JOBS AND EDUCATION IN INDIA

A Dissertation

Submitted to the Faculty of the National Graduate Institute for Policy Studies (GRIPS)

in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY IN DEVELOPMENT STUDIES

by

Nakajima Maki

March 2017

Abstract

Jobs and education, key drivers of economic development, are so intertwined that their effects on people's lives cannot be distinguished. While education enhancement is itself a development goal, education also equips people for productive employment and increased household income jobs can increase investment in children's education (World Bank, 2012).

India, world's fastest growing major economy, had average annual growth of about 7% in last two decades. Most of this remarkable growth was driven by development of the service sector. However, the recently growing modern service sector, most of which is skill-requiring and labor-saving, cannot absorb all available workers due to population growth. As a result, unskilled workers often can only find work as low-return casual workers. This is one major cause of persistent poverty, notably in rural areas.

The objective of this dissertation is to inquire into the role of jobs and education in economic development by empirically examining job choice and decision making concerning investment in education. The analysis in chapter 2 provides a characterization of the mechanism underlying the diversification of activities of rural households in four eastern states of India. These four states are poor and agrarian, and rural households there commonly work off-farm to mitigate risks and improve

livelihood. This study analyzes survey data for over 25,000 individuals to investigate the contribution of individual, household, community level characteristics to the securing of lucrative jobs and higher income. It is found that workers with higher human capital and more wealth have better access to high-return jobs, while low social status workers have largely casual and low-paying jobs. Based on the findings reported in chapter 2, I propose that factory jobs are a potential occupation for disadvantaged workers.

Chapter 3, focused on education, identified in chapter 2 as one of the most important determinants of obtaining high-return jobs, identifies the determinants of school progression in a dynamic framework by means of analysis of panel data for approximately 1,000 children in Andhra Pradesh state. Child ability at young age is found to have long-lasting positive effects on school progression; higher wages for casual jobs are found to lead to increased school drop-out; and the presence of a factory has a positive effect on school progression in a community.

An integrated approach is proposed combining manufacturing sector development and the provision of enhanced job opportunity information access, which could promote household investment in education and access to high-return jobs, two keys to the fostering of human capital, and in turn, of improved livelihood.

Acknowledgement

First and foremost I would like to express my sincere gratitude to my principal advisor, Professor Keijiro Otsuka, for his continuous guidance and encouragement during the entire period of the PhD program. I have learned not only from his constructive comments but also from his aspiration for research.

I would also like to thank the committee members, Professors Tetsushi Sonobe, Jonna Estudillo, Yoko Kijima, and Nazmul Chaudhury for their kind support which improved my thesis considerably.

In addition, my sincere gratitude goes to Professor Takashi Yamano for his generous support in conducting the survey in eastern India and for his advices in writing this dissertation. I also would like to thank International Rice Research Institute for permitting me to use their data.

I have learned from the instructions by Professors Kaoru Sugihara, Tomoya Matsumoto, and Chikako Yamauchi on economic history, micro economics, econometrics, and other subjects related to my study.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1 Significance of the Issue and Objectives of the Study

Jobs and education are undoubtedly the key factors of economic development, and their association with economic growth and poverty reduction have been extensively researched for decades. While education outcome is a development goal in itself, education also equips people for productive employment (World Bank, 2012). At the same time, increased income from jobs can lead to increased household investment in children's education. "Jobs are the most important determinants of living standards" (World Bank, 2012) and "changes in labor earnings are the largest contributor to poverty reduction" (World Bank, 2013).

India has undergone a rapid economic growth accompanied by changes in labor market and improvement in access to education over the past few decades. The stride made in the country is brought about by formal and informal institutions which are intertwined with labor market and household in an economy. Government program to enhance access to primary education (Sarva Shiksha Abhiyan) and to secondary education (Rashtriya Madhyamik Shiksha Abhiyan) have played an important role in reducing the cost of schooling and led to a remarkable success in school enrollment.

Large scale intervention, midday meal scheme, also significantly promoted schooling especially among poor households. At the same time, increase in household income encourages investment in children schooling, girls in particular since girls have been disadvantaged in educational attainment. Social stratification in India is based upon multiple categorizations such as caste class, gender, and areas, and its long history has placed the people who sit in the low stratum in disadvantage. Government intervention such as Enforcement of the Right of Children to Free and Compulsory Education Act and Mahatma Gandhi National Rural Employment Guarantee Scheme (aimed to provide employment to lower caste groups in rural area) are effective means to narrow the persistent inequality both in education and in occupation.

Recent emergence of new jobs in factory and IT-related industries in India provides new employment opportunities, but its beneficiaries are disproportionate. Narrowing the gap in educational attainment provides wider population with changes to find stable and well-paid jobs which have been occupied by privileged stratification in the society.

In developing countries, most people engaged in agriculture reside in rural areas where poverty rates are high. For the poor in rural areas, increased agricultural productivity is one means of moving out of poverty. Diversifying into non-agricultural jobs is another, and it is commonly observed in developing countries. The share of

income from the nonfarm sector in rural areas has been increasing and in 2007 accounted for about 35% of income in Africa, and 50% in Asia and Latin America by the mid-2000s (Reardon et al., 2007). Nonfarm employment is critical to poor households as a means of coping with insufficient non-farm income due to factors such as decrease in size of operational farm land, soil degradation, production shock resulting from natural disaster. The smoothing of consumption over time is another means, as income tends to fluctuate because agricultural income is nodrmally generated only a few times a year.

In India, the employment share of agricultural workers has been declining since the mid-1970s, and its pace accelerated substantially after 1999 (Eswaran et al., 2009; Lanjouw and Murgai, 2009). This explosive population growth led to rapid growth of the labor force, which could not be absorbed by agriculture alone due to low elasticity of demand for food. Alternative sources of employment for rural households are nonfarm work and migration. Since only a limited number of workers who can migrate, rural nonfarm economy activity is the key source of jobs for rural workers.

There are two distinctive trends in nonfarm sector growth in India. One is that the driving force of nonfarm growth has been the service sector (Fig 1.1.) and the other is the increase in casualization of nonfarm employment (Himanshu et al, 2013). The share of casual employment grew from 24% in 1983 to 29% in 2004 and increased even more

rapidly, to about 40%, in 2009-2010 (Himanshu et al., 2013). Concerns about casual nonfarm jobs include low wages, absence of long-term contacts, and hazards. The trend towards casualization is particularly alarming because disadvantaged households with less assets, less education, and lower social status are more likely to be found in casual employment (Lanjouw and Murgai, 2009; Deininger et al., 2013). This scenario, where the lucrative jobs are mostly taken by better-off households, makes it difficult for the poor to move out of poverty, and will tend to widen economic inequality.

Therefore, the key strategy is to create jobs that are high-paying yet accessible to the rural poor; this calls for deep insight into the mechanism underlying the job choice of households.

The main objective of the analysis in chapter 2 is to investigate the determinants of job choice, which should shed light on the abovementioned issues impacting a significant number of rural households in India. The analysis is an attempt to clarify what factors increase the probability of acquiring high-return jobs and increase earnings. In particular, the findings from this analysis will provide an understanding of job choice by the disadvantaged, through a focus on the roles of human capital, gender, and social group, all of which impact negatively on disadvantaged workers.

There is robust evidence of the importance of the role played by education in the selection of occupation. Enhancement of education attainment has been a worldwide

target. Universal primary education achievement was set as one of the goals of the Millennium Development Goals. Huge progress was made towards this goal and by 2015, net school enrollment rate in developing regions had reached 91%.

Accumulated evidence suggests that education attainment has a positive effect on one's people's welfare. In terms of returns to education, each additional year of schooling is found to increase earnings substantially (Psacharopoulos and Patrinos, 2004; Fulford, 2014). Education can also increase the probability of acquiring a better job. For example, Duraisamy (2002) finds that in India higher level of education, particularly a college degree, increases the likelihood of acquiring a regular job. This is consistent with the findings of the discussion of job choice in the previous section. that education increases the probability of securing high-return jobs.

The positive impact of education on children's future jobs and earnings motivate parents to increase investment in children's schooling. However, poor households in India often face various other constraints. Although the Constitution of India provides free and compulsory education to all children of age six to fourteen, and enrollment in primary school has improved dramatically in the past 50 years, many children do not attend school beyond upper primary (Table 1.1), largely because of the high drop-out rate, even in primary school (Table 1.2).

A large number of empirical studies attempt to identify the factors causing Indian children to drop out of school (Dre`ze and Kingdon, 2001; Filmer and Pritchett, 2001; Kochar, 2001; Duraisamy, 2002; Foster and Rosenzweig, 2004; Dostie and Jayaraman, 2006).[1] These studies, using data from the 1980s and 1990s, generally find household characteristics, particularly parental education and wealth, and school quality have a significant impact on school enrollment. However, school participation has increased substantially in recent years, especially for girls. Girls' gross enrollment ratio in lower primary school increased from 64.1% in 1981 to 100.6% in 2013-14 (Table 1.1). In other words, gender disparity, as still large in the 1980s, has narrowed dramatically in the past few decades.

