partial q_H}{\partial q_S} + \mu(1-\delta)\right](\frac{\partial \bar{q}_S}{\partial a}) > 0$$

The statement of the first part of the lemma follows from Equation (2.14).

(ii) From the definition of $\psi(a)$,

$$\frac{\partial \psi}{\partial c_0^G} = \frac{\partial \bar{v}_{SH}}{\partial c_0^G} = \theta\left[\gamma\delta \frac{\partial q_H}{\partial q_S} + \mu(1-\delta)\right](\frac{\partial \bar{q}_S}{\partial c_0^G}) + \frac{\partial c_0^H}{\partial c_0^G}\left(\theta \frac{\partial q_S}{\partial c_0^H}\left(\mu(1-\delta) + \delta\gamma.\frac{\partial q_H}{\partial q_S}\right) - 1\right) + \theta(\frac{\partial c_1^H}{\partial q_S}.\frac{\partial \bar{q}_S}{\partial c_0^H}.\frac{\partial c_0^H}{\partial c_0^G} + \frac{\partial c_1^H}{\partial q_S}.\frac{\partial \bar{q}_S}{\partial c_0^G})(\delta\gamma.\frac{\partial q_H}{\partial c_1^H} - 1)$$

By application of equations (2.5) and (2.12)

$$\frac{\partial \psi}{\partial c_0^G} = \frac{\partial \bar{v}_{SH}}{\partial c_0^G} = \theta\left[\gamma\delta \frac{\partial q_H}{\partial q_S} + \mu(1-\delta)\right]\left(\frac{\partial \bar{q}_S}{\partial c_0^G}\right) = \theta\mu \frac{\partial \bar{q}_S}{\partial c_0^G} + \theta\delta \frac{\partial \bar{q}_S}{\partial c_0^G}\left(\gamma \frac{\partial q_H}{\partial q_S} - \mu\right) > 0 .$$

The statement follows from the implication of equation (2.13) and Assumption 2.1.

By similar logic, $\frac{\partial \psi}{\partial c_1^G} = \theta\delta\gamma.\frac{\partial q_H}{\partial c_1^G} > 0$ and $\frac{\partial \psi}{\partial Q_H} = \theta\mu \frac{\partial \bar{q}_S}{\partial Q_H} + \theta\,\delta \frac{\partial \bar{q}_S}{\partial Q_H}\left(\gamma \frac{\partial q_H}{\partial q_S} - \mu\right) + \theta\delta\gamma \frac{\partial q_H}{\partial Q_H} > 0.$

□

Proof of Observation 2.4: A child with $a \geq \bar{a}$ is sent to higher education if and only if $\bar{v}_{SH} \geq \bar{v}_{NS}$ i.e. $\psi\,(a) \geq \underline{w}(1-\theta)$. Since from Lemma 2.2, $\psi\,(a)$ is continuous and monotonically rising in a for all values of $a \in [\bar{a}, A]$, if $\psi(\bar{a}) \geq \underline{w}(1-\theta)$, it must be true that $\psi\,(a) \geq \underline{w}(1-\theta)$. for all values of $a \in [\bar{a}, A]$. Therefore, the first part of the observation follows. However, if $\psi(\bar{a}) < \underline{w}(1-\theta) < \psi(A)$, lemma 2.2 implies that there exists a threshold of ability level

$\hat{a} \in (\bar{a}, A]$ and for all $a \geq \hat{a}$, $\psi(a) \geq \underline{w}(1-\theta)$. holds. For all values of $a < \hat{a}$, $\psi(a) < \underline{w}(1-\theta)$. holds. Therefore, the second part of the observation follows. A similar argument follows for the third part of the observation as well. □

Proof of Observation 2.5: (i) From Equation (2.16), we have

$$\frac{\partial c_0^H}{\partial a} = -\frac{\frac{\partial^2 q_S}{\partial c_0^H \partial a}}{\frac{\partial^2 q_S}{\partial {c_0^H}^2}}.$$

The above statement follows from equation (2.2) and other assumptions of the model. The proof of the rest of the first part of the observation is similar.

(ii) Similar to Observation 2.3 (ii). □

Proof of Lemma 2.3: (i) From the definition of $\tau(a)$,

$$\tau'(a) = \frac{\partial \bar{v}_{S0}}{\partial a} = \theta\mu \frac{\partial \tilde{q}_S}{\partial a} + \left(\frac{\partial c_0^H}{\partial a}\right)\left(\theta\mu \frac{\partial \tilde{q}_S}{\partial c_0^H} - 1\right)$$

Using the first-order condition (2.16), $\tau'(a) = \frac{\partial \bar{v}_{S0}}{\partial a} = \theta\mu \frac{\partial \tilde{q}_S}{\partial a} > 0$

The above statement follows from equation (2.17).

(ii) From the definition of $\tau(a)$,

$$\frac{\partial \tau}{\partial c_0^G} = \frac{\partial \bar{v}_{S0}}{\partial c_0^G} = \theta\mu\left(\frac{\partial \tilde{q}_S}{\partial c_0^G}\right) + \frac{\partial c_0^H}{\partial c_0^G}\left(\theta\mu \frac{\partial \tilde{q}_S}{\partial c_0^H} - 1\right)$$

Using the first-order condition (2.16), $\frac{\partial \tau}{\partial c_0^G} = \frac{\partial \bar{v}_{S0}}{\partial c_0^G} = \theta\mu \frac{\partial \tilde{q}_S}{\partial c_0^G} > 0$. (By equation (2.17))

By a similar argument, $\frac{\partial \tau}{\partial c_1^G} = 0$ and $\frac{\partial \tau}{\partial Q_H} = \theta\mu \frac{\partial \tilde{q}_S}{\partial Q_H}$ □

Proof of Observation 2.6: A child with $a < \bar{a}$ is sent for completion of school if and only if $\bar{v}_{S0} \geq \bar{v}_{NS}$ i.e. $\tau(\bar{a}) \geq \underline{w}(1-\theta)$. Since $\underline{w}(1-\theta) > 0$ and independent of a, lemma 2.3

implies if $\tau(\bar{a}) > \underline{w}(1-\theta) \geq \tau(0)$ holds, there must exist a value of a in $[0, \bar{a})$ such that $\tau(\tilde{a}) = \underline{w}(1-\theta)$. holds. The first part of the statement of the observation follows. The proof of the second and third parts of the observation also follows from the application of lemma 2.3.

□

Proof of Lemma 2.4: (i) Under assumption 2.2, from observation 2.1 we know that $\bar{a} \in [0, A]$ solves $\varphi(\bar{a}) = 0$. This must also hold in the final equilibrium,

$\theta[\,\mu(1-\delta)\mu\bar{q}_S(\cdot) + \delta\gamma q_H(\bar{q}_S(\cdot), a, c_1^H(\cdot), c_1^G) - c_1^H(\cdot)] - c_0^H(\cdot) - [\theta[\mu\bar{q}_S(\cdot)] - c_0^H(\cdot)] = 0$

(Since they are indifferent between higher and no higher education at $\bar{a}$, $c_0^H(\cdot)$ is the same in both. Thus q_S is also the same in both)

Thus, at $\bar{a}$ it must be $\psi(\bar{a}) = \tau(\bar{a})$.

(ii) At $a = \bar{a}$, using Lemma 2.2 (i) and Lemma 2.3 (i),

$\psi'(\bar{a}) - \tau'(\bar{a}) = \theta\gamma\delta\frac{\partial q_H}{\partial a} + \theta\delta\frac{\partial \bar{q}_S}{\partial a}\left(\gamma.\frac{\partial q_H}{\partial q_S} - \mu\right) > 0$ under a similar logic in Lemma 2.4(i).

□

Proof of Lemma 2.5:

Plugging in the optimum values of $\bar{q}_S(\cdot)$ and $c_0^H(\cdot)$ from equations (2.12) and (2.13) into equation (2.1), we get the optimum value of $q_H(c_0^G, a, c_1^G, \delta, \gamma, \mu, Q_H)$

Thus, by equation (2.9), we have $Q_H^0 = \frac{1}{A}\int_{\hat{a}}^{A} q_H(c_0^G, a, c_1^G, \delta, \gamma, \mu, Q_H^0)da$.

$$\frac{\partial Q_H^0}{\partial c_0^G} = \frac{\int_{\hat{a}}^{A}(\frac{\partial q_H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial c_0^G} + \frac{\partial q_H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial c_0^H}\frac{\partial c_0^H}{\partial c_0^G} + \frac{\partial q_H}{\partial c_1^H}\frac{\partial c_1^H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial c_0^G} + \frac{\partial q_H}{\partial c_1^H}\frac{\partial c_1^H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial c_0^H}\frac{\partial c_0^H}{\partial c_0^G})da - q_H(\hat{a})\frac{\partial \hat{a}}{\partial c_0^G}}{A - (\int_{\hat{a}}^{A}\frac{\partial q_H}{\partial Q_H} + \frac{\partial q_H}{\partial c_1^H}\frac{\partial c_1^H}{\partial Q_H} + \frac{\partial q_H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial Q_H} + \frac{\partial q_H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial c_0^H}\frac{\partial c_0^H}{\partial Q_H} + \frac{\partial q_H}{\partial c_1^H}\frac{\partial c_1^H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial Q_H} + \frac{\partial q_H}{\partial c_1^H}\frac{\partial c_1^H}{\partial q_S}\frac{\partial \bar{q}_S}{\partial c_0^H}\frac{\partial c_0^H}{\partial Q_H} - q_H(\hat{a})\frac{\partial \hat{a}}{\partial Q_H})}. \quad (2.A.7)$$

From Observation 2.7 (ii), at $\hat{a} \in (\bar{a}, A]$ we have, $\psi(\hat{a}) = \underline{w}(1-\theta)$. Thus,

$\frac{\partial \hat{a}}{\partial c_0^G} = -\frac{\theta\mu\frac{\partial \bar{q}_S}{\partial c_0^G} + \theta\delta\left(\gamma\frac{\partial q_H}{\partial q_S} - \mu\right)\frac{\partial \bar{q}_S}{\partial c_0^G}}{\psi^{'}(a)} < 0$ by Assumption 2.1, Lemma 2.2, and equation (2.13). Using this expression in (2.A.7), we see that $\frac{\partial Q_H^0}{\partial c_0^G} > 0$

The second part of the proof follows similarly □

Derivation of conditions in Proposition 2.1:

The necessary condition for proposition 2.1 to hold, is: $\left(\frac{\partial \tau}{\partial c_0^G} + \frac{\partial \tau}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_0^G}\right)dc_0^G + (\frac{\partial \tau}{\partial c_1^G} + \frac{\partial \tau}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_1^G})dc_1^G > 0.$

From lemma 2.3(ii), lemma 2.5, and equation (2.21) it follows:

$\left(\frac{\partial \tau}{\partial c_0^G} + \frac{\partial \tau}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_0^G}\right)dc_0^G + (\frac{\partial \tau}{\partial c_1^G} + \frac{\partial \tau}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_1^G})dc_1^G = \mu(\frac{\partial \tilde{q}_S}{\partial c_0^G} + \frac{\partial \tilde{q}_S}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_0^G} - \frac{\partial \tilde{q}_S}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_1^G}[\left(\frac{A}{A-\hat{a}}\right) - \frac{c_1^G}{A-\hat{a}}\frac{d\hat{a}}{dc_0^G}])dc_0^G.$

Therefore, $\left(\frac{\partial \tau}{\partial c_0^G} + \frac{\partial \tau}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_0^G}\right)dc_0^G + (\frac{\partial \tau}{\partial c_1^G} + \frac{\partial \tau}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_1^G})dc_1^G > 0$ if and only if $\frac{\partial \tilde{q}_S}{\partial c_0^G} + \frac{\partial \tilde{q}_S}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_0^G} - \frac{\partial \tilde{q}_S}{\partial Q_H}\frac{\partial Q_H^0}{\partial c_1^G}[\left(\frac{A}{A-\hat{a}}\right) - \frac{c_1^G}{A-\hat{a}}\frac{d\hat{a}}{dc_0^G}] > 0.$

From figure 2.4, the sufficient condition under which proposition 2.1 holds, is: $\tau_1(\hat{a}) > \underline{w}(1-\theta)$. Substituting for $\tau_1(\hat{a})$, by application of assumption 2.1 of the model, the condition can be written as:

$\theta\mu\tilde{q}_S(\cdot) - c_0^H(\cdot) > \underline{w}(1-\theta).$

Or $\theta w_I\left(\tilde{q}_S(c_0^{G'}, c_1^{G'}, \hat{a})\right) - c_0^H(c_0^{G'}, c_1^{G'}, \hat{a}) > \underline{w}$

Or $w_I\left(\tilde{q}_S(c_0^{G'}, c_1^{G'}, \hat{a})\right) > k_1$

where $k_1 = \frac{\underline{w}}{\theta} + \frac{c_0^H(c_0^{G\prime}, c_1^{G\prime}, \hat{a})}{\theta} > 0$. □

Proof of Proposition 2.2: It follows from Observation 2.6 (i), the implementation of no school dropout requires satisfaction of the following condition: $\tau(0) > \underline{w}(1-\theta) \geq 0$. However, $\tau(0) = -c_0^H(\cdot) < 0$ for all possible values of (c_0^G, c_1^G). The statement of the proposition follows. □

CHAPTER 3

DOES PRIVATE TUITION CROWD OUT PRIVATE SCHOOLING?: EVIDENCE FROM INDIA

3.1.Introduction

The private tuition market has developed in several countries across the world as a supplement to the formal education system. Two characteristics associated with this market as discussed by Bray (1999) are supplementation – these services are intended to supplement students' formal lessons, and privateness – they are offered by private individuals or enterprises mainly for the profit motive. The proliferation of this market has stemmed from the various deficiencies in mainstream education. While ample factors drive demand for private tuition including peer pressure, intense student competition, and high-stakes examination, deficiencies in teaching quality emerge as a primary source of demand (Bray and Lykins, 2012; Sen, 2010). Given the availability of such a market, students can easily supplement the paucity of classroom teaching with the help of private tutors[24]. A 2008 survey on private centres offering coaching to students estimated the private tuition sector to evolve as a thriving business in India worth $6.4 billion, growing at an annual rate of 15 percent (Vora and Dewan, 2009). The ASSOCHAM (2013) survey covering 12 major cities in India shows that around 92 percent of tutors reported that parents depend on the private tuition market to compensate for the perceived deficiencies in the state school system.[25] The schools imparting good quality teaching extinguish the need for private tuition. But this may not imply that the children are always sent to schools imparting good teaching quality. Since the deficiency in teaching quality can be compensated by private tuition, the parents may choose to send their child to a low-cost

[24] The literature frequently uses the nomenclature 'shadow education' to refer to educational activities outside the purview of the formal system of education (Stevenson and Baker, 1992; Bray, 1999; Buchmann, 2002). According to Bray and Lykins (2012), the phenomenon of shadow education is widespread across diverse nations like Bangladesh, Nepal, Sri Lanka, Mauritius, Japan, China etc. Tutorial classes under the name of 'juku' have mushroomed in Japan to provide supplementary tutoring and also preparation for entrance exams (Harnisch, 1994). Liu and Bray (2020) use nationally representative data in China to show the recent surge in private tuition demand among primary and lower secondary students.

[25] See MHRD(2016), Bray and Silova (2006)

option if it saves them money. However, the choice of school also depends on a host of school-specific attributes other than teaching quality such as the medium of instruction,[26] discipline, infrastructural facilities, etc. The way the existence of the private tuition market interacts with the teaching quality, the fees, and the other school-specific attributes in the parents' choice of schools for their child, is not clear either in the existing theoretical or the empirical literature, and therefore, stands as an important research question. The chapter addresses this issue. It deals with the school choice of children in the age group 13-18 at the secondary and higher secondary level where the possibility of a mismatch between perception and actual experience about the quality of teaching at schools, which can happen at the elementary level as pointed out by Lahoti and Mukhopadhyay (2019), Blimpo, Evans, and Lahire (2015), is reduced to the minimum through experienced learning and commonly available information about schools. First, it shows with Indian data that controlling for other factors affecting school choice, a more expensive private tuition market lowers the probability that at the secondary and the higher secondary level a child is sent to a government school. Therefore, a less expensive private tuition market crowds out private schooling. Second, it constructs a theoretical model to explain the empirical finding. The theoretical model shows that the empirical finding of the chapter is consistent with the conventional wisdom that in India the teaching quality imparted by government schools falls short of the teaching quality imparted by private schools.

There are several empirical studies analyzing the choice between government and private schools in developing countries. The literature investigates the impact of various school-specific attributes and household characteristics on school preferences. Owing to

[26] Preference for English as medium of instruction in schools stem from the usefulness of English in communication and in securing white collar jobs. In the age of globalization, the English language may have been potentially accepted as the lingua-franca. The literature shows that there is high returns associated with English medium instruction in schools (Azam et al. 2010; Singh, 2013; Munshi and Rosenzweig, 2006). According to Chakraborty and Bakshi (2016), the switch from English to Bengali medium of instruction imposed by the communist government in Indian state of West Bengal in 1993 significantly reduced wages for that age cohort. On the contrary, Muralidharan and Sundaraman (2015) shows that switching to private English medium schools under the lottery based school choice programme in Andhra Pradesh turned out to be detrimental for first generation learners.

quality considerations and efficient management, there is a tendency among households to prefer private schools. For instance, Alderman et al. (2001) show that school choice even among low-income households in Pakistan is highly sensitive to school-specific attributes, leading to a shift in demand towards more efficient and better-quality private schools. The findings of this chapter are consistent with the literature on school quality which reveals a general distrust and dissatisfaction towards government schools as a byproduct of the fallacies in the government education system. Inefficient quality of teaching (Glick and Sahn, 2006; Nishimura and Yamano, 2013) is one of the dominant reasons. In the Indian context too, there exists a vast literature on the inferior quality of government schools indicating a high rate of teacher absenteeism (Kingdon, 1996; Chaudhary et al., 2006; Desai et al., 2008; Kremer et al., 2005, Glewwe and Kremer, 2006)[27]. The literature also shows the relative effectiveness of private schools compared to government schools in terms of the achievement of students at the primary level (French and Kingdon, 2010; Singh, 2015; Muralidharan and Kremer, 2007; Desai et al., 2008)[28]. Besides, the recent development of low-fee private schools in many states has reinforced preferences toward private schools (Kingdon, 2020; Tooley and Dixon, 2007; Tooley, 2009)[29]. However, the literature still points out several flaws in the private schooling system infested by serious equity concerns and exclusion of many children from access to private schooling (Harma, 2011; Woodhead et al., 2013), which is also not fulfilled by their low fee counterparts. In a recent study, Lahoti and Mukhopadhyay (2019) while analyzing the parents' choice of schooling for their children in India show that the household perceptions of

[27] A study by Azim Premji Foundation (2017) shows that absenteeism of government school teachers without any reason was much lower at 2.5 percent compared to absence due to official and administrative duties. In a following paper, Kundu (2019) shows that though the onus of poor learning achievement lies on the government school teachers, the government must focus more on allocation of teachers and teacher training programmes. In reality, there is a huge shortage of government school teachers at elementary and secondary levels.

[28] However, according to Desai et al. (2008), government school students perform better in states like Maharashtra, Tamil Nadu, Delhi, Haryana and North eastern states after controlling for parental characteristics. Singh (2015) shows no significant impact of private school premium on achievement in urban areas.

[29] According to the UDISE Plus (2020-21) estimates, out of a total of approximately 15 lakh schools in India, 23 percent are private unaided. The proportion of private schools have increased by 30 percent over the period 2012-13 to 2020-21 (UDISE, 2012).

teaching-learning constituted 33 percent of the reasons to send children to a particular type of school[30]. The other reasons included discipline (11 percent), security (9 percent), medium of instruction (11 percent), and expenses (8 percent).

Despite being heavily critiqued, there is a huge proportion of the student population enrolled in government schools in India. The National Sample Survey 75th Round 2017-18 shows that in India a high proportion of 13-18-year-old students in secondary and higher secondary education levels attend government schools, especially in states like West Bengal (94 percent), Tripura (99 percent), Bihar (94 percent), Assam (92 percent), Orissa (93 percent). Figure 3.1 brings out the simultaneous high incidence of private tuition in these states. Clearly, the states with a high proportion of government school attendees have a more flourishing demand for private tuition. In the 22 major states considered for the analysis in this chapter, we find that among the students opting for private tuition, only 27.43 percent attend private schools, and the rest 72.57 percent attend government schools. This signals some kind of a deficiency in the government schooling system which leads to a spurt in demand for private tuition in these states.[31] This chapter points out that to some extent private tuition crowds out private schooling.

[30] In fact, Lahoti and Mukhopadhyay (2019, pg. 52) find in their survey at the elementary level of education that "parental perceptions vis-à-vis school realities gathered from the school…..shows a huge mismatch between the two in low-fee private schools. Although parents report that children are going to English medium schools, the reality for most such children is that they are not being taught in English. Similarly, although parents report selecting schools because they care about the quality of teachers, on average, they end up picking schools that have lesser qualified teachers than other schools."

[31] This observation lies in concordance to the estimates for rural households provided by Annual Status of Education Report (ASER, 2019), according to which among students in grades 1-5, 16.6 percent are government school students taking tuition while only 9 percent are private school students taking private tuitions in the year 2018. The corresponding rates for classes 6-8 are 21 percent and 7 percent respectively..

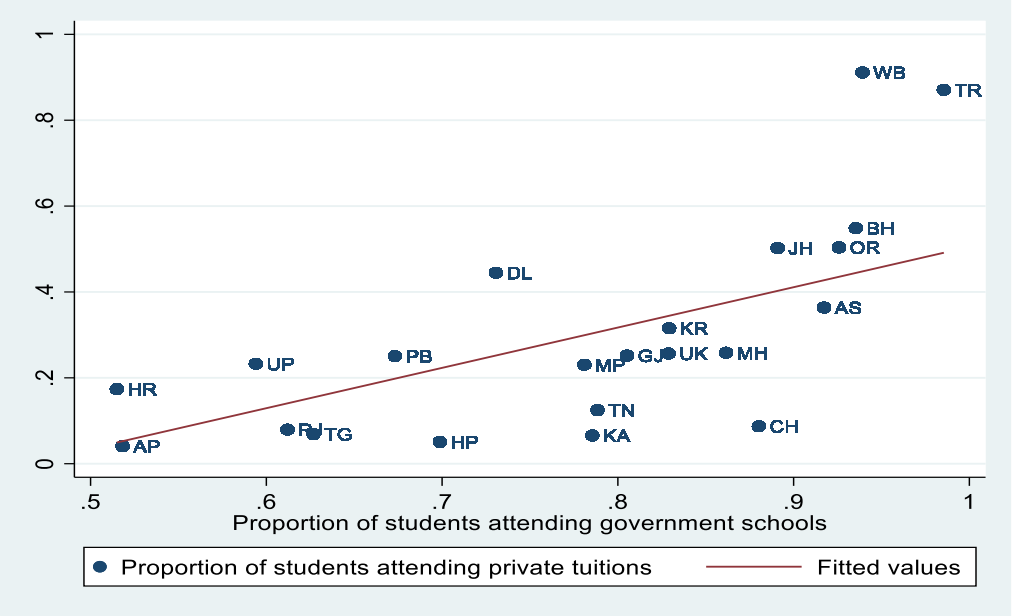


Figure 3.1: Relationship between proportion of students going to government schools and proportion of students attending private tuitions in the 13-18 age cohort statewise.