In light of this trend of increased access to basic education and decreased gender inequality, Kajisa and Palanichamy (2010) examined the way in which the effects of constraints recognized in the past literature had changed over time. They find that the influence of household resources, parental education, and gender on basic education has weakened over time. They also find that a main constraint faced by farmers now is poor availability of insurance to equip them to mitigate unexpected shock. Other empirical studies capture the effect of the dynamic labor market, which has been undergoing

[1] Many empirical studies attempt to identify the determinants of investment in schooling outside India (for example, Angrist and Krueger, 1991; Patrinos and Psacharopoulos, 1997; Quisumbing, Estudillo and Otsuka, 2004; Ersado, 2005; Carvarlho, 2012).

changes as a result of the recent globalization in India. Outsourcing has created new opportunities for semi-skilled and skilled workers and has heightened incentive to invest in children's schooling (Oster and Steinberg, 2013). These new job opportunities have also led to increased school participation by socially disadvantaged girls because they are not tied to traditional occupations and hence were able to quickly respond to newly emerging opportunities (Munshi and Rosenzweig, 2006).

The major contribution of this thesis is that it addressed two vital issues for an economy to develop; education and jobs and offer new insights to the relevant literature. One chapter investigates from a viewpoint of jobs and gains an valuable insight into how the rural households diversify their income earning activities. It confirms an important role played by education in acquiring well-paid jobs. At the same time, it suggests the possibility of factory jobs to be an accessible well-paid occupation for disadvantaged households. Another chapter focuses on education by analyzing school progression using a panel data from longitudinal cohort study. The sequential logit model allows one to fully exploit the unique data and offer findings on the long-term effects of children's early age attributes and local labor market on school progression by different educational stage.

The objective of chapter 3 is to understand the causes of Indian children dropping out of school, or in other words, the determinants of school progression. Data from a

panel survey from 2003 to 2013 are examined to elucidate the effects of gender and job opportunities, which are expected to have changed over time.

1.2 Overview of Study Area and Data

This dissertation uses two sets of data for analysis in chapter 2 and chapter 3 respectively. Chapter 2 makes use of primary data from a survey conducted in 2015 covering rural areas in four eastern states: Bihar, Odisha, Uttar Pradesh, and West Bengal. These states are poor agrarian regions of India. Average monthly per capita expenditure in these four states is much lower than the country average for both urban and rural areas. All four states have recently experienced a large decline in agriculture share in gross state domestic product (GSDP) recently. The four states average share of agriculture in GSDP dropped from 27% in 2004-05 to 19% in 2013-14. Comparison to the national level decline from 19% to 14% during the same period reveals that these states underwent a drastic economic transformation. Most of the decline in agriculture share in 2004-05 was absorbed by an increase in the share of non-manufacturing industry at the national level as well as in three of the four states: Odisha, Uttar Pradesh, and West Bengal. This was not the case in Bihar, where the share of service increased significantly. At the country level, the share of service in employment stayed constant even though its share of GDP increased. This may have been due to the fact that the service sectors that have grown in recent years are mostly knowledge-intensive and do

not use as much labors as traditional services. The manufacturing sector saw neither increase or decrease in GDP share and employment share. This is consistent with the long-term trends observed in Figure 1.1. The government set the ambitious target of increasing the manufacturing contribution to GDP from 15% in 2015 to 25% by 2025. A program ”Make in India” was launched in 2014 to promote industrialization through deregulation of industrial investment, creation of special economic zones (SEZ), and development of so-called industrial corridors. These efforts are expected to lead to extensive development of the manufacturing sector, which is more labor-intensive than modern service sectors. Therefore, the factory jobs created through industrialization could benefit the workers with relatively little education and skill.

This study is one of the first to analyze the data used in Chapter 2 since it has not been made available to public; it can fully make use of its advantages. First, it covers areas where poverty is prevalent and poor farmers commonly diversify their activities. Second, detailed information about the activities can be classified into fine categories which allow a detailed job choice estimation. Third, the data includes occupation and wage information for all individuals in each household, including those who are currently migrant workers. These advantages enable more a comprehensive analysis of the mechanism underlying activity diversification.

Chapter 3 uses a dataset from a survey conducted by Young Lives in Andhra Pradesh state.[2] Andhra Pradesh is the fourth largest state by area in India, with a population of over 84 million in 2011. Its economy is more developed than most states of India. Its average monthly per capita consumption is higher than the country average (Table 1.3), and the service sector accounts for more than half of the state's GDP. The GDP share of agricultural workers is more than half even though that of service sector workers is only about a quarter. This implies that Andhra Pradesh has been successfully developing labor-saving modern services. Although the data has been used by many study on schooling, to the best of author's knowledge, this study is the first to investigate the long-term school progression and to examine the relation with dynamic labor market.

The data covers both rural and urban areas in the state, which gives variation in labor market and schooling environment across the sample households. The design of the survey is unique in that it follows two same age groups for a long period of time, from 2003 to 2013. Using this data, it is possible conduct a detailed analysis of school progression following specific individuals from primary to secondary school. It will afford valuable insights as to the drivers of completion of and transition to different levels of education.

[2] Young Lives is an international study of childhood poverty following 2000 children in India.

1.3 Organization of the dissertation

Chapter 2, “Jobs off the farm: An inquiry into the role of wealth, human capital, and caste in rural areas of eastern India,” analyzes job choice by rural households. Chapter 3, “Dynamics of school progression in rural and urban Andhra Pradesh: The role of gender and job opportunities in a transforming economy,” investigates the factors affecting school progression in the state of Andhra Pradesh. Chapter 4 summarizes the major findings of the previous chapters and discusses the policy implications.

Table 1.1 Gross enrollment ratio at primary and secondary school by gender in India (from 1951 to 2014)

	Lower Primary I-V (6-10 Years)			Upper primary VI-VIII (11-13 Years)			Secondary IX-XII (14-17 years)		
	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
1950-51	60.6	24.8	42.6	20.6	4.6	12.7	N.A.	N.A.	N.A.
1960-61	82.6	41.4	62.4	33.2	11.3	22.5	16.7	4.1	10.6
1970-71	95.5	60.5	78.6	46.3	20.8	33.4	26.8	10.2	19
1980-81	95.8	64.1	80.5	54.3	28.6	41.9	23.1	11.1	17.3
1990-91	94.8	71.9	83.8	80.1	51.9	66.7	33.9	10.3	19.3
2000-01	104.9	85.9	95.7	66.7	49.9	58.6	39	28.4	33.9
2010-11	114.9	116.3	115.5	87.5	82.9	85.2	55.7	48.5	52.2
2013-14	98.1	100.6	99.3	84.9	90.3	87.4	61.9	62.1	62.0

Source: Government of India (2014a)

Table 1.2 Drop-out rates at primary and secondary school by gender in India (from 1951 to 2014)

	Primary I-VIII (6-13 Years)			Secondary IX-XII (14-17 years)		
	Boys	Girls	Total	Boys	Girls	Total
1960-61	61.70	70.90	64.90	NA	NA	NA
1970-71	64.50	70.90	67.00	NA	NA	NA
1980-81	56.20	62.50	58.70	79.80	86.60	82.50
1990-91	40.10	46.00	42.60	67.50	76.90	71.30
2000-01	39.70	41.90	40.70	66.40	71.50	68.60
2010-11	29.00	25.40	27.40	50.20	47.70	49.20
2013-14	21.20	18.30	19.80	48.10	46.70	47.40

Source: Government of India (2014a)

Table 1.3 Average monthly per capita expenditure in urband and rural areas of states under the study (in rupee, 2009-2010)

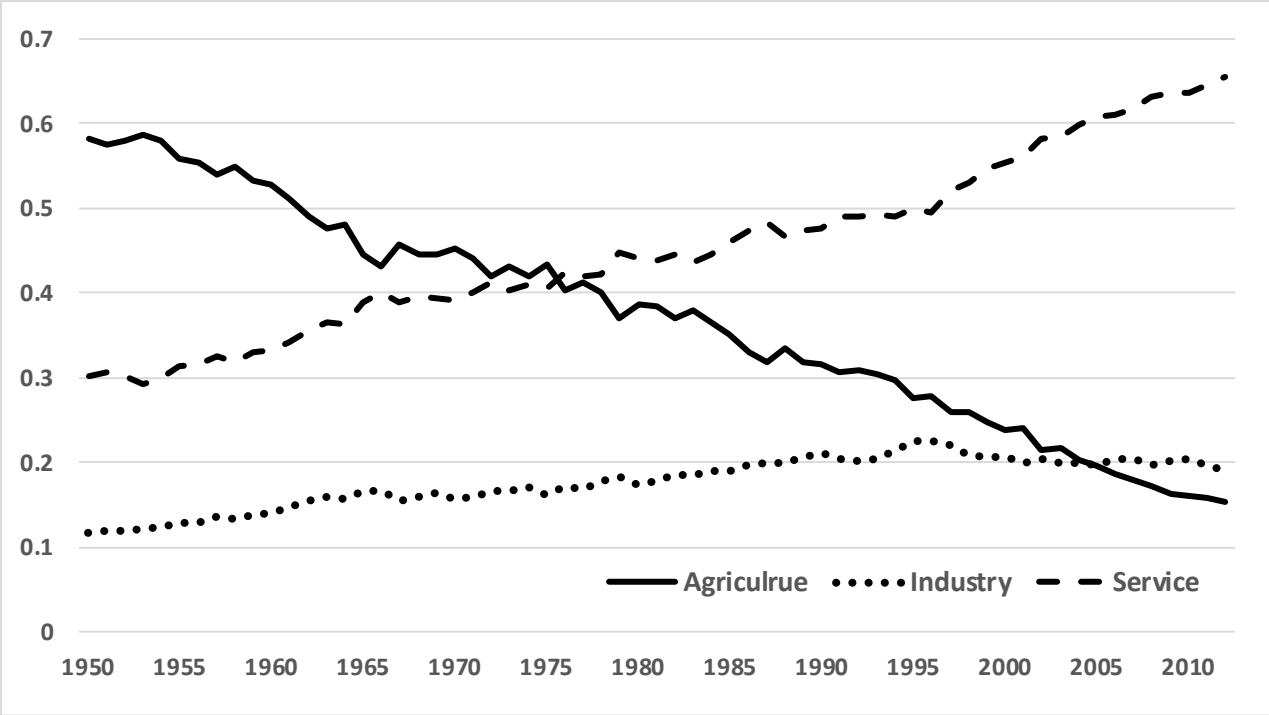
	Rural	Urban
Bihar	780	1,238
Odisha	818	1,548
Uttar Pradesh	899	1,574
West Bengal	952	1,965
Andhra Pradesh	1,234	2,238
All India	1,054	1,984

Table 1.4 GDP share and employment share by sector in the states under the study (2004-05)

		GDP share by sector		Employment share by sector	
		2004-05	2013-14	2004-05	2013-14
Bihar	Agriculture	0.32	0.19	0.77	0.64
	Industry	0.14	0.18	0.08	0.16
	Manufacturing	0.06	0.04	0.05	0.05
	Services	0.55	0.63	0.08	0.20
Odisha	Agriculture	0.23	0.16	0.67	0.62
	Industry	0.34	0.35	0.16	0.21
	Manufacturing	0.12	0.16	0.09	0.09
	Services	0.42	0.49	0.17	0.17
Uttar Pradesh	Agriculture	0.30	0.22	0.66	0.60
	Industry	0.23	0.21	0.16	0.21
	Manufacturing	0.13	0.12	0.11	0.10
	Services	0.47	0.57	0.18	0.19
West Bengal	Agriculture	0.24	0.17	0.49	0.43
	Industry	0.22	0.18	0.22	0.26
	Manufacturing	0.11	0.10	0.17	0.18
	Services	0.54	0.65	0.29	0.30
Four States Average[a]	Agriculture	0.27	0.19	0.65	0.57
	Industry	0.23	0.22	0.16	0.21
	Manufacturing	0.12	0.11	0.11	0.11
	Services	0.50	0.59	0.16	0.22
Andhra Pradesh	Agriculture	0.25	0.19	0.53	0.51
	Industry	0.24	0.24	0.19	0.25
	Manufacturing	0.12	0.12	0.12	0.11
	Services	0.51	0.57	0.28	0.24
All India	Agriculture	0.19	0.14	0.58	0.52
	Industry	0.28	0.26	0.18	0.23
	Manufacturing	0.15	0.15	0.12	0.11
	Services	0.53	0.60	0.24	0.24

[a] The figures are average of four states: Bihar, Odisha, Uttar Pradesh, and West Bengal.

Figure 1.1 Sector-wise gross value added in India from 1950-2012 (at constant 2005 national price in million)



Source: GGDC 10 Sector Database (2014)

Chapter 2

Jobs off the farm:

Wealth, Human Capital, and Social Group in Rural Eastern India

2.1 Introduction

India has undergone rapid economic growth in the past few decades, but poverty persists particularly in rural areas.[3] Three main approaches to improve the living standards of rural households are: intensifying agriculture, increasing income from rural nonfarm activities, and migration to find urban jobs.[4] Population pressure in India has caused land shortages, further constraining the development of the agricultural sector and giving farmers a strong incentive to look for job opportunities outside of agriculture. In India, unlike in other East Asian countries, it is the service sector rather than the manufacturing sector that drives the country's growth.[5] Specifically, the growing segments of the service sector in India are modern service sectors such as finance and

---

3. The proportion of the population under the poverty line is 30.9 per cent in the rural area and 26.4 per cent in the urban area in 2011-12 (Government of India, 2014a).
4. "Off-farm" refers to work other than self-employed farming and thus includes agricultural wage work; "nonfarm" refers to work outside the agricultural sector and thus excludes agricultural wage work.
5. The GDP shares of agriculture, industry, and service were 15 per cent, 19 per cent, and 66 per cent, respectively, in 2012 (GGDC, 2014).

IT-based work made up of highly skilled labour, as well as some low-return traditional service sectors. In September 2014, the Indian government launched a program called, ”Make in India,” to facilitate industrialization through the deregulation of industrial investment, including further liberalization of FDI, creation of special economic zones (SEZ), and development of so-called industrial corridors, all of which are expected to contribute to job creation.

The importance of nonfarm income for rural households has been steadily increasing in developing countries (Reardon et al., 2000; Lanjouw and Lanjouw, 2001; Haggblade et al., 2010; Estudillo and Otsuka, 2016). In rural India, income diversification by farmers is increasingly common (Lanjouw and Shariff, 2004; Himanshu et al., 2013) in the form of consumption smoothing measures to cope with agricultural shock (Kochar, 1999; Ito and Kurosaki, 2007). Recent studies on nonfarm employment in India find that socially and economically disadvantaged households engage more in low-return activities such as agricultural wage work and casual nonfarm work (Banerjee and Knight, 1985; Lanjouw and Shariff, 2004; Borooah et al., 2007; Madheswaran and Attewell, 2007; Lanjouw and Murgai, 2009; Deininger et al., 2013). Additionally, earlier studies observed that households with less wealth, human capital, and lower social status face entry barriers to the nonfarm labour market in India (Kijima

and Lanjouw, 2005; Kijima, 2006; Lanjouw and Murgai, 2009; Saha and Bahal, 2015). These studies also show that those who are likely to face entry barriers tend to rely on personal networks such as families, relatives, and friends to find nonfarm jobs, which are mostly casual and low-paid (Munshi and Rosenzweig, 2006; de Haas, 2010).

In this scenario, disadvantaged farm households continue to have less access to lucrative nonfarm jobs, which leads to a wider income disparity. This disparity is a major concern in eastern India where the economy is predominantly agrarian and poor. In light of this, our study examines the effects of wealth, human capital, and social group on occupational choice and on the monthly and annual income of farm household members, in order to gain an understanding of the increasing income diversification in rural areas of eastern India. In our study, monthly income is the income that workers receive in a month.[6] Furthermore, annual income is the monthly income multiplied by the number of months the workers actually worked in the previous 12 months.[7] Notably,

---

[6]. Monthly income is wage for those engaged in wage work. In case of self-employment, it is the value of production minus paid-out cost per month. For example, the daily wage rate of daily wage earners is converted to monthly by multiplying it by 25, assuming that they work approximately 25 days a month. Therefore, monthly income is a proxy for monthly wage rate in the analysis.

[7]. The questions regarding income are asked on a recall basis for the past year; this is bound to involve measurement errors. However, the question asks the frequency of payment and the number of times actually paid as opposed to total income for the year; this could serve to reduce measurement error.

the existing studies have two shortcomings: inadequate attention to job alternatives and inappropriate classification of nonfarm jobs.