Note: AP-Andhra Pradesh; TG-Telengana; HR-Haryana; RJ-Rajasthan; UP-Uttar Pradesh; TN-Tamil Nadu; HP-Himachal Pradesh; PB-Punjab; KA-Karnataka; MP-Madhya Pradesh; GJ-Gujarat; DL-Delhi; KR-Kerala; UK-Uttarakhand; MH-Maharashtra; CH-Chhattisgarh; AS-Assam; JH-Jharkhand; OR-Orissa; BH-Bihar; WB-West Bengal; TR-Tripura
Source: NSS 75th Round 2017-18 data on Education

Despite substantial literature existing on the global expansion of private tuition markets (Bray, 2013, 1999; Lee, 2013), research on private tuition is rather limited in India, especially for secondary and higher secondary students. Azam (2015) finds a much higher incidence of opting for private tuition among secondary and senior secondary students compared to primary. Using data on Muzaffarpur district in rural Bihar, Banerji and Wadhwa (2015) find that 76 percent and 81 percent of students in classes 9 and 10 respectively attend private tuition. Among tuition taking students in classes 9 and 10, almost all were enrolled in government secondary schools.[32] Sujatha (2014) uses data from four Indian states – Andhra

[32] The trend of higher incidence of private tuitions at higher grades is not specific to India (see Dang, 2007; Bray, 1999; Bray et al., 2013; Tansel and Bircan, 2006 etc.).

Pradesh, Kerala, Maharashtra, and Uttar Pradesh in her study to show that among government and private aided school students, the most propelling reason for opting for private tuition is the inability to understand school teaching, low teacher quality and the necessity to pass examinations. In a recent study on West Bengal, one of the Indian states, Ghosh and Bray (2020) too cited a lack of quality teaching as a predominant reason behind seeking private tutoring, which was brought up more frequently by government school students. They reported a higher pupil-teacher ratio, teacher absenteeism, and poor infrastructure. The results derived in this chapter are on the same lines, wherein students opt for private tuition to make up for the low quality of teaching in government schools. In a recent study, Berry and Mukherjee (2019) explored the role of the price of private tuition on attendance and dropouts using data from a field experiment in Delhi's slum areas. Lakshmanasamy (2017) and Azam (2015) looks into different factors affecting private tuition expenditure in Indian households. However, no study looks into the effect of the existence of the private tuition market on school choice[33]. This motivates the present chapter.

The study estimates the impact of private tuition expenditure on school choice for children in the age group of 13-18 years. It uses the National Sample Survey (NSS) of India's 75th Round survey on Household Social Consumption: Education, 2017[34]. It assumes that private tuition is institutionalized and the cost of sending the child to private tuition is common

[33] Banerji and Wadhwa (2015) discusses household school choices as well as private tuitions using evidence from a survey based in rural Bihar. However, it does not make an attempt to draw any correlation between the two.

[34] The Annual Status of Education (ASER) Report also provides data on private tuitions and expenditure incurred on it by households. However, this household-based survey only covers children in the age group 3-16 which excludes majority of the students in secondary and higher secondary levels, which is our primary focus group. Also, ASER is restricted only in rural areas. However, according to the NSS 75th Round data, 26 percent of students opt for private tuitions in urban areas. Also, according to the ASSOCHAM (2013) report, 87 percent of primary and 95 percent of high school students receive private tuitions in metro cities in India. This undermines the incidence of private tuitions in ASER data. For all these above-mentioned reasons, we chose to opt for NSS data instead of ASER in our analysis. See 'Private coaching poaches mainstream education' (2013, June 25). *Times of India*. Retrieved from: https://timesofindia.indiatimes.com/city/kolkata/private-coaching-poaches-mainstream-education/articleshow/20763491.cms

knowledge. Thus, while deciding about the type of school for their child, the households are well aware of the possible share of private tuition expenditure in their household per-capita expenditure. The school choice is influenced by the change in the expenditure share. The regression of the school choice variable on private tuition expenditure in the study considers the selection bias that arises due to exclusion of children not currently attending education in the first place. However, the results show an absence of selection bias, which leads us to carry out the subsequent estimation using a simple probit model. The normalization of the private tuition expenditure on a child by the monthly per capita consumption expenditure of the household takes care of the potential endogeneity of private tuition expenditure due to unobserved household characteristics[35]. The regression results show a negative significant effect of the relative private tuition expenditure on government school enrolment. It is also evident that a rise in private school fees relative to government schools raises the probability of selecting government schools. However, the effect does not have a significant impact on the probability of selecting a government school. We also conduct subsequent regressions controlling for district-level variables such as the proportion of government schools in the district and also variables capturing district-level quality differences between private and government schools (Sahoo, 2017) like relative proportions of government schools having English as a medium of instruction vis-à-vis private schools and relative proportions of teachers with professional qualification in government schools vis-à-vis private schools. The effect of private tuition expenditure remains robust to the addition of these controls.

The theoretical model presented in the chapter attempts to explain the empirical results by exploiting a household's utility maximization behaviour by choice of schooling when the government and private schools differ in terms of their fee structure, school-specific

[35] See Dang (2007), Dang and Rogers (2008) and Dongre and Tewary (2015) in this context. Kingdon (1996) reports a possibility of biased estimates of incidence of taking private tuitions in the regression for achievement scores due to reverse causality.

attributes, teaching quality and private tuition market exists that eliminates the gap in teaching quality between the schools. The model assumes that the private school has higher fees compared to the government schools as we have found in our data and allows for both cases where the school-specific attributes are better in the private schools vis-à-vis the government schools and vice-versa. In the model, the households differ from each other in their valuation of the school-specific attributes, which remain as their private information. For some households, it is very important for the choice of schools, and for others, it is not. The model exploits the variation in the households' preference for school-specific attributes to identify the marginal households, who will be indifferent between sending their child either to a government school or to a private school. This helps us to derive the equilibrium enrolment in government schools and private schools as a function of the cost prevailing at the private tuition market and the fee difference between the two types of schools. The comparative static exercise results suggest that independent of the difference in the school-specific attributes, as long as the private schools impart better teaching quality compared to the government schools, a higher cost at the private tuition market leads to a fall in government school enrolment, which exactly matches our empirical finding. The model constructed in the chapter bears some semblance to the demand-supply framework introduced by Dang and Rogers (2008) and the theoretical model in Kim and Lee (2010), where private tuitions are expected to meet the high educational preferences of households, which the supply-constrained public education system is unable to satisfy. However, given the host of school-specific dimensions, the role of private tuitions in bridging the teaching quality gap between private and government schools and facilitating heterogeneous school choice behaviour is not something that has been explored in these theoretical models, which we do in this chapter.

The intuition behind the result that the increase in private tuition expenditure as a share of household expenditure has a negative impact on government school enrolment,

obtained from the theoretical model, is the following. First, let us consider the case where private schools have better attributes than government schools. If the teaching quality is better in the private schools as well, the only reason a household would send its child to a government school is that it has a lower fee structure. The gap in teaching quality is compensated by sending the child to the private tuition market. By taking such a decision the household can be better off if and only if the private tuition does not cost much. If the cost of private tuition starts rising, the marginal household would be worse off by sending the child to a government school. Therefore, government school enrolment would fall. Now let us consider the complementary case where the government schools have better attributes than the private schools in addition to their lower fee structure. In this situation, the only reason a household may send its child to a private school is that its teaching quality is better than the government schools. The households sending their child to government schools would compensate for the gap in teaching quality by sending their child to the private tuition market. By taking such a decision the household can be better off if and only if the private tuition does not cost much. If the cost of private tuition starts rising, the marginal household would be worse off by sending the child to a government school. Therefore, government school enrolment would fall. Had the teaching quality in the private schools been worse than that in the government schools, in the first case, the rise in the cost of private tuition would increase the enrolment in the government schools, and in the second, no household would have sent their child to private schools. However, our empirical results do not support these observations. Therefore, our regression results regarding the effect of the cost of private tuition on school choice also help us to additionally conclude that in India the teaching quality in government schools falls short of the teaching quality in private schools.

The contribution of the chapter in the existing literature is more than one. First, it shows that the existence of private tuition market affects the school choice decision of the

households. In India a rise in the cost of sending child to the private tuition market reduces enrolment in the government schools. Second, the results help us to conclude that in India on average the teaching quality at government schools falls short of the teaching quality at the private schools at the secondary level, which could not be concluded earlier because of the unobservability of teaching quality, although it was known that the government schools are better in terms of the teachers' professional educational qualification and worse in terms of their attendance (Muralidharan, 2021, Muralidharan and Sundaraman, 2015). There is a growing literature on teacher value added (outcome-based measures such as previous test scores of students in measuring teaching effectiveness) on student achievements. However, it is mostly concentrated in primary education levels except for a few studies like Azam and Kingdon (2014), Slater, Davies and Burgess (2012), Aaronson, Barrow and Sander (2007). Besides, there are also papers criticizing the value-added approach, owing to the fluctuations in teachers' effectiveness over different years (Baker et. al., 2010). The chapter contributes to the literature by predicting differences in teaching quality using data on expenditure in private tuition markets. The chapter can also explain the findings of ASER (2021), which reports on India's simultaneous rise in both government school enrolment and private tuition enrolment at all age groups. It seems that in COVID 19 period as the parents became poorer and the private tuition market became cheaper because of widespread unemployment, private tuition crowded out the private school enrolment.

The plan of the chapter is the following. The next section discusses the data. The third section describes the empirical model. The fourth derives the regression results. The fifth section introduces the theoretical model and explains the empirical results. The section following concludes.

3.2 Data and descriptive

The dataset used in this chapter is taken from the NSS 75th Round survey on Household Social Consumption: Education, 2017 conducted by Govt. of India. The survey extensively covers almost the whole of India, comprising 1,13,757 households from urban and rural areas and enumerating 5,13,366 individuals. The data contains qualitative and quantitative aspects of education for household members with information on educational attainment, expenditure, literacy, etc. for individuals in the age group 3-35. In this chapter, we consider children in the age group 13-18, for purpose of study, which consists of 96 percent of the total students currently enrolled in secondary and higher secondary levels of schooling, where private tuition is widespread. The survey reports that 19.8 percent of all students (primary and above) opt for private coaching, out of which, secondary and higher secondary students have the highest participation rates in private coaching, 30.2 percent, and 27.5 percent respectively. The corresponding figures for primary and upper primary levels are lower at 16.4 percent and 21.9 percent respectively. As the chapter focuses on the secondary and higher secondary classes (Class 9 to 12), any student in the aforementioned age group attending other grades is excluded from the analysis. Our analysis considers data from 22 Indian states; 21 major states along with one north-eastern state, Tripura, where approximately 87 percent of the selected age group who are currently attending school, opts for private tuition. Since the primary concern of the chapter is the analysis of choice between government and private schools, we only include those districts of the states in our study that have positive proportions of students attending each type of school. Besides, to comply with our main objective of looking into the effect of private tuition on school choice, we include districts only with a positive proportion of students attending private tuition in the relevant age group[36]. The final sample size used for the current

[36] Out of the 515 districts, 55 districts have children going to either type of school, but none opting for private tuitions. Among these, 22 districts are from the state of Telengana. However, we enquired about the validity of the data and confirmed from the local residents of the State that even in Telangana there is a prevalence of private

study is 31420. Before proceeding with our empirical analysis, it must be noted that the entire discussion of school choice is built upon two types of school management, namely, government and private unaided. Following the argument stated by Kingdon (2020), due to the similarities in the functioning of government and private-aided schools in India, we have clubbed the two under the umbrella of government schools. Private unaided schools are referred to as private schools.

The summary statistics of the variables used in the present analysis is shown in Table 3.1. From the data we find that in the 13-18 age group, 65 percent of children are currently attending school in classes 9-12. Among those children currently attending school, 76 percent are enrolled in government schools. While 55 percent of the sample is male and 74 percent comes from a rural background, 34 percent of the sample belongs to a household in the lowest consumption quartile. 27 percent of the children in the sample come from a household with an educated head. A t-test for comparison of means indicate a positive significant difference between school fees of private and government schools. The average monthly private tuition expenditure of households is Rs 151.77. This monthly expenditure on private tuition for a child is 4 percent of the monthly per capita expenditure (MPCE) of the household on an average.

Since the NSS dataset consists of demand-side variables on education, we have used the Unified District Information System for Education (UDISE) raw data for the year 2017-18 to extract district-level school information which constitutes supply-side variables in education. For instance, we compute the district-level relative proportion of government schools with English medium of instruction and teachers with professional degrees vis-à-vis

tuition. It suggests the existence of data aberration. We consider that the same holds for districts in other states of India. Even after considering such districts, we show that the main results hold (The marginal effects are given in Appendix Table 3.A2.1) However, if there are such districts where private tuition does not exist even when the two types of schools exist, they represent the case of demand-supply mismatch in the private tuition market leading to market failure.

private schools[37]. The data on the proportion of government secondary schools[38] are also extracted from the total schools with secondary and higher secondary classes. The table shows that on average, government schools have a relatively higher proportion of teachers with a professional qualification. There is a lesser number of English medium government schools compared to private schools. The differences in mean proportions between the two types of schools are statistically significant.[39] Also, district wise the proportion of government schools is higher than private schools on average.

Table 3.1: Summary statistics of the variables used in the study

VARIABLES	Source	N	Mean	SD
Child level variables				
Currently attending school	NSS	31420	0.65	0.48
Government school	NSS	23072	0.76	0.43
Male	NSS	31420	0.55	0.5
Monthly expenditure on private tuition (INR)	NSS	31420	151.77	406.04
Expenditure on private tuition as share of MPCE	NSS	31420	0.04	0.11
Household level variables				
MPCE (INR)	NSS	31420	2,006	1,472.09
Location-Rural	NSS	31420	0.74	0.44
Lowest income quartile	NSS	31420	0.34	0.47
Household head educated (Class 10 and above)	NSS	31420	0.27	0.44
District level variables				
Monthly fees difference between pvt and govt. (INR)	NSS	460	985	743
Monthly private school fee (INR)	NSS	460	1222.76	777.5
Difference in proportion of English medium schools between govt and pvt.	DISE	460	-0.28	0.22
Proportion of government English medium schools	NSS	460	0.16	0.23
Proportion of secondary government schools	DISE	460	0.58	0.22
Difference in proportion of teachers with professional qualification between govt and pvt.	DISE	460	0.10	0.14

Note: NSS – National Sample Survey 2017-18; DISE – District Information System on Education 2017-18; MPCE – Monthly Per Capita Expenditure

[37] Teachers' degree may not completely reflect the teaching quality because of shirking. The literature also has disputed evidence on using teacher qualification as a proxy for teaching quality (Hanushek, 1997; Hanushek and Rikvin, 2006; Azam and Kingdon, 2014; Kremer et al., 2005)).

[38] This includes schools under Department of Education, Tribal Welfare Department, Local Body, Social Welfare Department, Ministry of Labour, Kendriya Vidyalaya, Jawahar Navodaya Vidyalaya, Sainik School, Railway School, Central Tibetan School, Recognized Madrasa and Government aided schools.

[39] The literature on school choice has often considered measures like pupil teacher ratio, class size (Nishimura and Yamano, 2013; Glick and Sahn, 2006) on school choice. Though DISE provides data for these measures at different levels of schooling, there is no secondary and higher secondary level aggregate figure for the same. We have included the district level average difference in private and government school fees in the school choice regression similar to Sahoo (2017). In Sahoo(2017), its effect on gender gap in school choice is gauged.

3.3 Empirical model and results

In this section a multivariate regression model is constructed for analysing school choice at the household level. The specification is characterized by Equation (3.1), where the dependent variable is a binary variable indicating whether the child is enrolled in a government school ($Y_{chds} = 1$) or private school ($Y_{chds} = 0$).

$$Y_{chds} = \alpha + X_{chd}\beta + \varphi_s + \epsilon_{chds} \quad (3.1)$$

The subscript c refers to a child from household h, in district d in state s. The explanatory variable X depicts the vector of explanatory variables including the household expenditure on private tuition of the child as a share of MPCE. It also includes child-specific characteristics such as the gender of the child, household-specific characteristics such as the location of the household, household head's education and monthly per capita consumption expenditure of all household members, and all the district-level variables specified above that might explain the school choice. β is the coefficient vector associated with this set of explanatory variables. The model also includes state-fixed effects, φ_s, to capture the state-level unobserved factors affecting the choice of schooling.

However, Equation (3.1) does not account for children currently not attending school.[40] This can give rise to a potential sample selection bias since there might exist some correlation between the unobserved factors affecting school choice and current attendance. Thus, we re-estimate the school choice equation conditional on children currently attending education following the suggestion of Heckman (1979). The methodology applied to account for this typical problem is the probit model with sample selection. The regression model is estimated

[40] School dropout is a major concern in India, especially at secondary levels. According to UDISE Plus (2017-18, 2018-19), 18.93 and 17.9 percent students dropped out at secondary level in 2017-18 and 2018-19 respectively. Though the percentage of secondary dropouts has reduced to 12.6 percent in 2021-22 (UDISE Plus, 2021-22), there still remains a reason for concern.

by maximum likelihood (Van De Ven & Praag, 1981). Two equations are required for this purpose:

(i) Selection equation: A probit regression with current attendance as the binary dependent variable; "1" if the child is currently attending school and "0" if not.

(ii) Outcome equation: Probit regression with school choice as the binary dependent variable taking a value "1" if government school is opted for and "0" if private school is opted.

Formally, to estimate the probit model estimating the probability of attending government school assumes the relationship.

$$Y^*_{chds} = \alpha + X_{chd}\beta + \varphi_s + \epsilon_{chds}$$

$$Y_{chds} = 1[Y^*_{chds} > 0] \tag{3.2}$$

We consider Y^*_{chds} as the latent variable capturing the household decision to enroll a child in a government school. It represents the household propensity to enroll the child in a government school. What we observe is a binary outcome Y indicating whether the child goes to a government school ($Y = 1$) or private school, ($Y = 0$) only when $Y^*_{chds} > 0$. The main variable of interest is the expenditure on private tuition as the share of MPCE which is included in the vector of explanatory variables X.

The information on school choice is contingent upon the fact that children are currently attending school.

$$A^*_{chds} = \mu + \boldsymbol{Z}_{chd}\boldsymbol{\gamma} + \varphi_s + u_{chds}$$

$$A_{chds} = 1[A^*_{chds} > 0] \tag{3.3}$$

We can observe the school choice decision of households only when $A_{chds} = 1$. $\boldsymbol{Z}_{chd}$ is the vector of covariates affecting the decision of attending school. It includes the gender dummy,

the location dummy, the monthly per capita consumption expenditure of the household members, the dummy indicating that the household head has completed tenth grade, share of private tuition expenditure in MPCE, district-level fee differences between private and government schools, and the dummy indicating that the household belongs to the lowest consumption quartile after controlling for the household consumption expenditure (MPCE). The last variable is the exclusion restriction which is required for the identification of these kinds of models, where at least one variable is present in the selection equation, which is excluded while estimating the outcome equation (Cameron and Trivedi, 2009). After controlling for the MPCE in both equations, we include an identifying variable in the selection equation which indicates whether a household lies in the bottom quartile. Here we follow the papers like Sahoo (2015) and Kumar and Sahoo (2021), which argue that this variable identifies the poorest individuals for whom enrolment decision is important. However, the household is so poor that the choice of schooling becomes less important.[41] The errors of the selection and outcome equations follow bivariate normal density. $\epsilon \sim N(0,1)$ and $u \sim N(0,1)$ and $corr(u, \epsilon) = \rho$.

When $\rho \neq 0$, the error terms of the selection and outcome equations are correlated and standard probit estimates would result in biased estimates as in the case of omitted variables.

[41] In the sample considered in the study, we find that among children coming from the lowest consumption quartile, 41 percent does not attend school, 52 percent attend government schools, whereas only 7 percent goes to private schools. Thus, it is evident that even if children from the poorest background attends school, government school is the default choice for the majority as it is the cheaper option.

Table 3.2: Effect of share of private tuition expenditure in MPCE on the choice of govt school using heckprobit

VARIABLES	(1) Govt	(2) Currently attending
Share of private tuition exp to MPCE	-0.390**	2,033***
	(0.194)	(57.23)
Log of fee difference pvt vs govt	0.0498	-0.145**
	(0.0971)	(0.0668)
Male	-0.0735**	0.0388
	(0.0348)	(0.0287)
Rural	0.404***	0.101**
	(0.0495)	(0.0456)
MPCE	-0.000187***	5.18e-05
	(1.92e-05)	(3.26e-05)
Lowest consumption quartile		-0.137***
		(0.0510)
Head's education at least class 10	-0.245***	0.775***
	(0.0515)	(0.0394)
Constant	0.468	1.855***
	(0.735)	(0.512)
Observations	31,420	31,420
Prob>Chi2	0.8164	
StateFE	YES	YES

Robust standard errors clustered at district level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1

The heckprobit results in Table 3.2, show that the Wald χ^2 statistic denoting selection is insignificant. This indicates that the outcome and selection equations are unrelated and can be independently estimated. In the next exercise, we perform ordinary probit regression of school choice on the explanatory variables including private tuition expenditure using Equation (3.1). The results are given in Table 3.3 below.

Table 3.3: Effect of share of private tuition expenditure in MPCE on the choice of govt school using probit

VARIABLES	(1) Govt	(2) Govt	(3) Govt	(4) Govt	(5) Govt
Pvt tuition exp as share of MPCE	-0.422***	-0.416***	-0.423***	-0.430***	-0.420***
	(0.160)	(0.161)	(0.162)	(0.152)	(0.154)
Log of fee difference pvt vs govt	0.0514	0.0560	0.0578	0.0580	0.0764
	(0.0970)	(0.0985)	(0.0980)	(0.0963)	(0.0987)
Male	-0.0744**	-0.0749**	-0.0749**	-0.0689**	-0.0702**
	(0.0348)	(0.0347)	(0.0349)	(0.0346)	(0.0345)
Rural	0.404***	0.403***	0.403***	0.385***	0.381***
	(0.0495)	(0.0495)	(0.0494)	(0.0517)	(0.0511)
MPCE	-0.000188***	-0.000187***	-0.000188***	-0.000181***	-0.000180***
	(1.86e-05)	(1.87e-05)	(1.85e-05)	(1.86e-05)	(1.85e-05)
Head's education above class 10	-0.252***	-0.252***	-0.253***	-0.243***	-0.245***
	(0.0452)	(0.0452)	(0.0452)	(0.0453)	(0.0453)
English Medium difference		0.0800			0.159
		(0.232)			(0.238)
Professional training difference			-0.410		-0.536
			(0.332)		(0.335)
Prop of govt schools in district				0.922***	0.983***
				(0.237)	(0.222)
Constant	0.467	0.477	0.431	-0.205	-0.280
	(0.735)	(0.735)	(0.741)	(0.747)	(0.750)
Observations	23,072	23,072	23,072	23,072	23,072
State FE	YES	YES	YES	YES	YES

Robust standard errors clustered at the district level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1

The first column in Table 3.3 displays the results with no additional district-level controls for school quality. The coefficient for private tuition expenditure is negatively significant at 1 percent. This implies that the probability of getting enrolled in a government school decreases as the share of private tuition spending in MPCE rises. The marginal effects are shown in Table 3.A1.1 in the Appendix. We find that at the median MPCE, if the share of expenditure on private tuition increased by one unit, the probability of attending government schools decrease by 0.1. The sign and figure remain more or less unchanged upon the introduction of additional controls. It is found that fees difference between private and government schools has a positive impact on the probability of attending a government school. However, this effect is insignificant. Male children are less likely to attend government schools while children in rural areas are more likely to go to government schools. We also find a negative correlation between the monthly per capita consumption of the household and the probability of going to a government school. Also, households with an educated head (one who has at least completed class 10) are more likely to send their child to private schools. Next, we include the district-level control variables one by one to check their effects separately and find any change in the impact of the share of private tuition expenditure. The second column in Table 3.3 displays the results with the difference in the district-level proportions of government vs private schools with English medium instruction. Though this difference has an insignificant impact, the effect of the expenditure share of private tuition on school choice still remains highly significant at the 1 percent level. A similar result (Table 3.3 Column 3) holds when we include the difference in the proportion of professionally educated teachers between government and private schools[42]. The fourth column in Table 3.3 controls for the proportion of government schools in

[42] It is seen from the regression result in Table 3.3 Columns 3 and 5, that a rise in the proportion of professionally trained teachers in government schools reduce the probability of selecting government schools, though the effect is insignificant. Though it might appear counterintuitive, the result lies in conjunction with Kremer et al. (2005), which shows that more educated teachers and those with professional training are more likely to be absent. The underlying logic arises from the power differential that a better educate and trained teacher possesses, rendering them less vulnerable to sanctions on account of shirking.

a district. This is significant at the 1 percent level. This shows that given all other factors the choice of sending a child to a government school is significantly influenced by the relative availability of such schools in the district vis-à-vis the private schools. However, the impact of the share of private tuition expenditure in MPCE still remains robust. The final column (Table 3.3 Column 5) supports the robustness of the result related to the effect of the share of private tuition expenditure in MPCE with all the district-level controls in the regression[43].

3.4 The theory

Consider a single-child household that decides whether to send its child to a government school (g) or private school (p). The measure of such households is assumed to be 1. The utility function derived by the household from its child attending the i^{th} type of school is

$$U = t_i + \delta\mu_i + x, \qquad \forall i = g, p$$

where t_i represents the teaching quality in the i^{th} type of school, μ_i represents school-specific attributes such as medium of instruction, discipline, infrastructure, etc. $\delta \in [0,1]$ represents the value assigned by a household to school-specific attributes, which is assumed to be uniformly distributed over the households. If $\delta = 0$, a household does not attach any value to school-specific characteristics. The valuation increases with a rise in δ. The utility of a household also depends on the expenditure on other goods, x.

[43] All the results hold if we control for the proportion of English medium government schools and monthly private school expenditure instead of the differences in English medium between the government and private schools and the log of school fee differences respectively in the regressions. Table 3.A2.2 shows a heckprobit regression where we find the absence of any selection bias. Following that, the final probit results are shown in Table 3.A2.3. The corresponding marginal effects are given in Table 3.A2.4.

The school fee at the ith type school is given by c_i, $\forall i = g, p$. From Table 3.1, we know that in India, on average private school fees are higher than government school fees. Therefore, we assume, $c_p > c_g$.

If $t_i > t_j$, $i \neq j$, the jth school child is sent to the private tuition market for bridging the gap in teaching quality[44]. The cost of sending a child to the private tuition market is the price prevailing at the private tuition market, denoted by $\theta > 0$ and the outcome is $t_i = t_j$. A household takes θ as given.

A household that does not send its child to private tuition maximizes U subject to the budget constraint $x + c_i = m$ by choice of x and by the choice of school.

A household that sends its child to private tuition maximizes U subject to the budget constraint $x + c_i + \theta = m$ by choice of x and by the choice of school.

Case 1: $\mu_p > \mu_g$.

Case 1 corresponds to the situation that private schools are better endowed with school-specific attributes.[45]

Case 1A: $t_p > t_g$.

Since private schools impart better teaching quality than government schools, the households that send their child to government schools also send the child to private tuition for bridging the gap in teaching quality. Consequently, $t_g = t_p$ is achieved.

The optimum choice of x generates the following values of indirect utilities corresponding to the respective choice of private and government school:

[44] The theoretical model assumes the existence of the private tuition market.
[45] Table 3.1 shows that in India there are more private schools with English as medium of instruction relative to the government schools.

$$v_p = t_p + \delta\mu_p + m - c_p; \quad (3.4)$$

$$v_g = t_p + \delta\mu_g + m - (c_g + \theta). \quad (3.5)$$

The decision on school choice rests upon the comparison of v_p and v_g. From equation (3.4) and (3.5):

$$v_p - v_g = \delta(\mu_p - \mu_g) - \varphi, \quad (3.6)$$

where $\varphi = (c_p - c_g) - \theta$.

From (3.6), $v_p \geq v_g$ iff $\delta(\mu_p - \mu_g) \geq \varphi$, a private school is chosen; $v_p < v_g$ iff $\delta(\mu_p - \mu_g) < \varphi$, a government school is chosen.

Assumption 3.1: $\varphi \in (0, \mu_p - \mu_g)$.

Since $\mu_p > \mu_g$, $\delta(\mu_p - \mu_g)$ is monotonically increasing in δ with $\delta(\mu_p - \mu_g) \to 0$ as $\delta \to 0$. Assumption 3.1 ensures, there exists a $\delta = \bar{\delta} \in (0, 1)$ such that $\delta(\mu_p - \mu_g) = \varphi$ holds. We assume the marginal households with $\delta = \bar{\delta}$ sends their child to government schools. Therefore, all the households having $\delta \in [0, \bar{\delta}]$ send their child to the government schools and all the households having $\delta \in (\bar{\delta}, 1]$ send their child to private schools. If Assumption 3.1 is violated, all the children attend either the private schools or the government schools even if they have choice, which is clearly not the case in our data.

The following diagram depicts the equilibrium school choice.

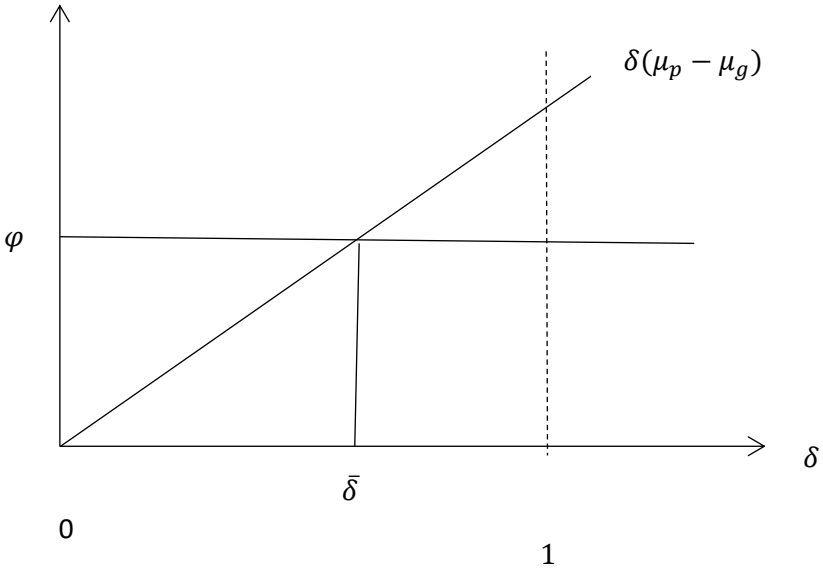


Figure 3.2: School choice decision with $\mu_p > \mu_g$ and $t_p > t_g$.

Since δ is distributed uniformly between [0,1], the number of students enrolled in government schools is given by:

$$\bar{\delta} = \frac{\varphi}{\mu_p - \mu_g}. \tag{3.7}$$

Since $\mu_p > \mu_g$ and $t_p > t_g$, in this case, the only reason a household opts for a government school is the money it saves by paying lower fees at the government school vis-à-vis the private schools, which can be used to have more x. If $(c_p - c_g)$ rises, φ rises. Therefore, from equation (3.7), $\bar{\delta}$ rises. On the other hand, if θ rises, φ falls, and therefore from equation (3.7), $\bar{\delta}$ falls. As the price of private tuition rises the marginal household sending its child to government school no longer finds it profitable to do so. Therefore, it switches to a private school where the child is no longer required to be sent to private tuition because of its higher teaching quality.

Case 1B: $t_p < t_g$.

Since government schools impart better teaching quality than private schools, the households that send their child to private schools also send the child to private tuition for bridging the gap in teaching quality. Consequently, $t_p = t_g$ is achieved.

The optimum choice of x generates the following values of indirect utilities corresponding to the respective choice of private and government school:

$$v_p' = t_g + \delta\mu_p + m - (c_p + \theta); \tag{3.8}$$

$$v_g' = t_g + \delta\mu_g + m - c_g. \tag{3.9}$$

The decision on school choice rests upon the comparison of v_p' and v_g'. From equation (3.8) and (3.9):

$$v_p' - v_g' = \delta(\mu_p - \mu_g) - \varphi', \tag{3.10}$$

where $\varphi' = (c_p - c_g) + \theta$.

From (3.10), ${v_p}' \geq {v_g}'$ iff $\delta(\mu_p - \mu_g) \geq \varphi'$, a private school is chosen; ${v_p}' < {v_g}'$ iff $\delta(\mu_p - \mu_g) < \varphi'$, a government school is chosen.

Assumption $3.1'$: $\varphi' \in (0, \mu_p - \mu_g)$.

When assumption $3.1'$ holds, by use of a similar argument developed in the previous case, one can conclude that in this case the number of students enrolled in government schools is given by:

$$\bar{\delta}' = \frac{\varphi'}{\mu_p - \mu_g}. \tag{3.11}$$

Since $\mu_p > \mu_g$ and $t_p < t_g$, in this case, the only reason a household opts for a private school is the superior school-specific attributes available at the schools. The deficiency in teaching quality at the private school is compensated by sending the child to private tuition. If θ rises,

φ' rises and therefore from equation (3.11), $\bar{\delta}'$ rises. As the price of private tuition rises the marginal household sending its child to private school no longer finds it profitable to do so. Therefore, it switches to a government school where the child is no longer required to be sent to private tuition because of the higher teaching quality imparted by the government school. If $(c_p - c_g)$ rises, the same argument holds leading to a rise in government school enrolment.

Case 2: $\mu_p < \mu_g$.

Case 2 corresponds to the situation that government schools are better endowed with school-specific attributes.

Case 2A: $t_p > t_g$.

Here the government school children go to private tuition. This case is similar to case IA except $\mu_p > \mu_g$ is replaced by $\mu_p < \mu_g$. The decision on school choice rests upon the comparison of v_p and v_g. From equation (3.6):

$$v_g - v_p = \delta(\mu_g - \mu_p) + \varphi, \qquad (3.12)$$

where $\varphi = (c_p - c_g) - \theta$.

From (3.12), $v_g \geq v_p$ iff $\delta(\mu_g - \mu_p) + (c_p - c_g) \geq \theta$, a private school is chosen; $v_g < v_p$ iff $\delta(\mu_g - \mu_p) + (c_p - c_g) < \theta$, a government school is chosen.

Assumption 3.1″: $\theta \in \left(c_p - c_g, ((\mu_p - \mu_g) + (c_p - c_g))\right)$.

Since $\mu_p < \mu_g$, $(\delta(\mu_g - \mu_p) + (c_p - c_g))$ is monotonically increasing in δ. It goes to $(c_p - c_g)$ as $\delta \to 0$ and goes to $((\mu_g - \mu_p) + (c_p - c_g))$ as $\delta \to 1$. Assumption 3.1″ ensures that there exists a $\delta = \bar{\delta}'' \in (0,1)$ such that $\delta(\mu_g - \mu_p) + (c_p - c_g) = \theta$ holds. All the households having $\delta \in [0, \bar{\delta}'')$ send their child to private schools and all the households having $\delta \in [\bar{\delta}'', 1]$ send their child to government schools.

Therefore, in this case, the number of students enrolled in government schools is given by:

$$1 - \bar{\delta}'' = 1 + \frac{\varphi}{\mu_g - \mu_p}. \quad (3.13)$$

If θ rises, φ falls. Therefore, from equation (3.13) the enrolment in government schools falls. If $(c_p - c_g)$ rises, φ rises. Therefore, from equation (3.13) the enrolment in government schools rises.

Since $\mu_p < \mu_g$ and $t_p > t_g$, the reason a child is sent to a private school is its better teaching quality. The child is sent alternatively to a government school if the cost of sending the child to private tuition is not too high. If the price rises further in the private tuition market, the marginal household saves money by sending its child to a private school instead of sending her to a government school. Therefore, government school enrolment falls.

Case 2B: $t_p < t_g$.

In this situation, on no ground does a household choose a private school over a government school for their child. Government schools are at an advantage over private schools with respect to all three factors determining the choice of schooling – teaching quality, school-specific attributes, and affordability. Since the question of school choice is irrelevant under this scenario, we leave it aside in our analysis.

We sum up the theoretical results in the following table:

Table 3.4: Summary of results from Theoretical Model on school choice of households

Relative school-specific attributes	**Relative teaching quality**	**A rise in the cost of private tuition (θ)**	**A rise in fee difference ($c_p - c_g$)**
$\mu_p > \mu_g$	$t_p > t_g$	Government school enrolment falls	Government school enrolment rises
	$t_p < t_g$	Government school enrolment rises	Government school enrolment rises
$\mu_p < \mu_g$	$t_p > t_g$	Government school enrolment falls	Government school enrolment rises
	$t_p < t_g$	Everyone goes to govt. schools. Enrolment does not respond.	Everyone goes to govt. schools. Enrolment does not respond.

The empirical results presented in section 3.3, clearly conform to the case $t_p > t_g$, shown in Table 3.4. independent of the relative status of school-specific attributes. Therefore, the negative significance of the coefficient of share of private tuition in the monthly per capita consumption expenditure in the school choice regressions also reflects the relation between the type of school and its teaching quality. We may conclude that government schools in India fall short in teaching quality compared to private schools.

3.5 Conclusions

The chapter tries to establish a connection between the parallel education sector and the formal school choice of households. While, the superiority of private schools in terms of the availability of school-specific factors places a compelling reason for selection into private schools[46], other factors such as fee structure and affordability move the odds in favor of their

[46] According to the NSS 75th round data, among school goers (primary and above), 17.2 percent households select private unaided schools due to availability of special facilities and 17.7 percent due to availability of English

government counterparts. The chapter claims that the remaining concern of households regarding the quality of schooling is easily tractable, given the existence of the private tuition market. The private tuition market makes up for the gap in teaching quality that exists between the government and private schools, but the higher price of participating in the private tuition market induces a household to choose a school for its child that imparts a better quality of education. The regression analysis with Indian data, carried out in this chapter, shows that the higher share of private tuition expenditure on a child in the monthly per capita consumption expenditure of a household has a negative significant impact on the probability of sending the child to a government school. The result remains robust to the addition of district-level controls such as the proportion of English medium government schools vis-à-vis private schools, the relative proportion of government schools with professional degree-qualified teachers, and the proportion of secondary government schools in a district. The district-level relative school fee in private schools vis-à-vis government schools also has a positive but insignificant impact on the probability of selecting government schools.

It might seem that one major limitation of the chapter is the presumption that the sole objective of private tuition is to fill in the teaching quality differences between government and private schools since it ignores the use of private tuition by households for topping up the quality imparted at schools (Baker et al. (2001)). This may also include private tuition expenses incurred for preparing the child for competitive examinations. However, if that was the dominant reason stimulating demand for private tuition, according to the theoretical model in the chapter, any additional private tuition cost would be borne by both types of households sending their children to either government or private schools. This would wipe out the differential impact of private tuition expenditure on school choice. The presence of a significant

medium learning. Kingdon (2006) points out that a 'differentiated product' delivered by English medium private schools have become more lucrative option for households.

impact of the cost of private tuition on school choice justifies the existence of an underlying difference in teaching quality in government and private schools. The model assumes that private tuition does not have an impact on the quality of teaching at schools. There are papers in the literature like Jayachandran (2014), Bhorkar and Bray (2018), which point out that this is a possibility in presence of a private tuition market. However, if private tuition erodes the quality of teaching at the type of school, where the quality of teaching is already poor, the argument forwarded in this chapter gets reinforced. The model also does not differentiate between the existence of high and low-quality tuition in the private tuition market. If 'high' quality tuition is defined as the quality, which is adequate in bridging the gap in teaching quality in different types of schools, it must be costlier than low-quality tuition[47]. That way it reinforces the argument given in the chapter that the parents in such a situation would enroll their child in a high-teaching quality school if they afford it rather than enrolling their child in a low-teaching quality school and sending him/her to private tuition. Otherwise, private tuition would still be opted for, which would reduce the quality gap without eliminating it.

The results derived in the chapter can have important policy implications. It can predict the likely outcome of the imposition of regulation on the private tuition market both on school choice and the quality of education in a country. The recently implemented regulation on the private educational coaching industry in China is a case in point (Sheng, 2021). Apart from banning private tuition companies providing school teaching from making profits, the local governments are aiming at setting the after-school tutoring fee[48]. A complete ban on the private tuition sector, according to this chapter, is expected to crowd in private schooling in China as

[47] Low-quality private tuition market may lead to negative value addition for the children who are forced to go to low-quality schools. It may happen only in absence of awareness of the parents and may lead to the exacerbation of social inequality as referred to by Azam (2015) and Aslam and Atherton (2012).

[48] See 'China to set prices for after school tutoring' (2021, September 06). *Nikkei Asia*. Retrieved from: https://aus.libguides.com/apa/apa-newspaper-web.

households will switch to private schooling from government schooling. In the case of regulation of private tuition fees, the reverse is expected i.e. enrolment in government schools is expected to rise with cheaper access to private tuition. In both cases, the education quality remains unchanged. However, the empirical validation of these hypotheses with Chinese data is required, which remains in our future research agenda.

There have been some state-level initiatives in banning private tuition in India. For instance, in West Bengal, the state government issued a notice to some government schools and aided-school teachers in 2022 for providing private tuition, which goes against Section 28 of the Right of Children to Free and Compulsory Education Act 2009[49]. However, enforcement remains an issue, given the enormous demand for private tuition. The National Education Policy (NEP, 2020) envisages an ambitious plan of improving teaching quality in schools through continuous professional development programs, providing incentives, recognizing outstanding teachers, etc. This is also supposed to be effective in meeting the quality gap between government and private schools, and simultaneously attenuate the need for private tuition. However, the recent data does not imply so. According to the recent National Achievement Survey (NAS, 2021), a large-scale assessment conducted to draw information on the achievement of students, we find that in the majority of states, private schools perform better than their government counterparts in terms of achievement of Class 10 students in mathematics, social science, science, and English. Despite the poor condition of government schools, the ASER (2021) data shows a clear shift in enrolment towards government schools in 2021 compared to 2018 associated with a rise in the incidence of private tuition to almost 40 percent. From these observations from recent data, it becomes obvious that the demand for

[49] See 'Private Tuition: West Bengal govt asks 40 schools to probe 200 teachers' (2022, August 27). *Times of India*. Retrieved from: https://timesofindia.indiatimes.com/city/kolkata/private-tuition-west-bengal-govt-asks-40-schools-to-probe-200-teachers/articleshow/93809000.cms.

private tuition far from withering away still constitutes an important component of one's education.

Apart from studying the Chinese case, as proposed above, this chapter can also be extended in several other ways, too. One can study gender bias in private tuition expenditure and its effect on the selection of streams for higher education. For this, the theoretical model presented in this chapter may be tweaked by redefining the utility function as the utility of entering a particular stream of higher education. Also, to understand the gender-based differences, we may incorporate gender-specific preferences for school characteristics in the utility function to derive the results. Another interesting study would be finding the role of private tuition expenditure in reducing school dropouts at secondary and higher secondary levels. Here, we can include a third possibility of dropping out of school in the theoretical model and compare it with the other options. The rest of the analysis may follow as usual. These also remain as future work.

Appendix 3.A1

Table 3.A1.1: Marginal effects of share of private tuition expenditure in MPCE on the probability of choosing a government school

VARIABLES	(1) marginal effects	(2) marginal effects	(3) marginal effects	(4) marginal effects	(5) marginal effects
Private tuition exp as share of MPCE	-0.106***	-0.104***	-0.106***	-0.107***	-0.104***
	(0.0397)	(0.0399)	(0.0400)	(0.0375)	(0.0379)
Log of fee difference pvt vs govt	0.0128	0.0140	0.0144	0.0144	0.0190
	(0.0242)	(0.0246)	(0.0245)	(0.0240)	(0.0246)
Male	-0.0186**	-0.0187**	-0.0187**	-0.0171**	-0.0174**
	(0.00873)	(0.00870)	(0.00874)	(0.00862)	(0.00860)
Rural	0.101***	0.101***	0.101***	0.0958***	0.0948***
	(0.0121)	(0.0121)	(0.0121)	(0.0127)	(0.0125)
MPCE	-4.69e-05***	-4.68e-05***	-4.69e-05***	-4.51e-05***	-4.48e-05***
	(4.60e-06)	(4.62e-06)	(4.55e-06)	(4.62e-06)	(4.57e-06)
Head's education above class 10	-0.0629***	-0.0630***	-0.0632***	-0.0605***	-0.0610***
	(0.0112)	(0.0112)	(0.0112)	(0.0112)	(0.0112)
English Medium difference		0.0200			0.0396
		(0.0579)			(0.0593)
Professional training difference			-0.102		-0.133
			(0.0828)		(0.0833)
Prop. of govt schools in district				0.229***	0.244***
				(0.0583)	(0.0549)
Observations	23,072	23,072	23,072	23,072	23,072

Note: Marginal effects at the median value of MPCE are calculated using the margins command in Stata. Marginal effects are shown in Columns (1)-(5) correspond to Columns (1)-(5) respectively in Table 3.3.