The issue of dealing with limited job alternatives has been addressed, but those studies analyse only rural nonfarm jobs (Lanjouw and Shariff, 2004; Lanjouw and Murgai, 2009; Imai et al., 2015;), the urban labour market (Banerjee and Knight, 1985; Borooah et al., 2007; Madheswaran and Attewell, 2007; Ito, 2009; Deininger et al., 2013), and migratory work (Mosse et al., 2002; Jha, 2013; Agrawal and Chandrasekhar, 2015). However, when a household diversifies its income sources, it is common sense for them to compare various possible alternatives in the labour markets and hence, it is important to consider agricultural, rural nonfarm, and urban jobs as alternative occupational choices for rural households. The second issue concerning inappropriate classification is crucial because nonfarm jobs are highly heterogeneous. Although that heterogeneity of nonfarm jobs is recognized in some studies (Banerjee and Knight, 1985; Imai, et al., 2015), most studies of nonfarm jobs in India only classify those jobs in terms of regular/casual (Lanjouw and Shariff, 2004; Borooah et al., 2009; Lanjouw and Murgai, 2009). In this study, economic activities are grouped into eight categories to identify different determinants of employment across a variety of job types.[8]

8. The eight categories are self-employed farming including household activities, agricultural wage

Our study addresses these two issues by means of a large-scale household survey of 7,085 rural households in eastern India. The survey was conducted in 2015 and respondents were asked about local off-farm jobs as well as migration. Concerning migration, respondents in the survey were asked how workers obtained the jobs, i.e. through a network or by other means.[9] In order to address the issue of selectivity, a two-stage procedure is employed. In the first stage, occupational choice is estimated with a multinomial logit estimation; and in the second stage, monthly and annual income functions are estimated. Furthermore, it is hypothesized that: (1) wealth and human capital (measured by years of schooling and English language skills) have positive impacts on participation in skill-requiring jobs (service/self-employment); (2) Scheduled Caste (SCs) and Scheduled Tribe (STs) members have less access to high-return jobs (service/self-employment and factory work); and (3) there is no significant difference in monthly income between SC, ST members, and others in the same occupational category, as labour markets are likely to be competitive (Becker, 1957).

---

work, local construction, local factory, local service and self-employment, migratory construction, migratory factory, and migratory service and self-employment.

9. The 68th National Household Survey is the most recent (at the time of writing) nation-wide household survey on employment available to the public in India. However, the data does not provide information about migration and the means by which people found migration work. Furthermore, the survey was conducted in 2012 and the results are now too old to accurately describe current employment activities in rural areas.

The rest of the chapter is structured as follows. Section 2.2 discusses the characteristics of the present labour market in India, based on a review of relevant literature. In section 2.3, data are presented with descriptive analysis, followed by an explanation of the estimation framework in section 2.4. The results of regression analyses are examined in section 2.5. Finally, section 2.6 presents the conclusion and policy implications.

2.2 The Indian labour market

2.2.1. Social group and occupations in India

Indian society has a complex social stratification system that differentiates people in many ways, particularly by class, caste, religion, gender, and region. The caste system, one of the main stratifications, is based on Hinduism and consists of hierarchical categories linked to traditional occupations. Specifically, caste determines access to wealth, power, and privilege; since it is a hereditary and endogamous system, it has led to social and economic inequality between caste groups. Furthermore, Hinduism is the dominant religion in India, which is adhered to by 80 per cent of India's population; Muslims account for 14 per cent; and a small number of Sikhs and

Christians make up most of the remainder (Indiastat, 2016). Upon independence in 1947, the Constitution of India introduced a reservation system that aimed to equalize opportunities across castes and tribes by establishing two lists of groups to be provided with reserved places in government jobs, political representation, and educational institutions. The SC group consisted of the bottom of the caste hierarchy, while the ST group was made up of socially and economically marginalized tribes.[10] Some years later, an eligible group was expanded to include Other Backward Classes (OBCs), which consisted of 2,339 disadvantaged groups.[11]

The Indian government provides its rural poor with job opportunities through the Mahatma Gandhi National Rural Employment Guarantee Act (NREGA), a rural workfare program implemented in 2005 that offers up to 100 days of unskilled manual work for rural workers and guarantees a minimum daily wage of Rs. 60. Under this program, work was provided for 50 million rural households in 2012/13; SCs and STs accounted for 51 per cent of total employment (Government of India, 2012). Bihar, Odisha, Uttar Pradesh, and West Bengal are the four states included in our study. In these states, employment was provided for 20 million rural workers, equivalent to one

[10]. For the history and evolution of the reservation policy in India, see Osborne (2001).

[11] . Unlike STs and SCs, OBCs are not given reservations in legislature and the quota is lower than their population share.

quarter of total employment in these states. Additionally, the most common work taken up in these states was related to rural connectivity, flood control and protection, and water conservation and water harvesting (Indiastat, 2016). These political efforts to equalize the economic and social disparities among the caste groups have been deemed unsuccessful in improving the status of the disadvantaged classes, as economic inequality between castes persists (Deshpande, 2000; Borooah, 2005; Kijima, 2006; Ito, 2009). Kijima (2006) highlighted that the differences in living standards between STs and non-SCs/STs can be explained in part by geographical differences, as STs are for the most part concentrated in remote areas. Meanwhile, the differences between SCs and non-SCs/STs are largely marked by variations in education and occupational choice. Notably, in urban areas, the association of caste with occupation is not as strong as it used to be (Hoff and Pandey, 2006; Cain et al., 2010), which is why it is important to investigate households' current occupational diversification strategies in rural areas, with special attention to the disadvantaged castes.

2.2.2 Off-farm employment

Structural change in the course of economic development increases the importance of nonfarm wage employment, which has been demonstrated by many studies involving

developing countries (Reardon et al., 2000; Lanjouw and Lanjouw, 2001; Estudillo et al., 2008; Otsuka et al., 2008; Haggblade et al., 2010; Himanshu et al., 2013; Estudillo and Otsuka, 2016). The literature on off-farm employment contains many attempts to identify the determinants of participation of rural household members in nonfarm activities. The motives for this diversification can be categorized as “push” and “pull” factors (Barrett et al., 2001; Haggblade et al., 2007; Davis et al., 2009). Push factors, which lead to insufficient income from farming, include declining productivity in farming, incomplete rural markets for credit and insurance, and declining availability of farmland. These factors motivate households to diversify their income sources away from farming so as to raise their income levels and smooth their consumption over time. In India, population pressure, which produces an increasing scarcity of land, is a major factor pushing rural households into nonfarm jobs. Indeed, empirical studies in India show that participation in nonfarm employment is positively associated with household size and negatively associated with the size of land assets (Micevska and Rahut, 2008; Imai et al., 2015).

Comparatively, pull factors, which contribute to increased demand in and improved access to nonfarm labour markets, include nonfarm sector development, urbanization, and improved infrastructure. They also include household and household member

attributes such as wealth and human capital. Among these factors, two have been found that promote a consistently positive impact on off-farm participation; they are household wealth (Barrett et al., 2001, Corral and Reardon, 2001; Kung and Lee, 2001; Lanjouw and Shariff, 2002); and education level (Estudillo and Otsuka, 1999; de Janvry and Sadoulet, 2001; Cherdchuchai and Otsuka, 2006; Matsumoto et al., 2006; Eswaran et al., 2009; Huang and Rozelle, 2009; Takahashi and Otsuka, 2009). Some studies have identified differential impacts of education on job choice. Return on education has been found to be higher for regular jobs than for casual jobs (Dutta, 2006; Kurosaki and Khan, 2006; Ito, 2009); higher for migratory jobs than for local jobs (Yúnez-Naude and Taylor, 2001); and higher for urban jobs than for rural jobs (Cain et al., 2010). Furthermore, Lanjouw and Murgai (2009) analyse five rounds of nationally representative surveys in India and observe that (1) those with higher education and social status engage more in regular nonfarm wage employment; (2) those with lower education engage more in casual nonfarm wage employment; and (3) self-employment is undertaken by those with lower as well as higher education. These differential impacts of education and social status on the type of employment suggest the importance of accounting for heterogeneity of nonfarm jobs in the analysis of job choice. Our study carefully

categorizes occupation to show how rural households in India diversify their income-earning activities.

2.3 Descriptive Analysis

This survey was conducted by the International Rice Research Institute (IRRI) in four eastern Indian states between March and August of 2015.[12] Sample villages and households in the rural areas were randomly selected through the following procedures: (1) the number of sample villages in each state was chosen to be proportional to the total rice area, while keeping the total number of sample villages at 720; (2) sample villages were randomly selected from the village list in the 2001 Village Census; and (3) for each sample village, 10 households were randomly selected from a household list provided by the village head and other informants. Table 2.1 shows the number of districts, households, and individuals by state. A total of 6,980 households were selected from 86 districts, and among 33,106 individuals, the analysis covered 25,512 individuals over 16 years of age and not in school. The sampled individuals were categorized into four social groups of Hindus with the following caste statuses: Upper Castes (UP), Scheduled Castes (SCs), Scheduled Tribes (STs), Other Backward Classes

[12] . The four states are Bihar, Odisha, eastern Uttar Pradesh, and West Bengal.

(OBCs), as well as non-Hindus, consisting of mostly Muslim households and a small number of Christian and Sikh households. Many studies lump Upper Castes and non-Hindus together as "other," but this is a source of bias. The classification in this study avoids that bias by distinguishing the schema based on caste and other factors through the separate identification of Hindus and then further dividing them into four caste-based groups, as in Mohanty (2006).