Appendix 3.A2

Table 3.A2.1: Marginal Effect of share of private tuition expenditure in MPCE on the probability of choosing a government school using probit including districts with no private tuition

VARIABLES	(1) marginal effects	(2) marginal effects	(3) marginal effects	(4) marginal effects	(5) marginal effects
Pvt tuition exp as share of MPCE	-0.102**	-0.102**	-0.102**	-0.105***	-0.103***
	(0.0410)	(0.0411)	(0.0413)	(0.0386)	(0.0389)
Log of fee difference pvt vs govt	0.0173	0.0176	0.0186	0.0180	0.0218
	(0.0242)	(0.0245)	(0.0244)	(0.0239)	(0.0244)
Male	-0.0272***	-0.0272***	-0.0273***	-0.0262***	-0.0264***
	(0.00862)	(0.00859)	(0.00862)	(0.00854)	(0.00853)
Rural	0.105***	0.105***	0.105***	0.0978***	0.0969***
	(0.0118)	(0.0118)	(0.0118)	(0.0127)	(0.0125)
MPCE	-5.04e-05***	-5.03e-05***	-5.04e-05***	-4.78e-05***	-4.76e-05***
	(4.94e-06)	(4.97e-06)	(4.92e-06)	(4.86e-06)	(4.82e-06)
Head's education above class 10	-0.0677***	-0.0677***	-0.0680***	-0.0646***	-0.0651***
	(0.0109)	(0.0109)	(0.0109)	(0.0108)	(0.0108)
English Medium difference		0.00393			0.0312
		(0.0572)			(0.0584)
Professional training difference			-0.0841		-0.126
			(0.0835)		(0.0837)
Prop of govt schools in district				0.240***	0.254***
				(0.0537)	(0.0512)
Observations	24,858	24,858	24,858	24,858	24,858

Note: Marginal effects at the median value of MPCE are calculated using the margins command in Stata

Table 3.A2.2: Effect of share of private tuition expenditure in MPCE on the choice of govt school using heckprobit controlling for monthly average private school fee

VARIABLES	(1) Govt	(2) Curently attending
Share of private tuition exp to MPCE	-0.399**	1,976***
	(0.194)	(58.76)
Monthly private school fee	3.89e-05	-0.000125***
	(6.08e-05)	(3.87e-05)
Male	-0.0731**	0.0375
	(0.0348)	(0.0288)
Rural	0.406***	0.0889**
	(0.0494)	(0.0448)
MPCE	-0.000188***	5.36e-05
	(1.96e-05)	(3.33e-05)
Lowest consumption quartile		-0.140***
		(0.0514)
Head's education at least class 10	-0.247***	0.777***
	(0.0514)	(0.0396)
Constant	0.788***	0.956***
	(0.146)	(0.140)
Prob>Chi2	0.8386	
Observations	31,420	31,420
State FE	YES	YES

Robust standard errors clustered at district level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 3.A2.3: Effect of share of private tuition expenditure in MPCE on the choice of govt school using probit controlling for the monthly private school fee and proportion of government English medium schools

VARIABLES	(1) Govt	(2) Govt	(3) Govt	(4) Govt	(5) Govt
Pvt tuition exp as share of MPCE	-0.427***	-0.424***	-0.428***	-0.438***	-0.438***
	(0.161)	(0.163)	(0.162)	(0.153)	(0.155)
Monthly private school fee	4.00e-05	5.29e-05	4.26e-05	6.13e-05	7.13e-05
	(6.08e-05)	(6.09e-05)	(6.11e-05)	(6.11e-05)	(6.09e-05)
Male	-0.0738**	-0.0732**	-0.0743**	-0.0679**	-0.0681**
	(0.0348)	(0.0347)	(0.0349)	(0.0346)	(0.0346)
Rural	0.406***	0.402***	0.406***	0.389***	0.387***
	(0.0494)	(0.0490)	(0.0494)	(0.0514)	(0.0507)
MPCE	-0.000189***	-0.000187***	-0.000189***	-0.000183***	-0.000183***
	(1.90e-05)	(1.91e-05)	(1.88e-05)	(1.89e-05)	(1.87e-05)
Head's education above class 10	-0.252***	-0.252***	-0.254***	-0.245***	-0.247***
	(0.0453)	(0.0452)	(0.0453)	(0.0453)	(0.0454)
Prop of English medium govt school		-0.458*			-0.222
		(0.237)			(0.267)
Professional training difference			-0.408		-0.557*
			(0.332)		(0.331)
Prop of govt schools in district				0.948***	0.955***
				(0.248)	(0.251)
Constant	0.796***	0.949***	0.804***	0.126	0.206
	(0.139)	(0.163)	(0.138)	(0.219)	(0.265)
Observations	23,072	23,072	23,072	23,072	23,072
State FE	YES	YES	YES	YES	YES

Robust standard errors clustered at the district level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 3.A2.4: Marginal Effect of share of private tuition expenditure in MPCE on the probability of choosing a government school using probit controlling for the monthly private school fee and proportion of government English medium schools

VARIABLES	(1) marginal effects	(2) marginal effects	(3) marginal effects	(4) marginal effects	(5) marginal effects
Pvt tuition exp as share of MPCE	-0.107***	-0.106***	-0.107***	-0.109***	-0.109***
	(0.0398)	(0.0404)	(0.0401)	(0.0376)	(0.0381)
Monthly private school fee	1.00e-05	1.32e-05	1.06e-05	1.52e-05	1.77e-05
	(1.52e-05)	(1.52e-05)	(1.52e-05)	(1.51e-05)	(1.51e-05)
Male	-0.0185**	-0.0183**	-0.0186**	-0.0169*	-0.0169**
	(0.00874)	(0.00868)	(0.00874)	(0.00862)	(0.00860)
Rural	0.102***	0.100***	0.101***	0.0968***	0.0960***
	(0.0120)	(0.0119)	(0.0120)	(0.0127)	(0.0124)
MPCE	-4.72e-05***	-4.66e-05***	-4.72e-05***	-4.56e-05***	-4.54e-05***
	(4.66e-06)	(4.70e-06)	(4.62e-06)	(4.67e-06)	(4.62e-06)
Head's education above class 10	-0.0631***	-0.0629***	-0.0634***	-0.0609***	-0.0613***
	(0.0112)	(0.0112)	(0.0112)	(0.0112)	(0.0112)
Prop of English medium govt school		-0.114*			-0.0552
		(0.0591)			(0.0665)
Professional training difference			-0.102		-0.138*
			(0.0827)		(0.0821)
Prop of govt schools in district				0.236***	0.237***
				(0.0609)	(0.0618)
Observations	23072	23072	23072	23072	23072

Note: Marginal effects at the median value of MPCE are calculated using the margins command in Stata. Marginal effects are shown in Columns (1)-(5) correspond to Columns (1)-(5) respectively in Table 3.A2.3.

CHAPTER 4

DOES OVEREDUCATION EXIST IN INDIA? EVIDENCE USING RETURNS TO EDUCATION APPROACH

4.1 Introduction

The rising incidence of skill mismatch in the labour market is a cause of concern in both developed (McGuinness et al, 2017; Cultrera et al, 2022; Croce and Ghignoni, 2012; Leuven and Oostereek, 2011) and developing countries (Battu and Bender, 2020; Mehta et al, 2011, Quinn and Rubb, 2006). The concept of skill mismatch is very broad and includes several quantitative and qualitative disparities (International labour Organisation [ILO], 2019). One of its most discussed forms in the literature arises in terms of overeducation, where workers are considered to be overeducated/overqualified if their education level exceeds that which is required by the job. The phenomenon of overeducation was first brought to notice by Richard Freeman (1976) in his study of the US labour market when the over-supply of higher educated graduates resulted in a substantial reduction in the returns to schooling. In fact, this concept of overinvestment in education was intrinsic in the existing labour market theories as well. For instance, the job competition model postulated by Thurow (1975) argued that people invest in education as a defensive necessity to secure a better position in the job queue. Thus, education overinvestment translates into overeducation. According to Charlot and Decreuse (2005), self-selection into education is inefficient and causes overeducation due to too many individuals willing to invest in education. This is primarily because workers fail to internalize their education decisions on the employment opportunities of others. According to the Periodic Labour Force survey 2019-20 (PLFS, 2021), in India approximately 65 percent of the population is in the working age group (15-59 age group). This provides the opportunity of reaping benefits out of a potential demographic dividend. However, the rising education levels accompanied by a dearth of appropriate jobs might be a stumbling block in the process (Sharma and Sharma, 2017). The objective of this chapter is to estimate overeducation in the Indian labour market for elementary occupations, where education requirements are low.

According to the recent estimates provided by the ILO, around 258 million workers are overeducated across 114 countries from all regions and income levels across the world[50]. The discrepancy between the job requirements and education of workers in labour market might arise out of an education system that provides much emphasis on the number of higher educated graduates it produces, simply overlooking the job prospects (Gooptu and Mukherjee, 2022)[51]. The latest trend in the gross enrolment ratio in higher education reveals that there has been a consistent rise between 2013-14 and 2020-21 from 23 to 27.3 percent[52]. However, the employment prospects are grim. Recent data[53] shows that an additional 4.75 crore individuals joined the labour force between 2018-19 and 2019-20 which was almost thrice the employment generated between 2017-18 and 2018-19. From Table 4.1, we see that the total employment in high skilled occupations[54] such as 'Professionals' and 'Technicians and associate professionals' have reduced between 2017-18 and 2019-20. However, the PLFS (2019, 2021) data also shows that despite this, the proportion of higher-educated in the total pool of regular and casual workers has increased from 16.4 percent in 2017-18 to 17.3 percent in 2019-20. Also, there is an increase in the proportion of higher-educated employment in elementary occupations (non-agriculture). This is potentially costly for an economy as it leads to an economic loss on account of endowing individuals with unproductive education. Here, the true potential of workers is utilized sub-optimally. However, the above statistics reveal only a partial image of the Indian overeducation scenario. Simply observing the presence of tertiary-educated workers in elementary occupations and labelling them as overeducated, without

[50] "258 million workers in the world are overeducated for their jobs". https://ilostat.ilo.org/258-million-workers-in-the-world-are-over-educated-for-their-jobs/

[51] This provides an argument against Castello-Climent and Mukhopadhyay (2013) which emphasizes the expansion of tertiary education as a growth-enhancing factor in developing. Gruber and Kosack (2014) explains the 'tertiary tilt' in developing countries, where governments spend a substantial proportion of their education budget on higher education.

[52] AISHE (2018, 2021)

[53] Economic Survey, 2021-22.

[54] The skill classifications provided by National Classification of Occupations (NCO 2004) are provided in Table 4.A.1 in the appendix).

looking into the returns would be incorrect, given the heterogeneity in the nature of work and also the differences in the demand and supply of different types of labour. Instead of going by an exogenously given education level for calculating overeducation, the chapter adopts a top-down approach looking into disaggregated occupation categories within elementary and estimating overeducation at the finest possible classification with Indian data. Thus, instead of calculating overeducation by the official education requirements approach as specified in PLFS, we use the 'returns to education' method (Mincer, 1974) that estimates the market-determined rates of overeducation. This method also circumvents the supply side biases of the conventional techniques using the average education level of workers as a benchmark for defining overeducation. The advantage of the returns approach is that it caters to both the demand and supply sides of the labour market rather than focussing on the supply side alone. Due to gender related differences in the nature of work, we analyse gender specific occupation categories to get precise estimates.

Table 4.1: Distribution of higher-educated regular/casual labourers in the 15-59 age group in different occupation divisions.

Occupation Division	Description	Proportion of higher-educated working in this occupation		
		2017-18	2018-19	2019-20
1	Legislators, Senior Officials and Managers	12	11.43	11.13
2	Professionals	28.44	28.03	29.11
3	Technicians and Associate professionals	27.21	26.83	23.88
4	Clerks	13.53	14.44	13.4
5	Service Workers and Shop and market sales workers	8.08	8.12	9.1
6	Skilled Agricultural and Fishery Workers	.22	.27	.3
7	Craft and Related Trade Workers	4.01	3.99	4.43
8	Plant and Machine Operators and Assemblers	2.81	3.01	3.07
9	Elementary Occupations (non-agriculture)	2.8	2.8	3.24
	Elementary Occupations (agriculture)	0.9	1.06	1.14

Source: PLFS 2017-18, 2018-19 and 2019-20

The chapter delves into disaggregated categories within elementary occupations and calculates the rates of return for different education levels. We specifically focus on

overeducation among higher-educated which might be a very serious issue, given the substantial investment costs associated with it, coupled with the low education requirements in the elementary occupations they are engaged in. According to this approach, the education level within each occupation providing positive significant returns is identified. This is the required education in the occupation. In the next step, any worker with higher levels of education, getting no additional returns, is deemed overeducated. Since we are especially concerned about overeducation of higher-educated workers, we first identify the occupations providing insignificant returns to tertiary educated and then calculate the proportion of such workers within the occupation. In the rest of the chapter, we shall infer the presence of overeducation in an occupation, if there exist higher-educated workers in that occupation who do not get any positive significant returns to their education. In the broadest classification of elementary occupations as a whole, it is found that higher-educated workers are not overeducated, since they get a positive significant return over the previous levels. However, since this provides a very crude estimate given the varied nature of work within elementary occupations, the subsequent regressions are performed for further occupation subdivisions with gender-specific estimates. At the final level of disaggregation, we investigate the extent of overeducation in three occupation groups. For the category 'Domestic and related helpers, cleaners and launderers' we clearly find overeducation among tertiary educated as they get insignificant returns to education. The result is robust to household level controls and district fixed effects. The result for this occupation is restricted to females, which constitutes the majority of this occupation sample considered in the study. The chapter also identifies another occupation category 'Messengers, porters, doorkeepers and related workers' where tertiary overeducation does not exist despite being classified as elementary, since they get positive significant returns over the lower education levels. The result holds for all specifications considered in the study. On the contrary, for workers in 'Mining and construction', once

accounted for the district-fixed effects, there are no additional returns above the 'illiterate' category. Thus, all workers with any education 'below primary' and above are overeducated. This includes higher-educated workers too. This confirms the existence of overeducation in this occupation as well. The preponderance of males in the latter two occupations leads us to restrict the analysis of males in these occupation groups. The chapter also provides a brief discussion on the overall overeducation rates in each occupation, which might include lower education levels as well. This is compared with the pre-existing methods used in the literature that takes the modal education levels for calculating overeducation.

The next section of the chapter discusses the measurement of overeducation, followed by the empirical model and derivation of results. It explains the counterintuitive findings. The section following concludes and offers a sketch of possible extensions.

4.2. Measurement of overeducation

There exist various kinds of skill mismatches in the labour markets. The way each of them is perceived either on behalf of the employers or employees, their measurement, their consequences and determinants make each of them unique. The literature (McGuinness et al. (2017), McGuinness (2006)) broadly defines various concepts under skill mismatch. It includes sub-concepts like vertical mismatch (overeducation/undereducation/over-skilling/under-skilling), horizontal mismatch – the extent to which, workers are engaged in occupations unrelated to their principal field of study, skill gaps – workers lack the necessary skills to perform their current job and skill shortages where employers are unable to fill up vacancies due to lack of qualified candidates for a post. The most commonly researched area within the broad spectrum of skill mismatch is overeducation or over-skilling where an individual possesses more education/skill than what is required by the job[55]. As pointed out by Roy

[55] Mc Guinness et al (2017) points out that most of the research on skill mismatch encompasses overeducation and overskilling. While the latter might provide a more precise estimate of skill mismatch going by the actual

Chowdhury et al. (2021), one of the major demand-side reasons behind the existence of overeducation in elementary occupations is the lack of absorptive capacity of the organized sector, while from the supply side, it is the low employability of the 'educated pool' of workers[56].

According to the literature, overeducation is empirically estimated by comparing the acquired education level of a worker to the level required by the occupation he is working in (Duncan and Hoffman, 1981; Verdugo and Verdugo, 1989; Leuven and Oosterbeek, 2011, McGuinness et al, 2017, McGuinness, 2006). Subsequently, the literature applies three different methods for empirically arriving at the occupation-wise required level of education – *subjective method*, where the required education level depends on the workers' self-assessment of the level of qualifications required to get or do the job versus what they actually acquire (Duncan and Hoffman (1981), Battu et al. (1999), Galasi (2008)). The worker considers himself/herself to be overeducated, undereducated or matched if their level of education is above, below or equal to the required level. In the *job evaluation* method (Rumberger, 1981, Hartog and Oosterbeek, 1988), the same task is bestowed upon professional job analysts who calculate estimates for the required education level for creating occupation dictionaries. The third method commonly followed in the literature is referred to as the *realised matches or empirical method*, where the mean or modal level of education in an occupation is taken as the threshold for comparing with the actual education level and arriving at an estimate for overeducation (Verdugo and Verdugo, 1989; Kiker et al, 1997). Applying the realised matches (RM) method, Kukreja (2018) estimated the overeducation rate of 67.61 per cent in the textile

skill levels of workers, it is quite difficult to measure the same. First, it is very difficult to frame apt questions to get information on the correct type of skills required for a job. This also makes comparison across datasets very cumbersome. Second, it is difficult to realize the worth of underused skills derived from experience, innate ability, training, schooling etc.

[56] According to Battu and Bender (2020), the informal sector acts as a depository for workers who fail to secure formal jobs. Herrera-Idárraga, López -Bazo and Motellón (2015) showed that in a developing country like Columbia, informal workers are more likely to be overeducated.

and clothing industry in India using the NSSO 68th round data (NSS, 2018) on employment and unemployment, which is much above the world average. While the self-assessment method might be subject to bias with lower response rates and exaggeration of educational qualifications, the job evaluation method entails a huge cost. Besides, occupational requirements change frequently, which calls for frequent updates in the occupation dictionaries. In that sense, the RM method is an improvement over the other two. However, one of its major drawbacks is its inability to address the demand for education. As this method entirely relies on the supply side of labour[57] (education level of workers) in a particular occupation, it might result in underestimation of overeducation rates. For instance, in a region, due to the low cost of education or cultural/traditional factors there might exist a huge supply of higher-educated individuals in comparison to demand, driving the national mean education level in a particular occupation to be high. As a result, workers in the same occupations in other regions/states, where absorption is high, will be considered to be undereducated. In this chapter, an alternative to the conventional methods used in the literature – the 'returns to education' approach is adopted where the wages are regressed on the education levels. This methodology is similar to Mehta et al. (2011), where overeducation was measured in four developing countries (Mexico, Philippines, Thailand and India) by first identifying jobs which provided returns to secondary and higher education lower than a 7 percent threshold and then calculated the proportion of workers with these education levels in those jobs. In contrast to Mehta et al. (2011), we focus specifically on the 'graduate and above category' category due to reasons stated in the previous section[58]. However, since the 7 percent threshold is not backed by any economic logic, we just use statistical significance to estimate overeducation.

[57] According to Hartog (2000), the RM method does not cater to the demand side and technological requirements of the job. Besides, Mehta et al. (2011) mentions the problems that arise with multimodal distribution of education in an occupation.

[58] Mehta et al. (2011) considered that primary education was required for unskilled jobs since most workers completed that education level. Thus, they only considered secondary and tertiary educated workers to be overeducated, if they received low returns, or because the employers hired so few of them, that the returns could

The study proceeds from the broadest classification of elementary occupations to the narrowest definition of occupations within elementary to get precise estimates. According to the National Classification of Occupations (NCO 2004), occupations have initially been assigned to ten occupational Divisions (with the last division including workers not classified by any occupation), identified by the initial digit in their NCO 2004 occupation code number (Table 4.A.1) on the basis of their job descriptions. The next level of disaggregation is the 30 Sub-Divisions identified by the two digits of the NCO 2004 code number where the first digit indicates the Division in which the Sub-Division falls under. The Sub-Divisions have next been assigned to 116 Groups (first three digits of the code). The groups have been assigned among 439 Families (first four digits). The final level of disaggregation are the 2945 Occupations (last two decimal digits of the code number). The PLFS dataset however provides the information on occupation classifications up to 3 digits. For instance, Elementary Occupations (Divisions and Sub-Divisions are provided in Table 4.A.2) are assigned under Division 9 according to the NCO 2004 classification. Among Elementary Occupations, there are three Sub divisions, each with a unique 2-digit code (with first digit as the Division code), namely – 'Sales and Service Elementary Occupations' (91), 'Agricultural, Fishery and related labourers' (92) and 'Labourers in mining, construction, manufacturing and transport' (93). The next level of disaggregation comprises of Groups. For example, the Sub-division 91 has 7 groups within it, each with a unique 3-digit NCO 2004 code – Street vendors and Related Workers (911), Shoe cleaning and other street services (912), Domestic and Related helpers, cleaners and launders (913), Building caretakers, Window and related cleaners (914), Messengers, Porters, Door Keepers and related workers (915) and Garbage collectors and related labourers (916). A further disaggregation of these 7 Groups into

not be calculated. Under a similar logic, we focus on overeducation among graduates only. In our study too, overeducation might exist among lower education levels. We also discuss briefly an alternate measure of overeducation, that includes lower levels of education too, later in the chapter.

occupation 'Families' is possible. However, due to data limitations, it is not possible to explore the extent of overeducation beyond Groups. The analysis in this chapter focuses on three of these Groups – 'Domestic and Related helpers, cleaners and launders' (913), 'Messengers, Porters, Door Keepers and related workers' (915) and workers in 'Mining and Construction' (931).

The chapter focuses specifically on 'Elementary occupations', which are the so called low-skilled jobs classified by Division 9 according to the NCO 2004 (Table 4.A.1), requiring low levels of education[59]. It follows a top-down approach which looks into narrower occupation categories step by step. Similar to Mehta et al. (2011), this chapter tries to shed light on jobs that are technologically stagnant where rising education level is a reason for concern[60]. One such occupation considered in the present study is the occupation category (within elementary occupations) 'Domestic and related helpers, cleaners and launderers'. In the remaining chapter, we shall refer to this set of workers as 'Domestic workers'[61]. In fact, the study shows that no domestic workers get any returns to education for any education 'below primary' and above. This also justifies the elementary nature of work in this occupation. Since our focus is primarily on graduate overeducation, we can conclude that all tertiary educated

[59] The NCO 2004 occupation codes are based on the International Standard Classification of Occupations (ISCO 88), an occupation classification system published by the ILO. According to the ISCO 88, the lowest skill level appropriate for doing simple and routine manual tasks, (Table 4.A.1) requires primary education. Adjusting to the Indian labour market, the NCO 2004 shows that the highest level of required education is 'up to 10 years of formal schooling'. Thus, secondary education can be the highest education requirement for jobs within Elementary occupations in India. This is however based upon the average job description within an occupation Division.

[60] In jobs which are subject to technical change, a rise in education may not indicate a rise in overeducation, since there is a rise in the educational requirements with advancement of technology. According to Duncan and Hoffman (1981), overeducation can only be a serious issue if production is not redesigned and job requirements are not upgraded. Thus, in effect, when the average education level rises, many educated workers end up in jobs requiring lower education and skills than their own attainment.

[61] Only regular/salaried and casual workers for the study and not self-employed individuals are considered since their salaries might include an unobserved capital component, biasing returns to education estimates. Table 4.A.3 in the Appendix gives the distribution of workers' status for different occupation categories respectively. For instance, among 'Street vendors and Related workers' around 81 percent are self-employed. A more suitable sample in this regard is the domestic labourer category, where, around 82 percent are regular/salaried or casual workers.

workers in this occupation are overeducated. However, we identify another occupation category within elementary, namely - 'Messengers, porters, doorkeepers and related workers', where tertiary graduates secure positive significant returns over the previous levels, negating the existence overeducation. Thus, we observe that there are some elementary occupations where overeducation does not exist, which has not been captured by Mehta et al. (2011). In an earlier paper Roy Chowdhury et al. (2021), which also uses the returns to education approach, finds no additional returns to education above 'below primary' level for elementary workers in mining and construction for the year 2011-12 using the 68th Round NSS data. The present chapter even though uses the 2018-19 data and controls for district level fixed effects, finds similar results. Thus, even in this occupation, all tertiary graduates are overeducated. There is however, a difference in the calculation of wages in the present chapter vis-à-vis Roy Chowdhury et al. (2021), which shall be explained in detail later.

Since the chapter shows that despite the low education requirement in elementary occupations as given by the official categorisation, there are occupations under the elementary category, where workers with tertiary degrees might secure positive significant returns, the basis of the official education requirement categorization can be questioned. It must be noted in this context, that the NCO 2004 education requirements are defined on the basis of the academic qualifications in an average job description within an occupation. First, the 'returns to education' method circumvents this by estimating overeducation by disaggregated occupation categories. Second, even if the education level prescribed holds for all sub-categories within elementary occupations, it cannot obscure away from the positive productivity signal that a higher education level generates to employers, who might prefer higher-educated workers even though the job does not need so. To them, a worker with more education might give a signal of being more productive and adept at performing the job-related tasks, and thus, secure a higher return.

The analysis in the chapter adds to the literature on the returns to education approach (Becker, 1964, 1975; Mincer, 1974; Psacharopoulos, 1981; Psacharopoulos and Patrinos, 2004, 2018; Patrinos and Psacharopoulos, 2020; Peet et al., 2015; Card, 2001), which is based on the human capital theory. The Mincerian wage (Mincer, 1974) which provides the main empirical framework for the human capital theory subsumes that wages are completely determined by the education levels of workers. In other words, productivity is fully embodied and is not dependent on the job requirements. The job competition model proposed by Thurow (1975) on the contrary argued that wages are solely determined by job requirements. The present analysis is midway between the two approaches. Following Mehta et al. (2011), the model in the analysis in this chapter uses the Mincerian wage equation, to determine the returns to different education levels. However, using this 'returns to education' approach, it also calculates the required education level for a particular occupation, which is the education level providing positive significant returns. As a consequence, all those with higher education degrees and getting no additional returns are overeducated. This captures Thurow's argument that earnings are determined only by the required education levels, with no reward for surplus education (Groot, 1996)[62]. The literature mostly concentrates on identifying overeducated workers first, using the conventional methods and then calculating the wage effects of overeducation for such workers (Darko and Abrokwa, 2020; Bahl and Sharma, 2020; Cultrera et al., 2022; Tsai, 2010). However, under the present approach used in the chapter, the technique is the reverse. By endogenously finding out the benchmark education level for calculating overeducation, it circumvents the existing biases of the RM method discussed earlier in this section. This chapter can also be considered to be an important contribution to

[62] In studies like Duncan and Hoffman (1981), Darko and Abrokwa (2020), Cohn and Khan (1995), Sicherman (1991), Verhaest and Omey (2012), Galasi (2008) though overeducated workers secure a positive significant return, their wages are lower compared to matched workers. On the contrary, Groot (1996) shows that overeducated workers face a negative return to overeducation.

the literature on skill mismatch in India (Kukreja, 2018; Sengupta, 2017; Mukherjee and Paul, 2012; Sharma and Sharma, 2017), which is relatively scarce.

4.3. Data and estimation strategy

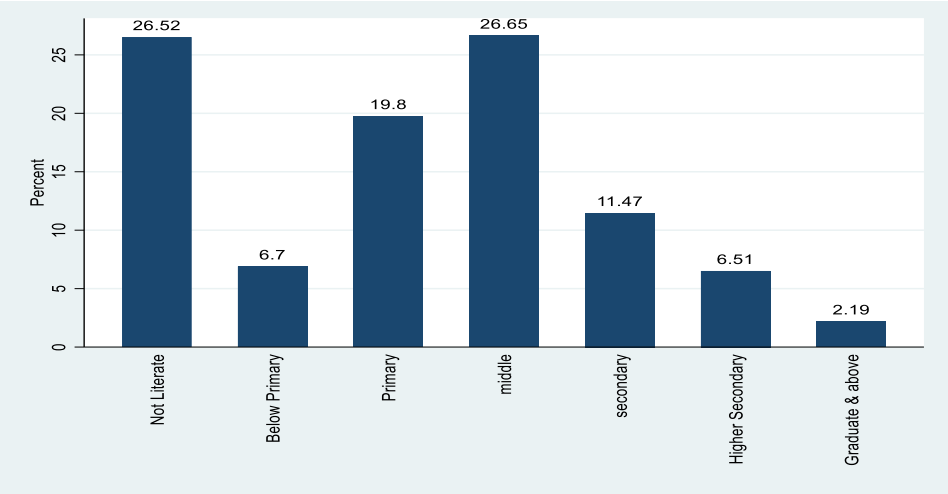
The data in this chapter is drawn from the Periodic Labour Force Survey (PLFS) conducted by the National Statistical Office (NSO) during the period July 2018 to June 2019[63] over the whole territory of the Indian Union. Given the stark mismatch arising out of the presence of higher-educated workers in elementary occupation despite its low education requirement (Table 4.1), the chapter stresses upon "elementary occupations" where 42.5 percent of the regular/salaried and casual workers[64] in the working age population (15-59 years) are employed.

Since information on wages is only available for the current status of workers, the sample consists only those, for whom working status (regular or casual), industry (two-digit NIC 2008 codes) and occupation (three-digit NCO 2004 codes) of the usual principal occupation matches that of the current occupation. In the PLFS terminology, the usual principal status refers to the activity in which the person was employed for the longest time in the reference period of 365 days. The current status on the contrary, provides information on employment and wages in the reference period of the last seven days before the survey period. For casual labourers, the information on wage earnings is provided on a daily basis. The weekly wage earned from the principal industry is obtained from the daily wage earnings provided under current daily activity particulars when the principal activity (*status × 2 digit NIC 2008*

[63] The period 2018-19 has been selected for the purpose of study due to the fact that it is the most recent dataset not susceptive to any external shock in the economy. For instance, the fourth quarter of the next round PLFS 2019-20 data falls within the Covid pandemic affected period. To avoid any unforeseen impact of any pandemic induced variation in wages, the immediate earlier period has been selected.

[64] According to the PLFS terminology, regular workers are those working in others' farm or non-farm enterprises (both household and non-household) and regularly receiving salary or wages (not on the basis of daily or periodic renewal of work contract). On the other hand, casual labourers are those casually engaged in others' farm or non-farm enterprises (both household and non-household) and receiving wages on the basis of their daily or periodic work contract.

code) matches the economic activity in the current daily status. For regular workers on the other hand, the weekly wage is calculated from the monthly earnings under the current status if the regular/salaried wage employment identified by the *status × 2 digit industry code NIC 2008 × 3 digit occupation code NCO 2004* matches in the principal status and the current weekly status[65]. Consequently, the PLFS 2018-19 data provides wage information on 52699 individuals working either in regular/salaried or casual labour activities (except agriculture) in the working age group 15-59 years. Within this wage sample, 11885 or 22.55 percent works in elementary occupations. Figure 4.1 gives the distribution of education categories of workers in these occupations in our sample. Around 2.19 percent of them have tertiary education



Source: PLFS 2018-19

Figure 4.1: Distribution of education levels among elementary occupations (15-59 age group) for regular/salaried and casual wage earners

[65] Weekly wages are taken to capture the volatility in the earnings of casual labourers. For instance, they might be unemployed on some days within a week and might be looking for a job. Thus, even though their daily/hourly wages might be higher in comparison to regular wage earners, that may not reflect their true earnings. Roy Chowdhury et al. (2021) calculated the wages for a regular/salaried or casual worker only when their principal status matched the status given in the current daily particulars, irrespective of the industry or the occupation in which the worker is engaged in. In that sense, the current paper provides a more accurate estimate of individual earnings on the basis of the industry and occupation in which he/she is working.

The descriptive statistics of the three elementary occupations considered in this analysis – 'Domestic and related helpers, cleaners and launderers', 'Messengers, Porters, Doorkeepers and related workers' and workers in 'Mining and construction' are provided in Table 4.2 below. We concentrate on the gender specific occupations, which comprises the majority within each of these occupations. Thus, we focus on female domestic workers (60 percent), male 'Messengers, Porters, doorkeepers and related workers' (92 percent) and male workers in 'Mining and construction (92.3 percent)[66]. The summary statistics for these occupations for all genders combined is given in Table 4.A.4 in the Appendix. In each of these gender specific occupation groups reported in Table 4.2, there exists workers with education above school level. For instance, .3 percent, 9 percent and 1.1 percent of domestic workers, 'Messengers, Porters, Doorkeepers and related workers' and 'Mining and construction workers' have tertiary education. It is also seen that within each of these occupations, the majority proportion of workers have an informal employment[67].

[66] From here onwards in the paper, by domestic workers, 'messengers, porters, doorkeepers and relate workers' and workers in 'mining and construction', we shall mean the female counterpart of the first and male of the two latter groups.

[67] According to the PLFS reports, employees are engaged in informal employment if either they work in a household enterprise or do not receive any social security benefits or are not entitled to any job contract. (Bordoloi et al., 2020).

Table 4.2: Descriptive statistics of the occupation Groups in the study

	Domestic and related helpers, cleaners and launderers (females, n=835)		Messengers, Porters, Doorkeepers and related workers (males, n=670)		Mining and construction labourers (males, n=5773)	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Education levels						
Not Literate	.44	.5	.107	.31	.259	.44
Below Primary	.113	.32	.044	.205	.069	.253
Primary	.172	.377	.101	.301	.221	.415
Middle	.178	.383	.298	.458	.285	.451
Secondary	.08	.265	.196	.397	.104	.305
Higher Secondary	.018	.133	.163	.369	.051	.221
Graduate & above	.003	.054	.09	.286	.011	.106
Demographic Characteristics						
Rural	.266	.442	.369	.483	.82	.384
Regular	.913	.282	.923	.266	.03	.173
Informal	.985	.123	.79	.41	.99	.08
Experience*	29.74	9.97	25.86	12.13	23.20	11.51
Household head completed class 10	.044	.206	.16	.366	.035	.184
Scheduled Tribes	.082	.274	.082	.274	.119	.323
Scheduled Castes	.306	.461	.204	.403	.386	.487
Other Backward Classes	.403	.491	.397	.489	.355	.478
General	.209	.406	.316	.465	.140	.347
MPCE**	2208.303	1439.94	2514.07	1524.71	1476.01	769.684

Source: PLFS 2018-19.

* Following the literature (Duraisamy, 2002; Quinn and Rubb, 2006), the experience term in the regression is considered to be the potential years of experience, which is the age of the worker minus their years of education minus 5. The sample contained few observations where workers had 54 years of work experience. Since the highest age considered is 59, it implies that they joined the market at age 5 (experience = Age - years of education – age of entry into education), which is very unlikely. Thus, we have dropped those observations.
** Monthly Per capita consumption expenditure

Table 4.3: Wage information for elementary occupation Groups according to the wage sample

	Domestic and related helpers, cleaners and launderers (females)		Messengers, Porters, Doorkeepers and related workers (males)		Mining and construction labourers (males)	
	Weekly wage	Monthly wage	Weekly wage	Monthly wage	Weekly wage	Hourly wage
	(1)	(2)	(3)	(4)	(5)	(6)
Not Literate	1048.76	3852.48	1679.87	6314.59	1600.61	38.95
Below Primary	944.68	3346.1	2993.89	9966.61	1800.49	41.02
Primary	1515.48	5581.88	2326.68	8442.13	1711.36	38.91
Middle	1401.88	5233.18	2923.02	11107.74	1758.25	41.44
Secondary	1515.65	5624.79	3147.49	11848.5	1814.65	41
Higher Secondary	1517.22	6016.99	2557.28	10040.14	1766.98	41.32
Graduate & above	1315.12	5260.47	3392.55	13517.27	2213.627	34.72

Source: PLFS 2018-19

Note: All figures are reported in terms of Indian Rupee

Table 4.3 provides wage information on the different occupation groups considered in the present study. It is seen that in the occupation categories – 'Messengers, porters, doorkeepers and related workers' and 'Mining and construction', higher-educated workers get a higher weekly wage (for both regular and casual workers) over lower education levels. However, the wages for domestic workers in the 'graduate and above' category is lower than that of workers with education levels ranging from primary to higher secondary. According to the wage sample considered in the study, 91.3 percent of domestic workers are regular wage earners. The corresponding monthly wages for regular domestic help workers are also reported in the Table 4.3 in Column (2). Here also, a similar trend follows. Also, as almost everyone among Messengers, Porters, doorkeepers and related workers are regular (92.3 percent), the monthly wages for this occupation category are reported in Column (4). We find a higher monthly wage for the tertiary graduates over lower education levels. On the contrary, in the occupation 'Mining and construction', the majority (97 percent) are casual wage earners. Their hourly wages are given in Column (6). The hourly wages for the 'graduate and above' category is lower than the previous education levels. Thus, given these occupation wise variations in wages, a closer probe is required for calculating the returns to education, where several other factors affecting the demand and supply side of labour needs to be accommodated. Also, due to the differences in wage patterns between casual and regular wage earners, one needs to control for the status of work (regular/casual). Given this overall picture of the wage earnings for different occupation groups, we shall proceed with our main regression analysis to estimate returns to education for determination of overeducation.

The standard approach using Mincer (1974) equation is used to calculate the returns to education for individual i employed in sector j.

$$\ln Wage_{ij} = \alpha + \sum_{g \in [1,7]} \beta_g D_{g,i} +$$

$$\gamma_1 Experience_i + \gamma_2 Experience_i^2 + \rho Sector_j + \delta X_{ij} + \varepsilon_{i,j} \qquad (4.1)$$

where $Wage_{ij}$ is the weekly wage, $D_{g,i}$ is a binary variable indicating completion of the education level g. It takes a value of one if grade g is completed by individual i. Thus, the coefficient β_g determines the return to that particular level of education over the previous level.[68] The education levels considered are illiterates, below primary, primary, middle, secondary, higher secondary (including diploma below graduate level) and graduate and above degree. Illiterates constitute the base category. The education dummies are sequential in the sense that each of these levels is a precursor to the next level of education. X_{ij} represents the individual control variables such as social group, status of the worker (regular or casual), nature of employment (formal or informal) and location of the worker (rural or urban). The experience acquired by the worker in the labour market, the square of the experience term[69] and the sector where he/she is working are also controlled.

4.4.Results

Given that the main focus of the chapter is on higher-educated employment in low-skilled occupations, where the education requirement is the lowest, the analysis begins with Division 9 according to the NCO 2004 classification (Table 4.A.1)- Elementary occupations.

[68] For instance, if an individual has completed secondary education, the conditional mean of log wages is $E(lnWage_{ij} |.) = \alpha + \beta_1 + \beta_2 + \beta_3 + \beta_4$ (abstracting away from all other regressors). If the person has completed middle education, the conditional mean of log wages is $E\left(lnWage_{ij}\middle|.\right) = \alpha + \beta_1 + \beta_2 + \beta_3$. Thus, the coefficient β_4 represents the additional returns which secondary education provides over middle education.

[69] The square of the experience term is included to denote diminishing returns to experience.

Within Division 9, the second Sub-Division, 'Agricultural, Fishery and Related labourers' (92) is left out from the analysis due to the difference in the nature of agricultural to non-agricultural occupations. Besides, they are primarily based in rural regions. The results are given in Column (1) of Table 4.4. It is found that tertiary educated workers get significant positive returns to education over higher secondary education and therefore, to all the previous education levels in elementary occupations considered as a whole. However, before declaring that there exists no overeducation among elementary occupations, it is imperative to delve deeper, given the diverse nature of work among different occupation Sub-Divisions within the broad Division 9. Though the NCO 2004 provide an education requirement for each occupation Division based on the average nature of work in it, it would be more robust to look at further sub categories to get precise estimates. Consequently, we shall proceed to a more disaggregated level to two Sub divisions (two-digit NCO 2004 classification) within Division 9 – 'Sales and Service elementary workers' (91) and 'Elementary workers in mining, construction, manufacturing and transport' (93). Columns 2 and 3 in Table 4.4 shows that even in these two occupation categories, there is a positive significant return derived by higher-educated graduates over higher secondary workers and consequently over all workers with lower education levels. However, even this estimate of overeducation might not be precise owing to variability in the nature of work among different occupation Groups (3 Digit) in a Sub-division.

Table 4.4: Regression of log of wages on education levels in the Occupation 'Division' – Elementary Occupations and 'Sub Division'- Sales and Service workers and 'Mining, manufacturing, construction and transport workers'

	Division 9 Elementary workers	Sub Division 91 Sales and service	Sub Division 93 Mining, manufacturing, construction and transport
VARIABLES	Log of week wage	Log of week wage	Log of week wage
	(1)	(2)	(3)
Below Primary	0.0542*	0.146**	0.0425*
	(0.0279)	(0.0555)	(0.0231)
Primary	0.0583	0.0835	0.0423
	(0.0403)	(0.0681)	(0.0361)
Middle	0.0415***	0.115***	0.0322**
	(0.0152)	(0.0337)	(0.0148)
Secondary	0.0418*	0.0152	0.0455
	(0.0218)	(0.0408)	(0.0288)
Higher Secondary	0.0184	0.0494	0.00911
	(0.0157)	(0.0568)	(0.0225)
Graduate & above	0.164***	0.184***	0.146***
	(0.0338)	(0.0524)	(0.0395)
Constant	6.718***	6.659***	6.700***
	(0.158)	(0.172)	(0.197)
Observations	11,885	2,922	8,961
R-squared	0.388	0.523	0.361

Robust standard errors clustered at state level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1. The regressions include individual level controls like experience, social group, region, status of work, nature of employment, sector. State Fixed effects have been included

As already mentioned, the maximum extent of disaggregation in occupation classifications is obtained up to the three-digit NCO 2004 level, namely occupation Groups. The chapter now explores some of these occupation Groups to determine whether there exists overeducated or not. The first step is to investigate the extent of overeducation among Domestic workers (NCO 2004 3-digit Group 913). Also, it must be noted that though 64 per cent of the sample wage earners in the domestic worker category work in household enterprises, the regressions are not specifically restricted to that sector as it leads to a significant reduction of the sample size,

which might affect the precision of the results. Instead, sector dummies have already been included in the regression equation (4.1) controls for this heterogeneity across industries.

Table 4.5: Regression of log of wages on education levels in the occupation 'Group': Domestic Workers and 'Messengers, porters, doorkeepers and related workers' and 'Mining and construction workers'.

	Domestic workers	Messengers, Porters, doorkeeper and related workers	Mining and construction workers
VARIABLES	Log of week wage	Log of week wage	Log of week wage
	(1)	(2)	(3)
Below Primary	0.123	0.219	0.0693**
	(0.0820)	(0.184)	(0.0260)
Primary	0.113	-0.0444	0.0241
	(0.0955)	(0.153)	(0.0409)
Middle	-0.0137	0.211***	0.0314
	(0.0759)	(0.0610)	(0.0210)
Secondary	0.0514	-0.0459	0.00936
	(0.106)	(0.0671)	(0.0370)
Higher Secondary	0.00945	-0.00914	0.00722
	(0.145)	(0.0883)	(0.0223)
Graduate & above	-0.133	0.161**	0.118**
	(0.258)	(0.0645)	(0.0463)
Constant	6.135***	6.849***	6.854***
	(0.401)	(0.322)	(0.277)
Observations	835	670	5,773
R-squared	0.378	0.435	0.315

Robust standard errors clustered at state level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1. The regressions include individual level controls like experience, social group, region, status of work, nature of employment, sector. State Fixed effects have been included.

The regression results are given in Table 4.5. Column (1) in Table 4.5 provides the results for the domestic worker category. The results clearly confirm the presence of overeducation in this occupation as there is no return to 'graduate and above' category. Next, we look into another Group within Elementary Occupation (Division 9) and Sub-division 91- Messengers, porters, doorkeepers and related workers (Group 915). Under this basic Mincer specification in Table 4.5 Column (2), no overeducation is present since workers under the highest education category 'Graduates and above' get positive significant returns. Similarly, 'Mining and construction'

workers get a positive significant return from higher education despite being an elementary occupation.

The usual concern about omitted variable bias related to the workers' innate abilities and that the education coefficients might not correctly reflect the returns to education is likely to arise. However, it is very difficult to find proper instruments for workers' abilities[70]. This concern is lessened as the occupations considered for this study mainly involves menial work which has little to do with innate ability[71]. Still, to account for any ability differences that might affect returns, proxies are incorporated for the person's ability such as the household head's education level (whether completed class 10), and the consumption quartile in which the household falls into[72]. The results are shown in Table 4.6. For the domestic worker category, the results still hold. Thus, overeducation exists. For the other two categories too, we find tertiary graduates getting positive significant returns over lower levels of education[73]. This again reiterates the absence of overeducation in these occupations.

[70] There are studies using information on personal characteristics like the number of siblings, family background etc. (Harmon et al., 2003; Korpi and Tahlin, 2009) in the Mincerian equation for calculation of returns. However, the PLFS does not provide any information on these variables. Card (2001) and Griliches (1977) shows that such ability bias is of less importance.

[71] Factors like the health status of workers would appear to be good proxy for the unobservables driving returns. However, we do not have information on health status in the PLFS dataset.

[72] Kane and Rouse (1995) included parental education and income as measures of family background and ability.

[73] Following Roy Chowdhury et al. (2021), which shows a positive correlation between overeducation and share of industry in state GSDP, we incorporate the average share of industry in the state GSDP over the years 2013-17 as a state level control. Also, employment and returns might be positively linked to the service sector share too. Tables 4.A.5 shows the reports the results. Here, we find the results still hold through for the categories – 'Messenger, Porters, doorkeepers and related workers' and 'mining and construction' workers. For the domestic worker category, primary education provides a positive significant return. However, given that we considered more stringent specifications by introducing state and district fixed effects in Tables 4.6 and 4.7, we consider those to provide more precise estimates.

Table 4.6: Regression of log of wages on education levels after including household controls, for the occupation 'Group': Domestic Workers, 'Messengers, porters, doorkeepers and related workers' and 'Mining and construction workers'

	Domestic workers	Messengers, Porters, doorkeeper and related workers	Mining and construction workers
VARIABLES	Log of week wage	Log of week wage	Log of week wage
	(1)	(2)	(3)
Below Primary	0.0752	0.222	0.0653**
	(0.0950)	(0.175)	(0.0261)
Primary	0.0839	-0.0491	0.0228
	(0.0945)	(0.145)	(0.0395)
Middle	-0.0285	0.124**	0.0294
	(0.0722)	(0.0533)	(0.0213)
Secondary	0.0941	-0.0284	0.00560
	(0.130)	(0.0676)	(0.0364)
Higher Secondary	0.0158	-0.0135	-0.00202
	(0.201)	(0.0671)	(0.0295)
Graduate and above	-0.177	0.193***	0.104**
	(0.296)	(0.0695)	(0.0440)
Constant	6.132***	6.805***	6.821***
	(0.351)	(0.316)	(0.282)
Observations	835	670	5,773
R-squared	0.420	0.481	0.322

Robust standard errors clustered at state level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1. The regressions include individual level controls like experience, social group, region, status of work, nature of employment, sector, household level controls such as household head's education, consumption quartile of the household. State fixed effects have been included.