Table 2.2 presents the composition of the sampled workers, the proportion of off-farm workers, the proportion of migrants who found jobs through personal networks, and capital endowments by social group. Of the 25,512 individuals, 16 per cent worked off-farm locally and five per cent migrated. The top three destination states of migrants were Delhi, Maharashtra, and Gujarat, accounting for 20.6%, 16.7%, and 10.8% of total migrants, respectively. A larger proportion of SCs and STs worked in the home locality over the others. The personal network through which migrants found jobs consisted of families, relatives, friends, and former co-workers. Others found jobs through institutions such as contractors, recruitment firms, and vocational training centres. Personal networks reduce the transaction cost in labour markets in the presence of information asymmetry where employers do not know the applicants' true abilities

(Kajisa, 2007). Networks also facilitate information acquisition through job seekers.[13] The disadvantaged castes appear to rely more on personal networks than on other means (Table 2.2). This observation supports the observation that relatively poor and low-skilled migrants depend more on networks to expedite migration than wealthier and highly skilled migrants (de Haas, 2010). It also confirms the findings of Munshi and Rosenzweig (2006) that lower caste workers in traditional sectors are tied to social networks in seeking jobs. As a result, those workers have less access to information outside the traditional occupations in local areas to which they have historical connections.

The last two columns of Table 2.2 report the level of human capital in terms of years of schooling and English language skills. The evaluation of English language skills is based on self-report and assigning a value of one if the person who knows at least basic English; otherwise zero. The general castes in our sample are endowed with the highest levels of schooling and English skills, followed by OBCs, SCs, and STs. To gain greater insight into how rural household members choose their occupations, it is important to categorize these occupations appropriately. The job categorization performed here takes into account differences between urban and rural markets,

[13] . In the agricultural sector, Mano et al. (2011) found that workers in cut-flower farms in Ethiopia helped their friends and relatives obtain jobs at those cut-flower farms.

between manufacturing and service sectors, and between construction work and other jobs. Service sector jobs include government officers, call centre staff, hotel personnel, restaurant employees, shop workers, security guards, and drivers. This categorization offers an understanding of the influence of recent government efforts to boost the manufacturing sector through job generation and through the workfare program NREGA, which provides rural workers with unskilled jobs, mostly in construction.

Occupations are classified into eight categories: (1) self-employed farming, housewife, and inactive worker, (2) agricultural wage work, (3) local construction, (4) local factory, (5) local service and self-employment, (6) migratory construction, (7) migratory factory, and (8) migratory service and self-employment. Since all the sample households engage in farming activity, those who own land and do not engage in any off-farm work fall into category (1), and the landless households who only engage in agricultural wage work fall into category (2).[14] The service and self-employment category is not a perfect proxy for high-return jobs given its heterogeneity, However, the jobs included in this category are mainly well-paid jobs such as government officers, hotel, restaurant, shops, and security guards, which is why the average wages for this

[14]. Occupations, classified into eight categories, are assigned depending on the occupation in which each individual earns the largest amount of income. The questions regarding off-farm jobs and income are listed in Table A1 of the online appendix.

category is much higher than other categories as it will be discussed in this section. Therefore this job classification allows one to gain deeper insights into how households occupational choice.

An implicit assumption here is that women sometimes contribute to farming from their own household directly and sometimes indirectly by engaging in household chores, which helps male members spend more time farming. The other categories are those who engage in occupations in addition to farming. Local job categories include cases where workers commute from home, while migratory job categories include only migrants. Notably, we analysed the cases of individuals who participated in any of these activities over the last 12 months. There were 76 individuals engaged in both local and migratory off-farm activities; they are classified as migrants in this analysis.

Table 2.3 summarizes the basic characteristics of sampled individuals, households, districts, and regularity of jobs by occupational type. Most of those who work locally off-farm engaged in either self-employment/services, or agricultural wage work, while the proportion of factory workers is much higher among migrants. In the four states under this study, major industries in manufacturing sector are agro processing, food processing, textile and leather. An annual survey of industry finds that the 10 states with the highest manufacturing sector (net value-added) accounted for 79

per cent of that of the country in 2011-12, which may explain why 80 per cent of factory workers in the sample are migrants. Years of schooling and English language skills exhibit greater importance in high-return jobs, factory, and self-employment/service. The value of assets for household use is highest for self-employment/service and higher for migratory jobs, presumably due to the need for investment in the start-up costs of self-employment and migration. As for landholdings, the trend may not be immediately clear from the descriptive data, but we expect that households with less land endowments are pushed more to work outside their farms. As expected, the households with more workers had more members migrating. Additionally, drought duration and submergence damage, indicators of crop damage by natural disaster, should affect the coping behaviour of households, but this is not apparent in our descriptive data.

One main focus of this study is the examination of the impact of social group on job choice. It is shown that the proportion of STs and SCs in agricultural wage work and construction is higher than that of others, while the number of STs and SCs in the self-employed farming category is smaller because the limited landholdings of STs and SCs pushes them to work off-farm, yet they are forced to accept unskilled and low-return jobs, which may be an indication of job discrimination. It is also possible, however, that STs and SCs have less access to nonfarm job information. District

characteristics also revealed differences across occupations. Data on daily income show clear contrasts across types of occupations with agricultural wage work being the lowest paid and factory and self-employment/service pay being much higher, particularly outside the home locality. Therefore, we consider these jobs as "high-return" jobs, i.e., key occupations for a pathway out of rural poverty. Furthermore, the last variable in Table 2.3 is the dummy variable for job regularity. Because we do not have information about job regularity, each occupational category is classified as regular if salary payment is monthly or annual; or if salary payment is weekly and the individual worked for more than three months in the previous year. It is clear that agricultural wage work and local construction work are mostly casual, while other forms of employment are regarded as regular.

2.4 Estimation Models and Variables

2.4.1 Determinants of occupational choice

To identify the determinants of individual participation in off-farm activities and migration, we estimated the occupational choice model using the multinomial logit (MNL) estimation:

$$\Pr(z_i = j) = \frac{\exp(X_i\gamma_j + X_h\gamma_j)}{\sum_{j=0,\ldots,7}\exp(X_i\gamma_j + X_h\gamma_j)}\ , \text{j=0,…, 7,} \quad (1)$$

where z_i is an indicator variable denoting the choice of individual *i* in household *h* with respect to *j*th job; X_i is a vector of individual attributes including age, age squared, years of schooling, and dummy variables for English skills, and male, single, and household head dummies; X_h is a vector of household attributes including land size, value of assets, number of workers, distance to market, ratio of irrigated farm land, number of days of most severe drought in the previous year, percentage of crop damaged by submergence and dummy variables for SCs, STs, OBCs, and non-Hindus; and γ_j are vectors of coefficients to be estimated, associated with choice of *j*th job.[15] Market access is defined as the distance to the market town that farmers regularly visit to sell farm products and purchase farm inputs and consumer goods. In this multinomial logit model, "self-employed farming only, housewives, and inactive workers" is a base category. Notably, we only included those 16 years old and above and excluded those still in school even after age 16.

2.4.2 Estimation of monthly and annual income

In the second stage, functions explaining monthly and annual income from off-farm jobs were estimated.[16] It is important to note that income is wage for those engaged in

[15] . The information about English skills is based on self-report, which equals one if the respondent reported having basic skill in English speaking, reading, and writing, and zero otherwise.

[16] . Since some workers receive their wages daily, we estimated the determinant of daily income with the same specification as monthly income and found no qualitative differences with the

wage work and is the value of production minus paid-out cost in case of self-employment.[17] In estimating monthly and annual income, we encounter an issue of sample selection bias because monthly and annual income are estimated only for off-farm jobs. Importantly, the lack of information on farming income is a limitation of this study. To address this selection bias issue, we adopted a two-stage procedure; the first stage being a multinomial logit estimation on occupation choice and the second stage being models of monthly and annual income. Following Lee (1983), the correction term λ is calculated as:

$$\hat{\lambda} = \frac{\emptyset\left[\Phi^{-1}[\widehat{Pr}(Z_i=j)]\right]}{\widehat{Pr}(Z_i=j)}$$

where $\widehat{Pr}(z_i = j)$ is the predicted value of household *i* that chooses the *j*th job and $\emptyset$ [·] and Φ[·] are the density and distribution functions for a standard normal variable.[18] Our second-stage model for monthly income was estimated as:

$$lnW_{ij} = X_i\beta_j + X_h\omega_j + \rho_j\hat{\lambda}_{ij} + \varepsilon_{ij}, \qquad j=1,\ldots,7$$

and, for annual income:

---

monthly income estimation results. The results are in Table A2.

17 . Since the definition of income is different for those in service and those in self-employment, a separate analysis was conducted, dividing the group into two categories, service and self-employment. The results were qualitatively the same as the reported results.

18. This methodology, developed by Lee (1983), is widely used in analysis of occupational choice and wage (for example, Liu, Zhang, and Chong, 2004; Kurosaki and Khan, 2006; Senaratna Sellamuttu et al., 2014).

$$lnY_{ij} = X_i\theta_j + X_h\sigma_j + \tau_j\hat{\lambda}_{ij} + \eta_{ij}, \qquad j=1,\ldots,7$$

where lnW_{ij} is the log of monthly income of individual *i* from *j*th job; lnY_{ij} is log of annual income of individual i from *j*th job; β_j and θ_j are vectors of coefficients of X_i for *j*th job; and ω_j and σ_j are vectors of coefficients of X_h. $\hat{\lambda}_{ij}$ is a correction term to control for selectivity bias, and ε_{ij} and η_{ij} are zero mean error terms. Notably, drought duration and submergence damage are included in the first stage only as the second-stage regressions can be identified if at least one variable in X_h in equation (1) does not affect income directly, but rather indirectly through job choice (Kurosaki and Khan, 2006).

It is important to acknowledge that there is limit as to how much endogeneity and measurement error can be controlled for. Reporting of type of job and earning by respondents are bound to involve measurement errors. Despite the effort to minimize the selection bias in estimating the income for wage earners, there is still a possible bias which occurs in making a decision to migrate. Therefore, the results of the analyses need to be interpreted with care given these limitations.

2.5 Estimation Results

2.5.1 Occupational choice behaviour

Table 2.4 shows the results of the multinomial logit estimation on occupational choice. The coefficients are marginal effects evaluated for the mean values of the explanatory variables. The choices include eight occupational categories as described in the previous section, including the "self-employed farming, housewives, and inactive workers" group as the base category. [19] Additionally, male workers have a much higher probability than female workers of working for all nonfarm jobs. Since gender is one of the most important determinants of occupational choice, separate analyses of occupational choice and monthly and annual income are conducted for male workers only. The results are summarized in Table A2.3 through Table A2.5 in the online appendix.[20] Furthermore, land size shows negative impacts on agricultural wage work, as well as local and migratory factory work, which indicates that these jobs provide more employment opportunities for landless and marginal farmers. The marginal effects of asset value suggest that owning more assets increases the probability of engaging in local self-employment/service. This finding is consistent with our interpretation in the

[19]. Considering the possible correlations between years of schooling and English skills and between land size and asset value, four separate analyses were conducted without including respective variables whose results are presented in Table A2.6 – Table 2.17 of the online appendix.

[20]. Qualitative results of the estimation using only male samples are consistent with results using all samples.

descriptive analysis that an explanation for the above impact of assets on self-employment is the need for investment in set-up costs for self-employment and migration. [21]

Irrigation seems to be an important determinant of occupational choice as well. Its negative effects on agricultural wage work and all migratory work indicate that farmers with more irrigation have less need to work off-farm. Crop damage by submergence and drought are expected to increase farmers' engagement in nonfarm jobs. Indeed, they present positive impacts only on nonfarm jobs and negative impact on agricultural wage work; presumably because after natural disasters, farmers cannot find jobs in agriculture and thus seek non-farm jobs to compensate for the loss. We hypothesized that human capital variables have positive impacts on the choice of service/self-employment. In fact, marginal effects of both years of schooling and English skills are positive and

[21]. Land and asset distribution are largely unequal across social groups, as evident in Table 2.2. Therefore, additional analyses are conducted which include the interaction terms between social groups and land size, and social group and value of asset, whose results are reported in Table A2.18- Table A2.20 of the online appendix. Both land size and asset value show differential impacts on the occupational choice by social groups. The impact of land size indicate that: (1) its negative effect on the choice of local self-employment/service is distinctively larger for STs and non-Hindus than others, (2) its negative effect on the choice of migrant construction is larger for upper-castes than other social groups, (3) its negative effect on the choice of migrant self-employment/service is only present for upper-castes and STs while its effect is positive for Non-Hindus. Asset value increase the probability of choice of off-farm jobs primarily for non-Hindus (agricultural wage work, local construction work, local factory work) and for OBCs (local factory and local self-employment/service).

significant only for service/self-employment and not for other jobs. Our finding that the level of human capital, both English language skills, and years of schooling does not have a significant impact on the choice of factory work suggests that the disadvantaged group may have a better chance of getting jobs in factories than in the service sector.

As for social group, non-Hindus have a higher probability of working in local factories, local self/services, migratory construction, and migratory factory; this finding offers no clear evidence of discrimination against non-Hindus in the labour market. STs and SCs are more likely to work in low-return activities such as agricultural wage work and local construction; while STs have a lower probability of working in migratory factory jobs and SCs have a lower probability of working in migrant factory jobs. This finding is consistent with our second hypothesis. Furthermore, this result could be an indication of job discrimination against the lower castes; however, it is possible that the caste dummies captured other factors that affect job choice such as access to nonfarm job information and geographical remoteness. As discussed earlier, low-caste people tend to rely on personal network to find jobs, and personal network is important in unskilled work. Therefore, low-caste people may be trapped in a vicious circle in which their access to information on nonfarm jobs is limited, which in turn limits their employment opportunities. This same result would be the case for high-return jobs.

Another issue is that their access to institutions such as vocational training centres and recruitment firms is limited. While upper castes use those institutions to find high-return nonfarm jobs, it is difficult for low-caste members to do.

The marginal effect of SC on finding local factory jobs is positive and significant. This finding also suggests that factory jobs are accessible, high-return nonfarm jobs for disadvantaged groups. For STs, living in a remote area has a considerable impact on job choice. Although distance to market was controlled for in the estimation, geographical factors such as long travel time to the nearest city and the unavailability of public transportation may hinder STs from engaging in nonfarm jobs. In our sample, 37 per cent of ST households reside in villages occupied by only STs; therefore, the geographical differences between STs and non-STs can be a source of variation between the two groups regarding the probability of working in the nonfarm sector. Overall, the results of the first-stage analysis regarding occupational choice corroborate the existence of entry barriers. Access to service/self-employment jobs that are high-paying requires more wealth, human capital, and higher social status. Factory jobs, on the other hand, do not require significantly higher levels of wealth or human capital than casual off-farm jobs.

2.5.2. Monthly and annual income

Table 2.5 presents the results of the off-farm monthly income function for each occupational category. Land size and asset value have positive and significant impacts on income from service/self-employment and factory, suggesting the need for investments to initiate lucrative self-employment and migration jobs. Furthermore, years of schooling and English language skills have positive impacts on monthly income from self-employment/service. Combined with the first-stage results, this finding implies that higher human capital increases the probability not only of acquiring high-return jobs, but also of higher monthly income. Notably, social group dummy variables show no clear evidence of wage discrimination against STs, SCs, OBCs, and non-Hindus. This result is consistent with the findings of previous studies (Ito, 2009; Deininger et al., 2013). It also confirms our third hypothesis that there is no significant income difference between SCs/STs and others within the same occupational category. Therefore, even though STs and SCs have a higher probability of engaging in casual and low-return off-farm jobs after controlling for human capital and wealth, they do not receive a lower monthly income than non-STs/SCs workers once they engage in the occupation.

The results of off-farm annual income estimations are shown in Table 2.6. Since annual income is monthly income multiplied by the number of months worked, this analysis allows us to examine what factors increase the duration of off-farm activities. Human capital and wealth indicators show some positive impacts, though not systematically. One important observation that can be made from Table 2.6 is the positive and significant coefficients of OBCs' migratory factory work. In particular, while the probability of choosing this job and its monthly income are not significantly different for OBCs, the annual income is higher because the number of months worked is greater. For SCs, the coefficient is positive and approximately the same size as that of OBCs, although not significant. For STs, there is no significant difference in annual income among castes, although the marginal probability of choosing migratory factory jobs is negative and significant. These results indicate that members of disadvantaged households may be able to work for a long period of time as factory workers because the work is neither seasonal nor casual, unlike agricultural wage work and construction work.

Combining all of the estimation results, it seems reasonable to conclude that factory work potentially gives great benefits to socially and economically disadvantaged rural workers because: (1) the level of required human capital is not as high as others than for

high-return nonfarm jobs since most of the jobs are not skill-intensive; (2) their monthly income (wage rate in the case of factory jobs) is higher than that from other casual nonfarm jobs; and (3) most factory jobs are regular and therefore workers can engage in long-term work. From our data, however, it is clear that there are few factory job opportunities in the locality. Therefore, efforts to lower the costs associated with migration such as improvement of road infrastructure and public transportation services, especially for those who live in remote areas, can increase the probability of finding factory jobs. Another way of supporting disadvantaged workers is to improve their access to information on nonfarm jobs. As observed in Table 2.2, ST and SC workers find nonfarm jobs through personal networks; whereas upper caste workers make more use of institutions such as recruitment firms and vocational training centres. To analyse this empirically, a simple regression of whether migrants find jobs through formal institutions was run on individual and household attributes (results in Table 2.7). The coefficient of STs/SCs (equal to one if a migrant is either ST or SC and equal to zero otherwise) is significant and negative, which confirms that for finding nonfarm jobs, STs and SCs workers rely on personal networks more than upper castes do. Therefore, providing ST and SC workers with better access to formal employment institutions

could help them acquire information (information on high-return jobs) not available within their own networks.