Table 4.7: Regression of log of wages on education levels in the occupation 'Groups': Domestic Workers, 'Messengers, porters, doorkeepers and related workers' and 'Mining and construction' workers after controlling for district fixed effects

	Domestic workers	Messengers, Porters, doorkeeper and related workers	Mining and construction workers
VARIABLES	Log of week wage	Log of week wage	Log of week wage
	(1)	(2)	(3)
Below primary	0.00126	-0.00711	0.0434
	(0.127)	(0.225)	(0.0312)
Primary	0.148	0.0450	0.0343
	(0.124)	(0.236)	(0.0306)
Middle	-0.0263	0.0996	0.0188
	(0.115)	(0.0893)	(0.0234)
Secondary	0.0747	0.0507	0.00161
	(0.163)	(0.131)	(0.0315)
Higher Secondary	-0.0781	-0.145	-0.0163
	(0.192)	(0.169)	(0.0445)
Graduate & above	-0.456	0.360**	0.0387
	(0.297)	(0.155)	(0.0565)
Constant	5.253***	7.418***	7.501***
	(0.582)	(0.464)	(0.267)
Observations	835	670	5,773
R-squared	0.638	0.761	0.562

Robust standard errors clustered at district level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1. The regressions include individual level controls like experience, social group, region, status of work, nature of employment, sector and household level controls such as household head's education, consumption quartile of the household. District fixed effects have been included.

In the final specification, Table 4.7 includes the household controls and includes district fixed effect to account for district level heterogeneity[74]. For domestic workers and 'Messengers, porters, doorkeepers and related workers', the same results are preserved as in the former specifications. For 'Mining and construction' workers, we find that the 'graduate and above' category loses significance. Thus, using the 'returns to education' approach, according to our definition of overeducation, we confirm that all tertiary graduates among domestic workers and 'mining and construction' workers are overeducated, whereas, there exists no overeducation

[74] The results in the Table 4.7 with the full set of controls is reported in Table 4.A.6 in the Appendix.

among ‘Messengers, porters, doorkeepers and related workers’. From Table 4.2, we see the corresponding proportions are .3 percent of domestic workers and 1.1 percent of ‘mining and construction’ respectively. In the category ‘Messengers, porters, doorkeepers and related workers’, there is no overeducation[75].

From the analysis above, it is seen that despite being classified as an elementary occupation, the graduate and above category still receives positive significant returns for Messengers, porters, doorkeepers and related workers (Group 915). As argued for the broader occupation categories explored so far, an explanation for such an observation for these occupation Groups might be the heterogenous composition of occupation Families (NCO 2004 4 digit codes) and other sub-classifications within them. For instance, the category ‘Messengers, porters, doorkeepers and related workers’ consists of sub-classifications such as messengers, package and luggage porters, deliverers, vending machine money collectors, meter readers etc, where higher-educated workers may add some value, which is not reflected in the job requirements given by the official classification. The educational requirements specified in the NCO 2004 codifications for different occupation Divisions reflect only the average education levels required and does not take into account the heterogeneous nature of occupations at a disaggregated (Sub-Divisions, Groups, Families and so on.) level. However, due to data limitations, it is not possible to carry out the analysis at a disaggregated level beyond the education Groups. In this context, the usual question that might follow is whether this trivialises the existence or need for an official education standard for an occupation. Though these requirements create an upper limit of education requirement, it is plausible that workers are considered to be more productive by employers who hire on the basis of education levels.

[75] We also report separate regressions for regular and casual workers in the occupations studied in this analysis. From Table 4.2, we see that majority of domestic workers and ‘messengers, porters, doorkeepers and related workers’ are regular workers, while majority of ‘mining and construction’ workers are casual labourers. We have performed regressions on the log of monthly wages for the first two categories for regular workers and hourly wage on the third category for casual workers. The results remain unchanged (see Table 4.A.7 in the Appendix).

Thus, even though a higher degree is not a prerequisite, in reality it fetches a higher return as it might signal a higher productivity among workers.

4.4.1. Comparison with alternate measures of overeducation

The results of the regression analysis in Table 4.7 shows that in all the occupations studied, other than the 'graduate and above' category, there exists lower levels of education too, which does not get positive significant returns over and above illiterates. Thus, if we are concerned about the overall overeducation percentage within an occupation, these results must also be taken into account. Under this alternate definition of overeducation, we define a worker as overeducated if he/she has any education more than the level providing positive statistically significant results. This can include any education level starting from 'below primary' to 'graduate and above' providing insignificant returns. From the results in Table 4.7, we can conclude that all domestic workers with any level of formal education are overeducated since none of them get positive returns significantly higher than the illiterate workers. The same results hold for 'mining and construction' workers. For 'Messengers, porters, doorkeepers and related workers' on the contrary, there is no significant positive return up to higher secondary education deeming all of these as excess education over the illiterate category. However, the 'graduate and above' category provides positive significant returns. Thus, all workers in this category having education between 'below primary' and 'higher secondary' are overeducated. We find that 56 percent of the domestic workers, 74.09 percent of workers in 'mining and construction' and 80.23 percent of 'Messengers, porters, doorkeepers and related workers' are overeducated.

Next, the overeducation rates calculated by the 'returns to education' approach are compared with the RM method, conventionally used in the literature, where the mean/modal level of education in an occupation is taken as the threshold for comparison. If the modal level

of education within an occupation is taken (Kiker et al, 1997), it is found to be 0 years for Domestic workers[76]. Using the mode as the benchmark, we find 56 percent of the domestic workers to be overeducated. This estimate matches the endogenously determined overeducation under the 'returns to education' approach. For 'Messengers, porters, doorkeepers and related workers', overeducation rate calculated by the modal value of years of education (8 years) is 45 percent. On the contrary, the 'returns to education' shows that 80.23 percent of workers under this occupation category are overeducated. Thus, we can conclude that there is an underestimation of overeducation under the RM method. It is to be noted that in this occupation, 'graduate and above' workers are overeducated according to the RM method, but not under the returns approach. On the other hand, primary and below primary educated workers are overeducated under 'returns to education' approach, but undereducated under the RM method. Even for the occupation Group 'Mining and Construction' workers, there is an underestimation of overeducation among according to the RM method (16.7 percent), as against the returns to education approach (74.09 percent).

4.5.Conclusions

The chapter addresses one of the most relevant issues infesting labour markets, namely the problem of overeducated workers. The PLFS 2018-19 data reveals an appalling situation wherein 3.8 percent of the tertiary graduates in the working age group are working in elementary occupations. However, simply looking at the nominal figure would be inadequate in revealing the actual degree of overeducation, owing to the variability of work within such occupations. In the study, the sample of wage earners is dissected into different sub-categories of elementary occupations. The chapter follows a new approach of calculating overeducation

[76] The PLFS 2018-19 dataset has information on the actual number of years in formal education for each individual which includes repetition years if any. While calculating overeducation according to the RM method, we only need to compare the actual education attained and job required education levels for which we do not included the year repetitions. The number of years per education level is summarised in Table 4.A.8.

among regular/salaried workers, using the 'returns to education' method among occupations labelled as elementary, which according to the official standards require very low education. Under this method, we first identify the occupations that provide insignificant returns to tertiary graduates and then calculate the proportion of such workers in the occupation. The analysis starts with the broad definition of elementary occupations and proceeds with more disaggregated levels to see the extent of overeducated workers in different occupation sub-categories. The utility of the returns to education approach is the simultaneous consideration of both the demand and supply sides of the labour market. Thus, it is easier to circumvent the bias that might arise by taking predefined education benchmarks for estimating overeducation. The broad Division- 'Elementary occupation' in fact provides positive significant returns to the higher-educated. However, instead of concluding the absence of overeducation in elementary occupations from this observation, the analysis proceeds with further disaggregated occupation categories which are more narrowly defined. Moving one step further, it is seen that both the Sub-Divisions 'sales and service elementary occupations' and 'elementary workers in mining, construction, manufacturing and transport' provides significant returns to the tertiary educated. However, both 'Divisions' and 'Sub-divisions' are still broad categories and hence loosely defined. As a final step, the analysis looks into the most disaggregated occupation 'Groups' on which data is available. We further concentrate on gender specific groups in each of the occupation groups, which constitute the majority. Thus, the study concentrates on female domestic workers and male workers in 'Messengers, porters, doorkeepers and related workers' and workers in 'mining and construction'. Under the original Mincer specification, we confirm the presence of tertiary graduates among domestic workers, who fail to get positive significant returns. The results are robust to inclusion of household level factors and district fixed effects. Thus, the rate of tertiary overeducation among female domestic workers is 0.3 percent. On the contrary, for 'Messengers, porters, doorkeepers and related workers' the returns to education

for tertiary educated workers are positive and significant even after inclusion of the district fixed effects. This shows that despite being an elementary occupation, it rewards workers with higher education. The chapter points out two plausible reasons behind such a result. First, the category 'Messengers, porters, doorkeepers and related workers' subsumes jobs like messengers, package and luggage porters, deliverers, vending machine money collectors, meter readers etc, where higher education might signal a higher productivity and better ability to deliver tasks. Since the NCO 2004 classifications provide the education requirements for the average job description, it might fail to account for this heterogeneity in nature of work between these narrowly defined categories. Another reason might be because employers consider higher-educated workers to be more productive in delivering certain jobs. In that sense, they end up getting higher returns over lower education levels. In such cases, there is an overestimation of overeducation, going by the official education requirements. In other words, since overeducation is estimated by the 'returns to education' approach, which accommodates both demand and supply of labour, it is plausible that a more educated worker might be more demanded by an employer of an elementary occupation, even though its education requirement is low. Lastly, for the occupation 'Mining and construction', we find overeducation among tertiary graduates. The corresponding proportion is 1.1 percent.

We also calculate the overall extent of overeducation in these occupations, which includes any education level above the 'illiterate' category that fails to provide significant returns. Comparing the occupation wise estimates of overeducation with the conventional RM method we find a match in estimates of overeducation among female domestic workers (56 percent), whereas an underestimation of overeducation among 'Messengers, porters, doorkeepers and related workers' (45 percent by RM method versus 80.23 percent under returns to education) and 'Mining and Construction' workers (16.7 percent by RM versus 74.09 percent under returns to education).

The present analysis has certain limitations. First, the usual controversy regarding ability bias associated with Mincerian estimation of returns might pose a concern regarding validity of the estimates under the returns to education approach. Also, the way the education dummies have been created in the empirical model, makes it difficult to find an instrument for each. However, given the menial nature of jobs considered in this study, that becomes less of a concern. Second, it is difficult to incorporate the firm's behaviour in affecting employment, which might be one plausible reason behind higher-educated domestic workers getting a positive significant return over other education levels. To compensate for that, we have included the district fixed effects in the regression to cater to demand and supply-side heterogeneities at district level to some extent.

The model in this chapter can be extended in several ways. While it is difficult to capture the demand side complexities affecting returns and overeducation, it is possible to look into factors that drive greater participation into higher education levels. According to Khanna and Morales (2017), the IT boom in the mid 1990s followed by the establishment of the H1B visa programme granting visas for temporary employment in high skilled occupations in the United States, created a noticeable rise in the incidence of Indian students pursuing engineering to secure gainful employment in the US. Such a policy can act as an exogenous shock affecting overeducation situation in the country. It is also possible to study the impact of a demand side shock such as the Covid 19 pandemic, which resulted in a sudden economic contraction, which led to windfall job losses[77], on overeducation in elementary occupations. The analysis can also incorporate the effects of policies targeted to improve the condition of vocational education in

[77] "Which jobs were 'lost' during India's COVID-19 lockdowns? Evidence from online vacancy postings" https://www.theigc.org/blogs/covid-19/which-jobs-were-lost-during-indias-covid-19-lockdowns-evidence-online-vacancy

the country. For instance, the National Education Policy 2020, which aims at integrating vocational education into formal schooling curriculum may affect education decisions and occupation choices. The occupations studied in this chapter have a minimal proportion of workers with formal vocational training. Thus, the study can be extended to other elementary occupations too. Also, there might exist differences in the extent of overeducation across different caste groups in India. All these remain the future agenda of research.

Appendix 4.A

Table 4.A.1: National Classification of Occupations 2004 one digit classifications (Divisions) along with the Educational requirements

Division	Description	Skill Level	ISCO-88 Requirement	NCO 2004 Education Requirement
1	Legislators, Senior Officials and Managers	Not Defined	Not Defined	Not Defined
2	Professionals	IV	Post Graduate University Degree	More than 15 years of formal education
3	Technicians and Associate professionals	III	First University Degree	14-15 years of formal education
4	Clerks	II	Secondary Education	11-13 years of formal education
5	Service Workers and Shop and market sales workers	II	Secondary Education	11-13 years of formal education
6	Skilled Agricultural and Fishery Workers	II	Secondary Education	11-13 years of formal education
7	Craft and Related Trade Workers	II	Secondary education	11-13 years of formal education
8	Plant and Machine Operators and Assemblers	II	Secondary Education	11-13 years of formal education
9	Elementary Occupations	I	Primary education	Up to 10 years of formal education and/or informal skills

Source: *www.labour.gov.in*, *www.ncs.in*

Note: The International Standard Classification of Occupations (ISCO) is an occupational classification system published by the ILO. The ISCO 88 defined the kind of work and the skill required to perform the job. The NCO 2004 codification of occupation was based on the ISCO 88 pattern, modified to suit the Indian context after considering the informal skills and experience. The Skill level for each identified occupation was decided on the basis of academic and technical qualification and experience requirement and also the average job description.

* No Skill level has been specified for the occupation Division 'Legislators, Senior Officials and managers' due to the huge variation in the skill levels required for executing the tasks and duties in these occupations.

Table 4.A.2: Categories under Elementary occupations

Division	**9**	**Elementary Workers**
Sub -Division	**91**	**Sales and Service Elementary Workers**
Groups (3 digit NCO 2004)	911	Street and Service Elementary Workers
	912	Shoe Cleaning and other Street Services
	913	Domestic and related helpers, cleaners and launderers
	914	Building caretakers, Window and related cleaners
	915	Messengers, porters, door keepers and related workers
	916	Garbage collectors and related workers
Sub-Division	**92**	**Agricultural, fishery and related labourers**
Sub-Division	**93**	**Labourers in mining, construction, manufacturing and transport**
	931	Mining and construction labourers
	932	Manufacturing labourers
	933	Transport labourers and freight handlers

Source: www.labour.gov.in

Table 4.A.3: Distribution of principal working status among elementary occupations in 15-59 age group.

Status of work	Street and Service Elementary Workers	Shoe Cleaning and other Street Services	Domestic and related helpers, cleaners and launderers	Building caretakers, Window and related cleaners	Messengers, porters, door keepers and related workers	Garbage collectors and related workers	Agriculture and fishery	Mining and construction labourers	Manufacturing labourers	Transport labourers and freight handlers	Total
Self-employed	81.07	68.48	17.83	1.60	3.35	10.89	7.44	2.98	6.92	25.93	9.52
Regular	15.10	21.82	67.75	82.24	88.81	81.44	3.51	3.00	53.55	21.13	15.07
Casual	3.83	9.70	13.88	15.17	7.84	6.32	87.76	91.06	39.36	52.49	73.85
Casual (public)	0.00	0.00	0.53	1.00	0.00	1.36	1.29	2.96	0.16	0.45	1.55
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Source: PLFS 2018-19

Note- Self employed includes workers in household enterprise working as own account worker or as an employer or as a helper.

Table 4.A.4: Descriptive statistics of the occupation Groups considered in the study (all gender)

	Domestic and related helpers, cleaners and launderers (n=1324)		Messengers, Porters, Doorkeepers and related workers (n=729)		Mining and construction labourers (n=6188)	
	Mean	SD	Mean	SD	Mean	SD
Education categories						
Illiterate	.343	.475	.124	.33	.288	.453
Below Primary	.09	.286	.04	.21	.07	.251
Primary	.158	.365	.106	.31	.216	.411
Middle	.237	.425	.288	.453	.271	.445
Secondary	.117	.321	.186	.39	.1	.297
Higher Secondary	.04	.193	.162	.368	.05	.215
Graduate and above	.016	.126	.09	.286	.01	.102
Demographic characteristics						
Male	.4	.49	.921	.269	.923	.266
Rural	.347	.476	.366	.481	.813	.39
Regular	.84	.367	.928	.258	.029	.168
Informal employment	.94	.234	.8	.4	.99	.08
Experience	27.326	11.261	25.92	11.83	23.6	11.48
Household head completed class 10	.059	.236	.157	.364	.034	.181
Scheduled Tribes	.082	.274	.08	.271	.126	.332
Scheduled Caste	.28	.449	.2	.4	.383	.486
OBC	.366	.482	.41	.491	.358	.479
General	.272	.445	.314	.464	.132	.339
MPCE*	2260.062	1515.33	2497.32	1507.9	1474.352	767.4572

Source: PLFS 2018-19

*Monthly Per capita consumption expenditure

Table 4.A.5: Regression of log of wages on education levels after including state level variables for the occupation 'Group': Domestic Workers, 'Messengers, porters, doorkeepers and related workers' and 'Mining and construction workers'

	Domestic workers	Messengers, Porters, doorkeeper and related workers	Mining and construction workers
VARIABLES	Log of week wage	Log of week wage	Log of week wage
	(1)	(2)	(3)
Below Primary	-0.176	0.259	0.0705*
	(0.113)	(0.167)	(0.0407)
Primary	0.346**	-0.0848	0.0223
	(0.127)	(0.145)	(0.0387)
Middle	-0.0147	0.138**	0.0477***
	(0.0843)	(0.0576)	(0.0143)
Secondary	0.0970	-0.0166	0.0132
	(0.134)	(0.0562)	(0.0322)
Higher Secondary	0.113	-0.0175	0.0252
	(0.226)	(0.0721)	(0.0350)
Graduate and above	-0.341*	0.223***	0.0857*
	(0.185)	(0.0719)	(0.0457)
Constant	5.830***	6.256***	5.670***
	(0.752)	(0.329)	(0.570)
Observations	832	663	5,768
R-squared	0.263	0.428	0.215

Robust standard errors clustered at state level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1. The regressions include individual level controls like experience, social group, region, status of work, nature of employment, sector, household level controls such as household head's education, consumption quartile of the household and state level controls such as the average share of industry and service in the GSDP over the years 2013-17 (RBI Handbook, various years)

Table 4.A.6: Regression of log of wages on education levels in the occupation 'Groups': Domestic Workers, 'Messengers, porters, doorkeepers and related workers' and 'Mining and construction' workers with district fixed effects reporting full set of controls

	Domestic workers (female)	Messengers, Porters, doorkeeper and related workers	Mining and construction workers
VARIABLES	Log of week wage	Log of week wage	Log of week wage
	(1)	(2)	(3)
Below Primary	0.00126	-0.00711	0.0434
	(0.127)	(0.225)	(0.0312)
Primary	0.148	0.0450	0.0343
	(0.124)	(0.236)	(0.0306)
Middle	-0.0263	0.0996	0.0188
	(0.115)	(0.0893)	(0.0234)
Secondary	0.0747	0.0507	0.00161
	(0.163)	(0.131)	(0.0315)
Higher Secondary	-0.0781	-0.145	-0.0163
	(0.192)	(0.169)	(0.0445)
Graduate and above	-0.456	0.360**	0.0387
	(0.297)	(0.155)	(0.0565)
Experience	-0.000820	0.00830	0.00935***
	(0.0232)	(0.0155)	(0.00334)
Experience squared	6.33e-05	-0.000105	-0.000113*
	(0.000378)	(0.000283)	(6.42e-05)
Scheduled Castes (Base: Scheduled Tribes)	0.0893	-0.0112	0.0339
	(0.242)	(0.137)	(0.0366)
Other Backward classes (Base: Scheduled Tribes)	0.0483	0.123	0.0190
	(0.261)	(0.126)	(0.0343)
General (Base: Scheduled Tribes)	0.206	0.0367	-0.00814
	(0.247)	(0.124)	(0.0407)
Household head completed class 10	0.152	0.0371	0.0430
	(0.137)	(0.161)	(0.0394)
MPCE Q2	0.0574	0.123	0.0242
	(0.113)	(0.113)	(0.0234)
MPCE Q3	0.267**	0.178	0.0499*
	(0.115)	(0.131)	(0.0263)
MPCE Q4	0.400***	0.297*	0.168***
	(0.119)	(0.160)	(0.0482)
Rural	-0.0542	-0.0173	-0.0256
	(0.107)	(0.120)	(0.0214)
Regular	0.0871	0.0609	0.515***
	(0.261)	(0.147)	(0.0859)
Informal employment	-0.0982	-0.343***	-0.656***
	(0.219)	(0.131)	(0.167)
Constant	5.253***	7.418***	7.501***
	(0.582)	(0.464)	(0.267)
Observations	835	670	5,773
R-squared	0.638	0.761	0.562

Robust standard errors clustered at district level. *** p<0.01, ** p<0.05, * p<0.1. District fixed effects have been included.

Table 4.A.7: Regression of log of wages on education levels for Domestic Workers, 'Messengers, porters, doorkeepers and related workers' and workers in 'Mining and Construction' separately for regular and casual labourers

	Domestic workers (regular)	Messengers, Porters, doorkeeper and related workers (regular)	Mining and construction workers (casual)
VARIABLES	Log of month wage	Log of month wage	Log of hour wage
	(1)	(2)	(3)
Below Primary	-0.0925	0.0993	-0.0108
	(0.146)	(0.273)	(0.0405)
Primary	0.221	-0.0442	0.0152
	(0.147)	(0.282)	(0.0356)
Middle	-0.0384	0.0937	0.0155
	(0.126)	(0.0953)	(0.0292)
Secondary	0.148	0.0147	-0.0167
	(0.155)	(0.145)	(0.0244)
Higher Secondary	-0.135	-0.134	-0.0372
	(0.205)	(0.172)	(0.0399)
Graduate & above	-0.497	0.334**	-0.0401
	(0.332)	(0.161)	(0.0441)
Constant	8.007***	8.817***	4.114***
	(0.492)	(0.441)	(0.0948)
Observations	739	622	5,289
R-squared	0.672	0.752	0.464

Robust standard errors clustered at district level are reported in the brackets. *** p<0.01, ** p<0.05, * p<0.1. The regressions include individual level controls like experience, social group, region, status of work, nature of employment, sector, household level controls such as household head's education, consumption quartile of the household. District fixed effects are included.

Table 4.A.8: Years of Schooling for different formal education levels

Years of Schooling	**Education Level**
0	Illiterates and without any formal education
3	Below primary
5	Primary
8	Middle
10	Secondary
12	Higher Secondary
16	Graduate and above

Source: PLFS 2018-19; Years are calculated using distribution of years of education among the working (15-59 years) age population.

CHAPTER 5

CONCLUSION

The thesis investigates three different facets of education sectors in developing countries. Since the role of education is intrinsic in shaping the overall well-being of individuals and facilitating the nation's progress, it becomes imperative to look into different educational outcomes. The core chapters stress school dropouts on one hand, which disrupts the accumulation of human capital and on the other hand, overeducation, involving an excess supply of higher educated individuals in the labour market. Alongside formal school education, there has been an increasing trend of household investment in supplementary education in the form of private tuition for the child to recuperate poor school teaching quality. The thesis investigates each of these issues. First, the impact of a government reallocation of the education budget on school dropouts and overeducation; second, the role of the cost of private tuition on school choice; third, the estimation of overeducation among elementary occupations in India, where education requirements are minimal. The present chapter concludes the thesis by summarizing the results

5.1 Summary of Findings

Chapter 2 verifies the effectiveness of a government reallocation of education budget in favour of school education, on two distinct problems of education sectors in developing countries – school dropouts and overeducation. Accounting for path dependence in education investments, complementarities between household and government institutional investment and also between school and higher education quality, the theoretical model endogenously determines ability thresholds, based on which households decide on education choices for their child. We start with a benchmark equilibrium, where every household sends their child for higher education after school completion, whereas, lower ability ones drop out of school to work as unskilled labour. However, all high-ability children do not secure a skilled job after higher education. Some are forced to work in the semi-skilled market as overeducated workers. The model shows that a shift in government spending towards school education leads to a new equilibrium, with lower dropout rates and lower enrolment in higher education, under certain

conditions. The lower enrolment in higher education also translates into lower overeducation. Skewing government investment in favour of school education has a direct positive impact on school quality through better infrastructure. In the second round, better school quality improves higher education quality, which in turn poses a positive feedback effect on school quality. However, a shift in investment away from higher education leads to a deterioration of teaching quality. This has an indirect perverse impact on school quality. If the direct effect dominates the indirect effect, school quality improves and school dropout reduces. In other words, for the reallocation to be fruitful it must be the case that the school quality is more responsive to school infrastructural expenditure. The results also show that if the improved school quality sufficiently raises the semi-skilled wages, it draws in children on the margin who either dropped out of school or opted for higher education in the benchmark scenario. As a consequence, overeducation reduces too. The other observation that the chapter reveals is that school dropout is imminent and cannot be eradicated using a reallocation policy.

Chapter 3 explores the role of private tuition cost on the household choice between government and private schools at the secondary and higher secondary level. The chapter shows that private tuition cost, which is estimated by the household share of private tuition expenditure for a child in his monthly per-capita consumption expenditure has a negative significant effect on the probability of attending government schools, after controlling for individual and household level variables and the monthly average fee difference between private and government schools. Thus, as the private tuition cost rises, it crowds in private schooling. Since school choice of households also relies on school specific factors, we next incorporate district level school specific factors which constitute the supply-side variables in education. The result remains unchanged even after the inclusion of these controls such as the proportion of government schools in a district, the difference between private and government schools in terms of proportion of professionally trained teachers and also in terms of proportion

of availability of English as a medium of instruction in a district. However, since the empirical model does not shed light on the quality differences between private and government school teaching, we construct a theoretical model to corroborate the empirical results. The theoretical framework is set on the premise that household school choice depends not only on the teaching quality of schools but also on other factors such as school-specific attributes like discipline, medium of instruction, availability of special facilities and affordability. It shows that irrespective of the school-specific attributes, a rise in the cost of private tuition leads to a fall in government school enrolment when teacher quality is better in private schools. Thus, juxtaposing the empirical findings with the theory, we affirm the hard-to-find evidence that private school teaching is of superior quality compared to its government counterparts in the Indian context.

Chapter 4 uses the 'returns to education' approach to calculate overeducation within elementary occupations. Since the study is mainly concerned with the presence of tertiary graduate workers in elementary occupations, we conclude the presence of overeducation, if this category of workers fails to get positive significant returns in such occupations. The analysis starts from the broadest classification and then proceeds into different disaggregated categories and sub-categories within it. While considering elementary occupation as a whole, we find the 'graduate and above' category getting positive significant returns denoting the absence of any overeducation. However, owing to the large variability in the nature of jobs within elementary occupations, we look into the next level of disaggregation. For both the sub-divisions under the umbrella of elementary occupations – 'sales and service elementary occupations' and 'elementary workers in mining, construction, manufacturing and transport', we find the absence of any overeducation, since those with higher education secure positive significant returns. Thereafter, we move to the final level of disaggregation on which information is available. In the occupation group 'Domestic and related helpers, cleaners and

launderers', we find that workers with higher education do not get any positive significant returns. The estimation of overeducation is limited to females, which constitute the majority in this occupation. The result is robust to the inclusion of household-level controls and district-fixed effects. In another occupation, 'Mining and construction' workers with tertiary education gets positive significant returns. However, the significance is lost, once we incorporate the district-fixed effects. Thus, in these occupation groups, tertiary educated workers are overeducated. We also identify another occupation group, namely, 'Messengers, Porters, Door Keepers and Related Workers', where there exists no overeducation since higher-educated workers get positive significant returns. We focus on male workers in both these above occupation categories. The chapter points out two reasons behind this unconventional finding. First, the official occupation requirements specified by the NCO 2004 occupation classification provide education requirements according to the average job description in the broad occupation divisions. It fails to capture the heterogeneity in the nature of work and differences in the demand and supply of labour. Second, even if the education requirements are applicable for all categories and subcategories within elementary occupations, higher education levels might signal higher productivity under some occupation groups, and provide a significant positive return. The chapter also calculates the overall extent of overeducation in these occupations, which includes any education level above the 'illiterate' category getting insignificant returns. For instance, among 'Domestic and related helpers, cleaners and launderers' and 'Mining and construction' workers, there is no positive significant return for any education level over illiterates. Thus, any educated worker in this category is overeducated. For 'Messengers, Porters, Door Keepers and Related Workers', all workers with education between 'below primary' and higher secondary are overeducated, since they do not get any significant positive returns over illiterates. However, as higher-educated workers get a positive significant return, they are excluded. Thus, the study shows that 56 per cent of 'Domestic and

related helpers, cleaners and launderers', 74.09 per cent among 'Mining and construction' and 80.23 per cent of workers in 'Messengers, Porters, Door Keepers and Related Workers' in India, are overeducated. Compared with the modal method of overeducation, we find that the estimates match for 'Domestic and related helpers, cleaners and launderers. For the other two categories, there is an underestimation in the proportion of overeducation going by the modal method.

5.2 Limitations

The work has many limitations. Here we mention a few.

In Chapter 2 a major limitation is the assumption that households have perfect information about their child's ability. It treats households to be identical in their income levels. However, it is possible to check, how the interaction between the ability and the income influence education choices. Also, the model only explores choices about two broad categories, namely, school and higher education. Thus, the model abstracts away from choices that might exist even within these categories. For instance, households can choose between vocational and regular education at secondary level of education. In higher education too, one might consider undergraduate and postgraduate levels separately. This can lead to more choices at each stage of decision-making for households. However, the basic framework remains unchanged.

Chapter 3 assumes that private tuition is used only to make up for any teaching quality gap existing between private and government schools. However, in reality, the objective of taking up private tuition might not just be restricted to that. In fact, intense peer pressure and increasing standards of competitive exams might be a strong driving force as well. However, given that this kind of demand arises from both government and private school students, the results of the chapter would remain unaffected as the relative differences between private

tuition costs on school choice would be wiped out. Second, private tuition may have an impact on the quality of teaching in schools. As seen from the literature on private tuition, teachers often adopt corrupt measures by shirking in schools and at the same time coercing students to take tuition from them privately (Biswal, 1999; Bhorkar and Bray, 2018). This possibility has not been explored in the thesis. However, it can be predicted that such a possibility would reinforce the results as bad quality teaching would force students to take private tuition. Besides, there might also exist segregation in the private tuition market as well between good and bad-quality of private tutors, which has not been addressed separately in the analysis. To secure the appropriate quality from the private tuition market, a higher price might be required. Thus, if the price of good quality private tuition is too high, it might crowd in private schools. This observation goes in tandem with the results derived from the model constructed in the third chapter. The literature also shows some private schools implementing business models where they incorporate before or after-school tutoring in their packages (Bray and Ventura, 2022), from which we abstract away in the thesis. However, it can be predicted that this possibility would strengthen our results, since in that case, private schools would charge an even higher fee, while government school students would have to rely on private tuition markets. The chapter however does not explore the social inequality aspect that might arise due to the presence of households forced to send their children to low-quality schools. Since these households are unable to meet the cost of private tuition, the teaching quality gap persists for them.

Chapter 4 which explores overeducation in elementary occupations in the Indian labour market has one major shortcoming in terms of its inability to include individual ability measures in the regression for estimating overeducation. It might lead to an overestimation of the coefficients of education levels. However, that possibility is alleviated to some extent by the inclusion of household-level measures of ability in the regression such as the household's

consumption expenditure and the household head's education. Also, given the sequential nature of dummies for each education level, it is not possible to find an instrument for each level. This concern is alleviated to some extent since the jobs considered for the study are mostly dependent on physical labour. Thus, their dependence on inherent ability measures such as IQ becomes less important. Instead, measures of the workers' health would have been more beneficial. However, we do not have information on that in the PLFS dataset. Second, we do not have information on quality measures of education of the worker, which might have an impact on returns. Also, it is very difficult to analyse the employer's behaviour while hiring and how that affects the proportion of overeducation of workers, especially since most of the employment considered in this chapter is informal.

5.3 Policy implications

There are several policy implications derived from the thesis.

The results derived in Chapter 2 imply that given the poor quality of schooling in developing countries, there might be a large improvement in school outcomes due to a marginal improvement in the higher education quality and thus, the quality of teachers. Hence, a reallocation policy biased towards school education requires introspection. The deterioration in higher education quality might largely affect the quality of teachers, which might not fully be recuperated by greater investment in school infrastructure. The chapter also highlights the importance of semi-skilled wages in countering both the school dropout and the overeducation problem. It requires that this wage must be highly responsive to school quality. In this regard, policies implementing the expansion of vocational training become important, which has been rightly emphasized in the National Education Policy in India. The impact of such vocational training programmes on overeducation can also be tested empirically by extending the Chapter 4 analysis to other occupations as well.

Given the impact of private tuition cost on school choice derived in Chapter 3, there could be important policy implications of a private tuition ban or regulation on school choice and quality of education. In the year 2021, the Chinese private tuition market, in an attempt to eradicate social inequality, outlawed private coaching companies from reaping profits from this business. Besides, local governments also attempted to regulate private tuition fees. According to our results, a complete ban would force parents to shift to the private schooling option for their children as the private tuition market would cease to exist. A regulation of the private tuition fee would however increase enrolment in government schools as it remains to be a more affordable option. In India, though there have been several attempts to enforce a ban on private tuition, it is far from successful. This is due to the largely informal nature of the private tuition market (Bhorkar, 2023), where tutors and receivers can decide when and what kind of tutoring to choose. In fact, from anecdotal evidence, it is seen that a large proportion of tuition is provided from either the tutor's or tutee's homes. Sujatha (2014) shows a significant proportion of students taking tuition from their school teachers in four Indian states (Andhra Pradesh, Kerala, Maharashtra and Uttar Pradesh), despite the banning. Thus, it is less likely that a private tuition ban policy would be properly enforced in the Indian context.

5.4 Future Research Agenda

The research carried out in this thesis has opened up avenues for further investigation into specific issues like school dropouts, overinvestment in education leading to overeducation and the functioning of private tuition markets.

Chapter 2 of the thesis primarily links the households' education investment decisions to the supply side of the labour market. However, if the demand side is brought into the model, it is possible to look into the effect of exogenous demand shocks on investment at different levels of education. This might include an investigation into the impact of a changing profile of

Foreign Direct Investment on the education sector. The consequence of such a shock on overeducation can also be checked empirically from Chapter 4. Since, under such a situation, there exists a possibility of better job prospects, overeducation is likely to reduce. Also, the weak responsiveness of school quality to higher education quality, which is the necessary condition for the reallocation policy to reduce school dropouts and overeducation, can be empirically tested

Chapter 3 can be extended to study the role of gender bias in private tuition expenditure on the choice of streams at the higher secondary level. This can also be incorporated into the theoretical model by including gender-specific preferences for school facilities in the utility function. This can contribute to the literature, which already reveals a gender segregation in higher education stream choice (Sahoo and Klasen, 2021), where girls are less likely to take up science or commerce compared to humanities. The research on private tuition in this thesis focuses on the role of private tuition cost on school choice – government versus private. However, a third possibility of dropouts might arise. This case might be especially interesting in the Indian context, which is characterised by a high dropout rate at the secondary level.

As seen from the fourth chapter of this thesis, the returns to education approach cater to both the demand and supply sides of the labour market and endogenously determines the required education level which gives us the proportion of overeducated workers. In this regard, the sudden IT boom in the US associated with the advent of the H1B visa allowing temporary employment abroad might be a good explanatory variable. As Khanna and Morales (2017) show that this exogenous shock attracted a large proportion of young students to opt for computer science which had the potential of securing them high-skilled employment abroad. The study can also be extended with future rounds of PLFS data to study the impact of a demand-side shock induced by the Covid-19 pandemic, which led to massive job losses, on overeducation.

Bibliography

Aaronson, D., Barrow, L. & Sander, W. (2007). Teacher and Student Achievement in Chicago Public High Schools. *Journal of Labour Economics*, 25(1).

Abbott, B., Gallipoli, G., Meghir, C. & Violante, G.L. (2013). Education Policy and Intergenerational Transfers in Equilibrium. *Journal of Political Economy*, *127(6).*

AISHE (2018). *All India Survey on Higher Education 2017-18*, Ministry of Education, Government of India, New Delhi.

AISHE (2021). *All India Survey on Higher Education 2020-21*, Ministry of Education, Government of India, New Delhi.

Alderman, H., Orazem, P. & Paterno, E. (2001). School quality, school cost, and the public/private school choices of low-income households in Pakistan. *The Journal of Human Resources.* 36, 304 – 326.

Angrist, J. & Lavy, V. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics,* 114, 533-575.

Arcalean, C. & Schiopu, I. (2009). Public versus Private investment and Growth in a Hierarchical Education System. *Journal of Economic Dynamics and Control*, 34(4), 604-622.

ASER. (2019). *Annual Status of Education Report 2018*, Aser Centre, New Delhi.

ASER. (2021). *Annual Status of Education Report 2021*, Aser Centre, New Delhi.

Aslam and Atherton (2012). *The 'Shadow' Education Sector in India and Pakistan: The Determinants, Benefits and Equity Effects of Private Tutoring* (ESP Working Paper Series No. 38).

ASSOCHAM (2013). *Private Coaching poaches mainstream education.* Retrieved from: https://www.assocham.org/newsdetail.php?id=4050

Azam, M. (2015). *Private Tutoring: Evidence from India* (IZA Discussion Papers No. 8770) Retrieved from: SSRN: https://ssrn.com/abstract=2526594

Azam, M., Chin, A. & Prakash, N. (2010). *The Returns to English-Language Skills in India* (IZA Discussion Papers No 4802) Retrieved from: https://EconPapers.repec.org/RePEc:iza:izadps:dp4802.

Azam, M. & Kingdon, G. (2014). *Assessing Teacher Quality in India*. (IZA Discussion Paper No. 8622). Retrieved from: https://www.theigc.org/sites/default/files/2015/01/Azam-Kingdon-2014-Working-Paper-1.pdf

Azim Premji Foundation (2017). *Teacher Absenteeism Study*. Retrieved from: https://archive.azimpremjiuniversity.edu.in/SitePages/pdf/Field-Studies-in-Education-Teacher-Absenteeism-Study.pdf.

Bahl, S. & Sharma, A. (2021). Education-Occupation Mismatch and Dispersion in Returns to Education: Evidence from India. *Social Indicators Research*, 153, 251-298.

Baker, D.P., Akiba, M., Letendre, G. & Wiseman, A.W. (2001). Worldwide shadow education: outside-school learning, institutional quality of schooling and cross-national mathematics achievement, *Educational Evaluation and Policy Analysis*, 23, 1, 1–17.

Baker, Eva L., Paul E. Barton, Linda Darling-Hammond, Edward Haertel, Helen F. Ladd, Robert L. Linn, Diane Ravitch, Richard Rothstein, Richard J. Shavelson, and Lorrie A. Shepard. (2010). *Problems with the Use of Student Test Scores to Evaluate Teachers*. Economic Policy Institute Briefing Paper. Retrieved from: https://files.epi.org/page/-/pdf/bp278.pdf

Baland, J.M., & Robinson, J.A. (2000). Is Child Labour inefficient? *Journal of Political Economy*, 108 (4), 663-679.

Banerji, R. & Wadhwa, W. (2015). *Choice and Cost: The role of household decision making and spending on secondary education in Bihar.* International Growth Centre.

Baron, E.J. (2022). School Spending and Student Outcomes: Evidence from revenue limit elections in Wisconsin. *The American Economic Review*, 14(1), 1-39.

Battu, H., Belfield, C.R. & Sloane, P.J. (1999). Overeducation among Graduates: A cohort view. *Education Economics*, 7(1).

Battu, H., & Bender, K. (2020). Educational mismatch in developing countries: A review of the existing evidence. In S. Bradley, & C. Green (Eds.), *The Economics of Education: A Comprehensive Overview* (2 ed., pp. 269-289). Academic Press. Retrieved from: https://doi.org/10.1016/B978-0-12-815391-8.00020-3

Bearse, P., Glomm, G., & Patterson, D. M. (2005). Endogenous Public Expenditure on

Education. *Journal of Public Economic Theory*, 7(4), 561-577.

Becker, G. S. (1964) *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. National Bureau of Economic Research.

Becker, G. S. (1975) *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, Second Edition, Columbia University Press for National ureau of Economic Research

Berry, J. & Mukherjee, P. (2019). *Pricing Private tuition in urban India: Demand, Use and impact.* Retrieved from: https://www.povertyactionlab.org/sites/default/files/research-paper/Pricing-Private-Education_Feb2019.pdf

Bhorkar, S. (2023). Variegated Roles of and Relationships Between Private Tutoring and Schooling: Insights From the State of Maharashtra, India. *ECNU Review of Education,* 1-23. Retrieved from: https://doi.org/10.1177/20965311231167186

Bhorkar, S. & Bray, M. (2018). The expansion and roles of private tutoring in India: From supplementation to supplantation. *International Journal of Education Development*, 62, 148-156.

Biswal, B. J. (1999). Private Tutoring and Public Corruption: A Cost Effective Education System for Developing Countries. *The Developing Economies,* 222-40.

Björkman, M. (2004). *Public Funding in the Educational Sector and its Effect on Test Scores*. Retrieved from: http://neumann.hec.ca/neudc2004/fp/bjorkman_martina_juin_24.pdf.

Blankenau, William. (2005). Public Schooling, College Subsidies and Growth. *Journal of Economic Dynamics and Control*, 29, 487-507.

Blimpo, M. P., Evans, D. K. & Lahire, N. (2015). *Parental Human Capital and Effective School Management Evidence from Gambia*. (World Bank Working paper 7238). Retrieved from: https://ideas.repec.org/p/wbk/wbrwps/7238.html

Bordoloi. M., Pandey, S. & Farooqui, M.H. (2020). *A Comprehensive Overview: Defining and Measuring Informality in India's Labour Market*. Centre For Policy Research, New Delhi.

Bray, M. (1999). *The shadow education system: Private tutoring and its implications for planners, Fundamentals of Educational Planning No. 61*. Paris: UNESCO International

Institute for Educational Planning (IIEP). Retrieved from: http://unesdoc.unesco.org/images/0018/001802/180205e.pdf.

Bray, M. (2013). Shadow education: Comparative Perspectives on expansion and implications of private supplementary tutoring. *Procedia – Social and Behavioral Sciences*,77, 412-420.

Bray, M., & Lykins, C. (2012). *Shadow education: Private supplementary tutoring and its implications for policy makers in Asia*. Comparative Education Research Centre, The University of Hong Kong, and Asian Development Bank.

Bray, M., Shengli, Z., Lykins, C., Wang, D. & Kwo, O. (2013). Differentiated demand for private supplementary tutoring: Patterns and implications in Hong Kong secondary education. *Economics of Education Review*, 38(2014), 24-37.

Bray, M. & Silova, I. (2006). The Private tutoring phenomenon: international patterns and perspectives. https://www.opensocietyfoundations.org/reports/education-hidden-marketplace-monitoring-private-tutoring

Bray, M. & Ventura, A. (2022). Multiple Systems, multiple shadows: Diversity of supplementary tutoring received by received by private-school students in Dubai. *International Journal of Educational Development*, 92,102624.

Buchmann, C. (2002). Getting Ahead in Kenya: Social Capital, Shadow Education, and Achievement. *Research in Sociology of Education*, 13,133–59.

Cameron, A. C. & Trivedi, P.K. (2009). *Microeconometrics using Stata*. Stata Press.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5): 1127−1160.

Castello-Climent, A. & Hidalgo-Cabrillana, A. (2012). The role of educational quality and quantity in the process of economic development. *Economics of Education Review*, 31(4), 391-409.

Castelló-Climent, A. & Mukhopadhyay, A. (2011). Mass education or a minority well educated elite in the process of growth: The case of India. *Journal of Development Economics*, 105, 303-320.

Castro, J.F., Ortega, L., Yamada, G. & Mata, D. (2022). *The Magnitude and Predictors of Overeducation and Overskilling in Latin America: Evidence from PIAAC*. (IZA Discussion Papers, No. 15143).

Chakraborty, T. & Bakshi, S. K. (2016). English Language Premium: Evidence from a policy experiment in India. *Economics of Education Review*, 50, 1(16).

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., & Rogers. F.H. (2006). Missing in Action: Teacher and Health Worker Absence in Developing Countries. *Journal of Economic Perspectives*, 20(1), 91-116.

Chattopadhyay, S. (2012). *Education and Economics: Disciplinary Evolution and Policy Discourse*, Oxford University Press.

Charlot, O., & Decreuse, B. (2005). Self-Selection in Education with Matching Frictions. *Labour Economics*, 12, 251-267.

Chetty, R., Friedman J.N. and J.E. Rockoff (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *The American Economic Review*, 104 (9), 2593-2632.

Clotfelter, C. T., Ladd, H. F. & Vigdor, J.L. (2007). *Teacher Credentials and Student Achievement in High School: A Cross Subject Analysis with Student Fixed Effects* (Working

Paper13617). Retrieved from: https://www.nber.org/system/files/working_papers/w13617/w13617.pdf

Cohn, E., & Khan, S. P. (1995). The wage effects of overschooling revisited. *Labour Economics, 2(1), 67–76.* doi:10.1016/0927-5371(95)80008-l

Croce, G. & Ghignoni, E. (2012). Demand and Supply of Skilled Labour and Overeducation in Europe: A Country Level Analysis. *Comparative Economic Studies*, 54(2), 413-439.