2.6 Conclusion

This study analysed the diversification strategies of rural households' economic activities using large-scale household survey data from rural areas in eastern India. The first issue we addressed was how to include both local and migratory jobs as occupational choices for rural households. In light of this, we included both local and migratory off-farm jobs as employment alternatives for rural households to understand the mechanism underlying diverse occupational choices. The second issue addressed is related to the classification of occupation. We classified occupations into eight categories to enable the verification of differential impacts of wealth, human capital, and social group on occupational choice and monthly income across different occupational categories, taking into account the Indian government's recent efforts to develop the manufacturing sector and to provide the rural work force with unskilled jobs. This study contributes to the literature on occupational diversification towards off-farm jobs by providing evidence that after addressing selection bias, entry barriers to high-return activities limit the job opportunities of people with less wealth, less human capital, and low social status in rural areas.

The findings of this study point to the importance of factory work employment where monthly income is higher than casual nonfarm jobs, but where human capital requirements are not high, indicating that the manufacturing sector provides job opportunities for unskilled workers. However, it was observed that most factory jobs are available in the urban areas, which is not an easy option for many low-caste people. Therefore, such measures such as improvement of roads and public transport are expected to help the disadvantaged groups from remote areas find jobs in factories. Another way to provide the disadvantaged castes with better access to high-return jobs is to disseminate more information about nonfarm jobs. We found that low-caste workers seek jobs primarily through personal networks and generally find low-return jobs; while upper-caste workers generally engaging in higher-return jobs than lower-caste workers rely on recruitment firms and vocational training centres more than the low-caste workers do. These findings suggest that providing low-caste people with information on nonfarm job opportunities not available through their personal networks and giving them better access to employment support institutions could increase their chances of working in the high-return nonfarm sectors.

It should also be noted that almost all migrant workers in factories are men. Female labour participation is very low, compared with that in other developing countries.[22] This finding could largely be the result of persistent social norms observed in India. Nevertheless, women in India should be given more opportunities to engage in regular, salaried jobs such as factory work—a trend observed in some other countries such as Bangladesh (Heath and Mobarak, 2015). Finally, human capital is crucial to job seeking and income earning. Entry barriers faced by impoverished people may be overcome through education. Additionally, better access to factory work, to information on nonfarm jobs, and to higher human capital could improve the living standard of the socially and economically disadvantaged people in rural India, which in turn would lead to poverty reduction.

[22]. India was ranked 11th from the bottom in the world in terms of female labour participation rate (ILO, 2013).

Table 2.1 Number of sample districts, households, and individuals by state (Only those above 16 years old and not in school)

State	Districts	Households	Individuals
Bihar	34	1,491	5,868
Odisha	27	1,784	6,040
Uttar Pradesh	9	1,198	5,390
West Bengal	16	2,321	8,214
Total	86	6,794	25,512

Table 2.2 Composition, proportion of off-farm workers, migrants' job seeking network, years of schooling, and English skill, land size and asset value of workers by social group

	Social group composition		Proportion of off-farm workers by social group (%)		Proportion of migrants who found job through personal network among total migrants by social group (%)[a]	Years of schooling (mean value)	English skill[b] (mean value)	Land size (ha)[c]	Asset value (rupee)[d]
	Number	Percentage (%)	Local off-farm	Migratory off-farm					
Upper caste	6,413	25.1	12.9	5.0	63.3	7.8	0.9	1.0	34,053
OBCs	8,188	32.1	14.2	7.1	75.4	6.3	0.7	0.9	22,356
SCs	4,588	18.0	19.9	4.6	71.2	5.1	0.5	0.5	10,923
STs	3,418	13.4	19.3	3.4	79.3	3.5	0.4	0.4	7,149
Non-Hindu	2,905	11.4	20.9	4.0	84.6	5.4	0.6	0.6	16,895
All	25,512	100.0	16.3	5.3	73.0	6.0	0.7	0.8	20,764

[a] Network includes families, relatives, friends and previous co-workers while the rest are through recruitment firms, contractors, and vocational training center.

[b] English skill evaluation is based on self-report. The value is binary and equals one if one knows basic English.

[c] Land size of owned land, excluding the rented land.

[d] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table 2.3 Individual, household, social group, and district characteristics by occupational category

	Total	Self-employed farming, housewoves, discouraged workers	Self-employed farming + agricultural wage work	Self-employed farming + Local nonfarm work			Self-employed farming + Migratory nonfarm work		
				Construction	Factory	Self-employment /Service	Construction	Factory	Self-employment/ Service
Individual characteristics									
Male (=1)	0.55	0.46	0.63	0.98	0.98	0.94	0.99	0.99	0.97
Age	38.9	39.8	38.9	35.6	34.3	36.2	34.3	31.7	33.3
Single (=1)	0.18	0.17	0.15	0.18	0.28	0.21	0.20	0.27	0.27
Head (=1)	0.27	0.24	0.41	0.54	0.43	0.41	0.45	0.19	0.22
Years of schooling	6.0	5.7	3.8	5.0	7.6	9.4	5.8	8.6	10.4
English skill (=1)	0.65	0.58	0.56	0.43	0.57	1.33	0.43	0.87	1.52
Household characteristics									
Land size (ha)[a]	0.76	0.83	0.20	0.37	0.41	0.82	0.54	0.66	1.00
Landless (=1)	0.06	0.00	0.56	0.09	0.07	0.04	0.09	0.05	0.06
Asset value (rupee)[b]	20,764	20,865	8,166	9,314	14,026	32,176	12,128	22,014	41,414
Number of workers	4.4	4.5	3.8	3.7	4.1	4.4	4.2	5.3	5.3
Drought duration (days)[c]	6.4	6.4	4.8	5.8	9.1	7.0	10.2	7.3	8.8
Submerge damage (%)[d]	7.4	7.3	8.4	7.5	10.5	7.6	8.9	9.4	5.4
Social group									
Upper Caste (=1)	0.25	0.27	0.14	0.14	0.13	0.29	0.13	0.20	0.31
Other Backward Caste (=1)	0.32	0.33	0.19	0.25	0.37	0.33	0.41	0.45	0.42
Schedule Caste(=1)	0.18	0.16	0.31	0.30	0.31	0.17	0.22	0.16	0.13
Schedule Tribe (=1)	0.13	0.13	0.23	0.20	0.10	0.10	0.10	0.09	0.07
Non-Hindu (=1)	0.11	0.12	0.13	0.11	0.10	0.11	0.15	0.10	0.06
Worker's district of origin									
Distance to state capital (km)	218	214	231	224	249	232	230	228	227
Population density (per ha.)	16.6	16.8	16.3	16.5	15.6	16.0	15.8	15.0	15.4
Earnings									
Daily off-farm earnings (rupee)	361.0	.	189	242	370	405	356	423	664
Regularity of the job[e]									
Regular (=1)	0.10	0.00	0.06	0.13	0.84	0.66	0.82	0.77	0.73
Number of observation	25,512	19,238	2,199	666	134	1972	172	570	561

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

[c] Drought duration is the days the most severe drought continuted in the previous year

[d] Submerge damage is the percentage of crop damanged by being submerge in the presious year

[e] Jobs are calssified as regular if (1) the salary payment is monthly or annual, (2) the salary payment if weekly and worked more than 3 months in the previous year, and (3) if they migrate and worked more than 3 months in the previous year.

[f] The income information for self-employed farming is not available.

Table 2.4 Estimation results of the multinomial logit model on occupational choice at the individual level