Cultrera, L., Mahi, B., Rycx, F. & Vermeylen, G. (2022). *Education and Skill Mismatches: Unravelling their effects on Wages across Europe* (IZA Discussion paper No. 15108). Retrieved from: https://www.iza.org/publications/dp/15108/educational-and-skills-mismatches-unravelling-their-effects-on-wages-across-europe

Dang, Hai-Anh. (2007). The Determinants and Impacts of Private Tutoring Classes in Vietnam. *Economics of Education Review*, 26(6), 684–99.

Dang, H. & Rogers, F. H. (2008). The Growing Phenomenon of Private Tutoring: Does It Deepen Human Capital, Widen Inequalities, or Waste Resources? *World Bank Research Observer*, 23(2), 161-200.

Darko, C.K. & Abrokwa, K.K. (2020). Do you really need it? Educational mismatch and earnings in Ghana. *Review of Development Economics*, 1-28.

Desai, S., Dubey, A., Vanneman, R., & Banerji, R. (2008). Private schooling in India: A new educational landscape, *India Policy Forum*, 5(1), 1-58.

Dongre, A & Tewary, V. (2015). Impact of Private Tutoring on Learning Levels: Evidence from India. *Economic and Political Weekly*.

Duflo, E., Dupas, P & Kremer, M. (2012a). S*chool Governance, Teacher incentives, and Pupil-teacher Ratios: Experimental Evidence from Kenyan Primary Schools.* (NBER Working paper 17939). Retrieved from: https://ideas.repec.org/p/nbr/nberwo/17939.html

Duflo, E., Hanna, R. & Ryan, Stephan, P.R. (2012b). Incentives Work: Getting Teachers to come to School. *The American Economic Review*, 102(4), 1241-1278.

Duncan, G.J. & Hoffman, S. D. (1981). The Incidence and wage effects of overeducation. *Economics of Education Review*, 1(1), 75-86.

Duraisamy, P. (2002). Changes in returns to education in India, 1983–94: by gender, age-cohort and location. *Economics of Education Review*, 21, 609-622.

Economic Survey (2021-22). *Economic Survey 2021-22*, Ministry of Finance, Government of India, New Delhi.

Economic Survey (2022-23). *Economic Survey 2022-23*. Ministry of Finance, Government of India, New Delhi.

Freeman, R. (1976). *The Overeducated American*. Academic Press.

French, R. & Kingdon, G. (2010). *The relative effectiveness of private and government schools in Rural India*: Evidence from ASER data (DoQSS Working Papers No 10-030). Retrieved from: https://EconPapers.repec.org/RePEc:qss:dqsswp:1003.

Galasi, P. (2008). *The effect of educational mismatch on wages for 25 countries*. (Budapest Working Papers on the labour market 2008/8).

Galor, O., & Zeira, J. (1993). Income Distribution and Macroeconomics. *The Review of Economic Studies*, 60(1), 35–52.

Ghosh, P. & Bray, M. (2020). School systems as breeding grounds for shadow education: Factors contributing to private supplementary tutoring in West Bengal, India. *European Journal of Education*, 342-360.

Gibbs, B.G. & Heaton, T.B. (2013). Drop out from primary to secondary school in Mexico: A life course perspective. *International Journal of Educational Development*, http://dx.doi.org/10.1016/j.ijedudev.2013.11.005

Gilpin, G. & Kaganovich, M. (2009). *The Quantity and Quality of Teachers: A Dynamic Tradeoff*. (CESifo Working Paper Series No. 2516). Retrieved from: https://www.econstor.eu/bitstream/10419/26561/1/592489914.PDF

Glewwe, P., Hanushek, E.A., Humpage, S. & Ravina, R. (2014). School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010. In P. Glewwe (Eds.), *Education Policy in Developing Countries*, Chicago: University of Chicago Press.

Glewwe, P. & Kremer, M. (2006): Schools, Teachers, and Education Outcomes in Developing Countries. In Eric Hanushek and Finis Welch (Eds), *Handbook of the Economics of Education,* Elsevier.

Glick, P., & Sahn, D. E. (2006). The demand for primary schooling in Madagascar: Price, quality, and the choice between public and private providers. *Journal of Development Economics*, 79(1), 118–145.

Gooptu, S., & Mukherjee, V. (2022). School dropout and overeducation in developing economies: Feasibility of a budgetary solution. *Review of Development Economics*, 27(2), 825-852. Retrieved from: https://doi.org/10.1111/rode.12964

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica* 1-22

Groot, W. (1996). The incidence and returns to overeducation in the UK. *Applied Economics*, 28, 1345–1350.

Gruber. L., & S. Kosack (2014). The Tertiary Tilt: Education and Inequality in the Developing World, *World Development*, 54, 253 - 272.

Hanushek, E. A. (1997). Assessing the Effects of School Resources on Student Performance. *Educational Evaluation and Policy Analysis*, 19(2), 141-164.

Hanushek, E. A., Lavy, V. & Hitomi, K. (2006). *Do students care about school quality? Determinants of dropout behaviour in developing countries*. (NBER Working Paper 12737). Retrieved from: https://www.nber.org/system/files/working_papers/w12737/w12737.pdf

Hanushek, E. A. & Rikvin, S.G. (2006). Teacher Quality. In E. Hanushek, F. Welch (Eds.) , *Handbook of the Economics of Education* (pp. 1051-1078). Elsevier

Harma, J. (2011). Low-cost private schooling in India: Is it pro-poor and equitable? *International Journal of Educational Development*, 31(4), 350–356. doi:10.1016/j.ijedudev.2011.01.003

Harmon, C., Oosterbeek, H., & Walker, I. (2003). The Returns to Education: Microeconomics. *Journal of Economic Surveys, 17(2), 115–156.*

Harnisch, D. L. (1994). Supplemental education in Japan: Juku schooling and its implication. *Journal of Curriculum Studies*, 26(3), 323-334.

Hartog, J. (2000). Over-education and earnings: Where are we, where should we go? *Economics of Education Review*, 19(2), 131–147

Hartog, J. & Oosterbeek, H. (1988). Education, allocation and earnings in the Netherlands: Overschooling? *Economics of Education Review*, 7(2), 185–194.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161.

Herrera-Idárraga, P., López-Bazo, E. & Motellón, E. (2015) Double Penalty in Returns to Education: Informality and Educational Mismatch in the Colombian Labour Market. *The Journal of Development Studies*, 51:12, 1683-1701, DOI: 10.1080/00220388.2015.1041516

Hidalgo-Hidalgo, M., & Iturbe-Ormaetxe, I. (2012). Should We Transfer Resources from

College to Basic Education? *Journal of Economics, 105*(1), 1-27.

Hirakawa, Y. & Taniguchi, K. (2020). School dropout in primary schools in rural Cambodia: school-level and student-level factors. *Asia Pacific Journal of Education, 41(3), 527–542.* doi:10.1080/02188791.2020.1832042

International Labour Organization (2019). *Skill and job mismatches in low and middle income countries.*

International Labour Organization (2020). *258 million workers in the world are overeducated for their jobs.* https://ilostat.ilo.org/258-million-workers-in-the-world-are-over-educated-for-their-jobs/

Jackson, C.K. & Mackevicius, C. (2021). *The Distribution of School Spending Impacts.* (NBER Working Paper 28517). Retrieved from: https://www.nber.org/papers/w28517

Jaume, D. (2021). The labour market effects of an educational expansion. *Journal of Development Economics*, 149. https://doi.org/10.1016/j.jdeveco.2020.102619.

Jayachandran, S. (2014). Incentives to Teach Badly: After-School Tutoring in Developing Countries. *Journal of Development Economics*, 108, 190–205.

Kane, T.J, & Rouse. C.Z. (1993). *Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?* (NBER Working Paper 4268). Retrieved from: https://www.nber.org/system/files/working_papers/w4268/w4268.pdf

Khanna. G. & Morales, N. (2017). *The IT Boom and Unintended Consequences of Chasing the American Dream*. (Centre for Global Development. Working paper 460). Retrieved from: https://dx.doi.org/10.2139/ssrn.3015317

Kiker, B.F., Santost, M.C. & Oliveira, M.D. (1997). Overeducation and Undereducation: Evidence from Portugal. *Economics of Education Review*, 16(2), 111-125.

Kim, S. & Lee, J.H. (2010). Private tutoring and demand for education in South Korea. *Economic Development and Cultural Change*, 58(2), 259-296.

Kingdon, G. (1996): "The Quality and Efficiency of Public and Private Education: A Case Study of Urban India," *Oxford Bulletin of Economics and Statistics*, 58(1), pp 57–82

Kingdon, G. (2006). Teacher characteristics and student performance in India: A pupil fixed effects approach. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.530.9030&rep=rep1&type=pdf

Kingdon, G. (2020). The Private Schooling Phenomenon in India: A Review, *The Journal of Development Studies*, 56(10), 1795-1817.

Korpi, T., & Tåhlin, M. (2009). Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000. *Labour Economics, 16(2), 183–193*

Kracke, N., Reichelt, M. & Vicari, B. (2018). Wage Losses due to Overqualification: the Role of Formal Degrees and Occupational Skills. *Social Indicators Research*. 139, 1085-1108.

Kremer, M., Chaudhury, N., Rogers, F.H, Muralidharan, K & Hammer. J. (2005). Teacher Absence in India: A Snapshot. *Journal of the European Economic Association,* 3 (2-3), 658-667.

Kukreja, P. (2018). *Skill mismatch and returns to education in manufacturing: A case of India's textile and clothing industry* (ICRIER Working Paper, No. 364). Retrieved from: https://icrier.org/pdf/Working_Paper_364.pdf

Kumar, A. & Sahoo, S. (2021). *Social Identity and STEM Choice: Evidence from Higher Secondary Schooling in India*. (GLO Discussion Paper, No. 900). Retrieved from: https://www.econstor.eu/bitstream/10419/235971/1/GLO-DP-0900.pdf

Kundu, P. (2019). Deteriorating Quality of Education in Schools: Are Teachers Responsible? *Economic and Political Weekly,* 54(24).

Lahoti, R. & Mukhopadhyay, R. (2019). School Choice in Rural India: Perception and Realities in Four States. *Economic and Political Weekly,* 54(49).

Lakshmanasamy, T. (2017). Demand for Private Tuition: A Quantile Regression Analysis of Household Expenditure on Private Tutoring in India. *Journal of Educational Planning and Administration,* 31(2), 115-139.

Lee, J. Y. (2013). *Private tutoring and its impact on students' academic achievement, formal schooling and educational inequality in Korea*. PhD dissertation, Columbia University

Lee, K.G. & Polachek, S. W. (2017). Do school budgets matter? The effect of budget referenda on student dropout rates. *Education Economics*, 26(2):129–144

Leuven E. and Oosterbeek, H. (2011). Overeducation and Mismatch in the Labour Market, In E.A. Hanushek, S. Machin and L. Woessmann (Eds.), *Handbook of the Economics of Education*, (vol. 4, pp 283-326). Elsevier.

Levin, H. M. (1972). *The costs to the nation of inadequate education*. *A Report prepared for the Select Committee on Equal Educational Opportunity*. Center for Benefit-Cost Studies of Education. Retrieved from: https://repository.upenn.edu/cbcse/55?utm_source=repository.upenn.edu%2Fcbcse%2F55&utm_medium=PDF&utm_campaign=PDFCoverPages

Liu, J., & Bray, M. (2020). Evolving micro-level processes of demand for private supplementary tutoring: Patterns and implications at primary and lower secondary levels in China. *Educational Studies*, 46(2), 170–187. https://doi. org/10.1080/03055698.2018.1555452

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

Majumdar, T. (1983). *Investment in Education and Social Choice.* Cambridge University

Press.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, 20(3).

McGuinness, S. Pouliakas, K. and Redmond, P. (2017). *How Useful is the Concept of Skill Mismatch?* (IZA Discussion Paper Series No. 10786). Retrieved from: https://www.econstor.eu/bitstream/10419/170770/1/dp10786.pdf

McMahon, W.W. (2004). The Social and External Benefits of education. In G. Johnes & J. Johnes. (Eds.), *International Handbook on the Economics of Education*. https://doi.org/10.4337/9781845421694.00011

Mehta, A., Felipe, J., Quising, P., and Camingue, S. (2011). Overeducation in developing economies: How can we test for it, and what does it mean? *Economics of Education Review*, 30(2011), 1334-1347.

MHRD (2016). *Private Tuition: Extent, Pattern and Determinants*. Retrieved from: https://keithlewin.net/wp-content/uploads/2016/05/7.-Private-Tuition-Extent-Patterns-and-Determinants.pdf

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: Columbia University Press.

Morsy, H. & Mukasa, A.N. (2020). Mind the Mismatch? Incidence, drivers and persistence of African youths' skills and educational mismatches. *African Development Review*, 32, S5-S19.

Mukherjee. A. & Paul. S.B. (2012). Community identity and skill mismatch: A study on Indian labour market. Retrieved from: http://www.isid.ac.in/~pu/conference/dec_12_conf/Papers/AnirbanMukherjee.pdf

Munshi, K., & Rosenzweig, M. (2006). Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy. *The American Economic Review*, 96 (4), 1225-1252.

Muralidharan, K. (2021). *Personnel Management and School Productivity: Evidence from India* (NBER Working Paper Series 28336)) Retrieved from: https://www.nber.org/papers/w28336

Muralidharan, K., Das, J., Holla, A. & Mohpal, A. (2017). The fiscal cost of weak governance: Evidence from teacher absence in India. *Journal of Public Economics*, 145, 116-135.

Muralidharan, K. & Kremer, M. (2007). Public and Private Schools in Rural India. Retrieved from: https://scholar.harvard.edu/files/kremer/files/public_and_private_schools_in_rural_india_final_pre-publication.pdf.

Muralidharan, K. & Sundaraman, V. (2015). The Aggregate Effect of School Choice: Evidence from a Two-stage Experiment in India. *Quarterly Journal of Economics*, 130 (3).

Mussida, C., Sciulli, D. & Signorelli, M. (2018). Secondary school dropout and work outcomes in ten developing countries. *Journal of Policy Modeling*, https://doi.org/10.1016/j.jpolmod.2018.06.005

Nakajima, M., Kijima, Y. & Otsuka, K. (2018). Is the learning crisis responsible for school dropout? A longitudinal study of Andhra Pradesh, India. *International Journal of Educational Development*, 62, 245-253.

NAS (2021). *National Achievement Survey*, Ministry of Education, Govt. of India, New Delhi.

NEP (2020) *National Education Policy*. Ministry of Human Resource Development, Govt. of India, New Delhi.

Nishimura, M., & Yamano, T. (2013). Emerging private education in Africa: Determinants of school choice in rural Kenya. *World Development*, 43, 266–275.

NSS (2018). *Employment and Unemployment NSS 68th Round*, Ministry of Statistics and Programme Implementation, Government of India, New Delhi

NSS (2019). *NSS 75th Round on Education 2017-18*, Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

NSS (2020). *NSS 75th Round on Education 2017-18 Report No. 585*, Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

Nordbloom, K. (2003). Is increased public schooling really a policy for equality?: The role of within-the-family education. *Journal of Public Economics*, 87, 1943-1965.

OECD (2021), *Education at a Glance 2021: OECD Indicators*, OECD Publishing, Paris.

Patrinos, H.A. & Psacharopoulos, P. (2020). Returns to Education in Developing Countries. In S. Bradley & C. Green (Eds.) *The Economics of Education*. Academic Press.

Peet, E.D., Fink, G. & Fawzi, W. (2015). Returns to Education in Developing Countries: Evidence from living standards and measurement study surveys. *Economics of Education Review*, 49, 69-90.

PLFS (various years), *Periodic Labour Force Survey,* Ministry of Statistics and Programme Implementation, Government of India

PRS (2022). *Demand for Grants 2022-23 Analysis*. PRS Legislative Research, New Delhi.

PRS (2023), *Demand for Grants 2023-24 Analysis*. PRS Legislative Research, New Delhi.

Psacharopoulos, G. (1981). Returns to Education: An Updated International Comparison. *Comparative Education*, 17(3), 321-41.

Psacharopoulos, P. & Patrinos, H.A. (2004). Returns to investment in education: a further update. *Education Economics*, 12(2), 111-134.

Psacharopoulos, G. & Patrinos, H.A. (2018). *Returns to investment in education: A Decennial Review of the Global Literature*. (Woking Paper 8402). World Bank.

Quinn, M. A., & Rubb, S. (2006). Mexico's labor market: The importance of education-occupation matching on wages and productivity in developing countries. *Economics of Education Review*, 25(2), 147–156.

Ranjan, P. (2001). Credit constraints and the phenomenon of child labour. *Journal of Development Economics*, 64(2001), 81-102.

RBI Handbook (various years). *Handbook of Statistics on Indian Economy*, Reserve Bank of India.

Romer, P.M. (1990). Endogenous Technical Change. *Journal of Political Economy*. 98(5), 71-102.

Roy Chowdhury, S.,Bros, C., and Gooptu, S. (2021). Education, skill mismatch and wage differentials. In N.V. Varghese and M. Khare (Eds.) *Indian Higher Education Report.* New Delhi: National Institute of Educational Planning and Administration.

Rumberger, R.W. (1981). The rising importance of overeducation in the U.S. labour market. *Economics of education review*, 1(3), 293-314.

Rumberger, R.W. (1987). High School Dropouts: A Review of Issues and Evidence. *Review of Educational Research*, 57(2), 101-121.

Sahoo, S. (2015). *Essays on Economics of Education*. Indian Statistical Institute, Kolkata. Retrieved from: http://hdl.handle.net/10263/6178

Sahoo, S. (2017). Intra-Household Gender Disparity in School Choice: Evidence from Private Schooling in India. *The Journal of Development Studies*, 53(10), 1714-1730.

Sahoo, S. & Klasen, S. (2021). Gender Segregation in Education: Evidence from Higher Secondary Stream Choice in India. *Demography*, 58(3), 987-1010.

Schultz, T. (1961). Investment in Human Capital. *The American Economic Review*, 51(1), 1-17.

Schultz, T. (1963) *Economic Value of Education*. Columbia University Press.

Sen, A. (2010). *Primary schooling in West Bengal*. UNESCO.

Sengupta, A. (2017). *Mismatch between Skills and Jobs in Indian Labour Market During the Post-Reform Era: Estimates with Unit Level Data*. Paper presented at IARIW Conference on

Experiences and Challenges in Measuring Income, Inequality and Poverty in South Asia, New Delhi. Retrieved from: http://old.iariw.org/India/sengupta-abs.pdf

Sheng, Wen (2021, July. 04). *China puts lid on private tutoring business to ensure fairer education for all.* Global Times. https://www.globaltimes.cn/page/202107/1227758.shtml

Sicherman, N. (1991). Overeducation in the labour market. *Journal of Labour Economics*, 9(2), 101-122.

Siddhu, G. (2011). Who makes it to secondary school? Determinants of transition to secondary schools in rural India. *International Journal of Educational Development*, 31, 394-401.

Singh, A. (2013). *Size and Sources of Private School Premium in Test Sores in India* (Working Paper 98) Young Lives.

Singh, A. (2015). Private school effects in urban and rural India: Panel estimates at primary and secondary school ages. *Journal of Development Economics*, 113, 16–32. doi:10.1016/j.jdeveco.2014.10.004.

Sharma, S. & Sharma, P. (2017). Educational Mismatch and its impact on earnings: Evidence from Indian labour market. *International Journal of Social Economics*. 44(12), 1778-1795.

Slater, H., Davies, N. M., and Burgess, S. (2012). Do teachers matter? Measuring the variation in teacher effectiveness in England. *Oxford Bulletin of Economics and Statistics*, 74(5), 629-45.

Stevenson, D. L., & Baker, D. P. (1992). Shadow education and allocation in formal schooling: Transition to university in Japan. *American Journal of Sociology*, 97, 1639–1657.

Su, Xuejuan. (2004). The Allocation of Public Funds in a Hierarchical Education System. *Journal of Economic Dynamics and Control*, 28(12), 2485-2510.

Sujatha, K. (2014). Private tuition in India: Trends and issues. *Revue International d'Éducation de Sèvres*. Retrieved from: https://doi.org/10.4000/ries.3913

Tan, C. (2017). Private Supplementary Tutoring and Parentocracy in Singapore. *Interchange*, 48, 315-329.

Tansel, A. & Bircan, F. (2006) Demand for Education in Turkey: A Tobit Analysis of Private Tutoring Expenditures. *Economics of Education Review*, 25(3), 303–13.

Tooley, J. (2009). *The Beautiful tree: a personal journey into how the world's poorest people are educating themselves*. Cato Institute

Tooley, J., Dixon, P., & Gomathi, S. V. (2007). Private schools and the millennium development goal of universal primary education: A census and comparative survey in Hyderabad, India. *Oxford Review of Education*, 33(5), 539–560.

Thurow, L. C. (1975). *Generating Inequality*. Basic Books.

Tsai, Y. (2010). Returns to overeducation: A longitudinal analysis of the US labour market. *Economics of Education Review*, 29, 606-617.

UDISE (2012). *Unified District Information System for Education.* National Institute of Education Planning and Administration, New Delhi.

UDISE Plus (various years). *Unified District Information System for Education Plus*, Ministry of Education, Government of India, New Delhi.

UNESCO (2016). *Who Pays for What in Education?* UNESCO Institute for Statistics, http://uis.unesco.org/sites/default/files/documents/who-pays-for-what-in-education-national-revealed-through-accounts-2016-en_0.pdf

UNESCO (2019). *New Methodology Shows that 258 Million Children, Adolescents and Youth Are Out of School*. (Fact Sheet no. 56), UNESCO Institute for Statistics, Retrieved from:

http://uis.unesco.org/sites/default/files/documents/new-methodology-shows-258-million-children-adolescents-and-youth-are-out-school.pdf

United Nations (2020). *Ensure inclusive and equitable quality education and promote lifelong opportunities for all*. https://unstats.un.org/sdgs/report/2020/goal-04/

Van de Ven, W. P. M. M. & Pragg, B. M. S. Van. (1981). The demand for deductibles in private health insurance: A probit model with sample selection. *Journal of Econometrics*, 17, 229–252.

Vora, N. & Dewan, S. (2009). *Indian Education Sector: Long Way from Graduation!*. IDFC-SK Securities Ltd.

Verdugo, R.R. & Verdugo. N. T. (1989). The impact of surplus schooling on earnings: Some Additional Findings. *The Journal of Human Resources*, 24(4), 629-643.

Verhaest, D. and Omey, E. (2012). Overeducation, Undereducation and Earnings: Further Evidence on the Importance of Ability and Measurement Error Bias. *Journal of Labour Research*, 33, 76-90.

Woldehanna, T., Endale, K., Hamory, J. & Baird, S. (2021). Absenteeism, Dropout, and On-Time School Completion of Vulnerable Primary School Students in Ethiopia: Exploring the Role of Adolescent Decision Making Power in Household, Exposure to Violence, and Paid and Unpaid Work. *The European Journal of Development Research*, 33, 1349-1389.

Woodhead, M., Frost, M. & James, Z. (2013). Does growth in private schooling contribute to Education for All? Evidence from a longitudinal, two cohort study in Andhra Pradesh, India. *International s of Educational Development*, 33(1), 65–73.