	Marginal effects on the probability of choosing *j*						
	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment/service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 self-employment/service
Individual characteristics							
Male (=1)	0.0051	0.0855***	0.0148***	0.1239***	0.0257***	0.0813***	0.0503***
	(0.007)	(0.015)	(0.002)	(0.007)	(0.004)	(0.017)	(0.007)
Age	-0.0006	0.0024	0.0010*	0.0123***	0.0005	0.0026**	0.0025***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age-squared	0.0001	-0.0041*	-0.0013**	-0.0151***	-0.0009	-0.0041***	-0.0032***
	(0.000)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Single (=1)	-0.0077	-0.0112*	0.0013	-0.0143*	-0.0036	-0.0092	-0.0063
	(0.010)	(0.005)	(0.002)	(0.007)	(0.004)	(0.007)	(0.008)
Head (=1)	0.0237	-0.0051*	-0.0011	-0.0050	-0.0008	-0.0192***	-0.0155***
	(0.020)	(0.002)	(0.001)	(0.013)	(0.003)	(0.005)	(0.003)
Year of schooling	-0.0036*	-0.0024***	0.0000	0.0059***	-0.0004	-0.0006	0.0015**
	(0.002)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
English skill (=1)	-0.0052	0.0023	-0.0007	0.0160**	-0.0024	0.0072	0.0092**
	(0.020)	(0.005)	(0.001)	(0.005)	(0.003)	(0.005)	(0.003)
Household characteristics							
Land size (ha)[a]	-0.1535**	-0.0127	-0.0035***	0.0087	0.0001	-0.0032*	0.0025
	(0.054)	(0.008)	(0.000)	(0.005)	(0.001)	(0.002)	(0.002)
Log of Asset Value[b]	0.0027	-0.0004	0.0009	0.0128***	0.0006	-0.0002	-0.0006
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Number of workers	-0.0033***	-0.0031**	-0.0005	-0.0022	-0.0003	0.0022***	0.0014***
	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)	(0.000)	(0.000)
Distance to market	0.0584	-0.0714	-0.0022	0.0306	-0.0128	0.0308	0.0210
	(0.110)	(0.047)	(0.013)	(0.154)	(0.023)	(0.034)	(0.033)
Irrigated ratio (=1)	-0.1636***	0.008	(0.001)	0.0106*	-0.0038**	-0.0073**	-0.0107**
	(0.039)	(0.007)	(0.002)	(0.005)	(0.001)	(0.003)	(0.004)
Submerge percent	-0.0000*	0.000	0.000	0.000	0.000	0.0001**	(0.000)
	0.000	(0.000)	0.000	(0.000)	(0.000)	0.000	(0.000)
Drought duration (10 day	-0.0015*	0.0002*	0.0001*	0.0005	0.0002	-0.0002	0.0000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Social Group							
STs (=1)	0.0337**	0.0116*	0.0009	-0.0042	-0.0021	-0.0064***	-0.0050
	(0.012)	(0.005)	(0.002)	(0.005)	(0.002)	(0.002)	(0.003)
SCs (=1)	0.0430***	0.0172***	0.0047***	0.0091	0.0023	-0.0028	-0.0054*
	(0.009)	(0.002)	(0.001)	(0.010)	(0.003)	(0.005)	(0.002)
OBCs(=1)	0.009	0.0099*	0.0033*	(0.000)	0.002	(0.000)	-0.0032***
	(0.008)	(0.005)	(0.001)	(0.004)	(0.002)	(0.001)	(0.001)
Non-Hindu (=1)	0.0323	0.0019	0.0029*	0.0267***	0.0046**	0.0065**	-0.0053
	(0.022)	(0.004)	(0.001)	(0.008)	(0.002)	(0.002)	(0.006)
Observations	25512	25512	25512	25512	25512	25512	25512

Note: Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.
Base occupational category is self-employed farming.
Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers electric fan, and computers.

Table 2.5 Estimation results of off-farm monthly income estimation at the individual level

	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment/service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 self-employment/service
Individual characteristic							
Male (=1)	0.364**	-0.345	-0.090	0.233	-0.037	0.760	-0.079
	(0.113)	(0.439)	(0.493)	(0.230)	(0.588)	(0.655)	(0.467)
Age	0.009	-0.004	0.033	-0.000	0.063	0.049	0.001
	(0.007)	(0.011)	(0.028)	(0.023)	(0.032)	(0.031)	(0.032)
Age-squared	-0.014	0.008	-0.041	0.015	-0.076	-0.067	0.010
	(0.008)	(0.018)	(0.036)	(0.032)	(0.043)	(0.040)	(0.036)
Single (=1)	-0.117	-0.112*	-0.044	-0.037	-0.035	-0.119	-0.049
	(0.087)	(0.041)	(0.281)	(0.033)	(0.135)	(0.051)	(0.067)
Head (=1)	-0.053	0.036	-0.154	-0.132**	-0.114	-0.145	0.033
	(0.044)	(0.059)	(0.131)	(0.040)	(0.119)	(0.135)	(0.199)
Year of schooling	-0.001	0.006	0.014	0.036*	0.011	-0.001	0.052*
	(0.010)	(0.012)	(0.039)	(0.011)	(0.021)	(0.005)	(0.022)
English skill (=1)	0.015	0.028	0.213	0.102**	0.092	0.178	0.218
	(0.117)	(0.083)	(0.218)	(0.031)	(0.123)	(0.086)	(0.118)
Household characteristics							
Land size (ha)[a]	-0.088	0.102	0.275*	0.029	-0.079	-0.020	0.033
	(0.198)	(0.122)	(0.106)	(0.014)	(0.050)	(0.035)	(0.039)
Log of Asset Value[b]	0.006	0.037	-0.038	0.074*	0.045	0.029*	0.028
	(0.005)	(0.017)	(0.057)	(0.025)	(0.053)	(0.010)	(0.026)
Number of workers	-0.020	0.010	-0.071	-0.003	0.032	0.009	0.020
	(0.020)	(0.019)	(0.037)	(0.017)	(0.025)	(0.025)	(0.011)
Distance to market	-1.502	0.318	0.833	-0.118	0.597	0.256	1.958
	(1.141)	(0.294)	(3.969)	(0.859)	(0.585)	(1.132)	(1.374)
Irrigated ratio (=1)	-0.254	-0.103	-0.304	-0.046	0.111	0.017	-0.124
	(0.310)	(0.103)	(0.158)	(0.093)	(0.124)	(0.034)	(0.164)
Social Group							
STs (=1)	0.030	-0.072	-0.363	0.022	-0.059	0.013	0.062
	(0.047)	(0.122)	(0.235)	(0.060)	(0.294)	(0.144)	(0.153)
SCs (=1)	0.019	-0.140	-0.183	-0.162	0.085	-0.003	0.024
	(0.086)	(0.100)	(0.135)	(0.079)	(0.329)	(0.159)	(0.137)
OBCs(=1)	-0.004	-0.114	-0.241	-0.167**	0.014	0.036	-0.045
	(0.055)	(0.084)	(0.236)	(0.047)	(0.132)	(0.102)	(0.048)
Non-Hindu (=1)	-0.051	-0.022	-0.091	-0.017	0.026	0.066	0.014
	(0.043)	(0.066)	(0.169)	(0.086)	(0.248)	(0.253)	(0.025)
Constant	8.116***	9.673***	10.084**	8.032***	7.708***	6.527**	8.363**
	(0.185)	(1.075)	(2.439)	(1.348)	(0.739)	(1.985)	(1.936)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.133	0.106	0.242	0.243	0.108	0.153	0.303

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.
Sate dummies were included to control for the state level heterogeneity
See Table 4 for [a] and [b]

Table 2.6 Estimation results of off-farm annual income estimation at the individual level

	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment/service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 self-employment/service
Individual characteristic							
Male (=1)	0.725**	-0.170	-0.313	0.187	0.400	0.246	0.280
	(0.155)	(0.620)	(0.358)	(0.438)	(1.207)	(0.227)	(0.390)
Age	0.004	0.009	0.052	-0.026	0.216*	0.062*	0.031
	(0.022)	(0.023)	(0.022)	(0.036)	(0.077)	(0.021)	(0.061)
Age-squared	-0.016	-0.021	-0.047	0.049	-0.263	-0.072*	-0.034
	(0.027)	(0.020)	(0.033)	(0.049)	(0.117)	(0.023)	(0.085)
Single (=1)	-0.289	-0.100	0.242	-0.104	0.420	0.181*	-0.015
	(0.141)	(0.092)	(0.181)	(0.103)	(0.355)	(0.067)	(0.149)
Head (=1)	0.019	0.067	0.073	-0.198***	-0.509	-0.009	-0.235
	(0.103)	(0.055)	(0.178)	(0.008)	(0.321)	(0.050)	(0.252)
Year of schooling	-0.021	-0.004	0.073*	0.042	0.023	0.008	0.029
	(0.024)	(0.027)	(0.024)	(0.020)	(0.035)	(0.009)	(0.023)
English skill (=1)	0.439**	-0.062	-0.001	0.167	-0.054	-0.077*	0.309*
	(0.097)	(0.233)	(0.158)	(0.081)	(0.138)	(0.030)	(0.131)
Household characteristics							
Land size (ha)[a]	-0.258	-0.110	0.102	0.038*	-0.034	0.042	0.060
	(0.446)	(0.135)	(0.218)	(0.014)	(0.063)	(0.054)	(0.025)
Log of Asset Value[b]	-0.001	0.051	0.027	0.091	0.015	0.067***	0.030
	(0.028)	(0.039)	(0.079)	(0.049)	(0.071)	(0.008)	(0.022)
Number of workers	-0.002	0.084	0.042	0.009	0.024	-0.020	0.004
	(0.027)	(0.067)	(0.052)	(0.038)	(0.045)	(0.027)	(0.007)
Distance to market	-4.413***	-1.748	-2.217	-1.631	0.474	-3.432**	-0.357
	(0.697)	(1.302)	(4.320)	(1.082)	(1.235)	(0.681)	(2.195)
Irrigated ratio (=1)	0.003	0.182	-0.082	0.033	0.022	0.293***	-0.109
	(0.527)	(0.120)	(0.126)	(0.176)	(0.222)	(0.041)	(0.085)
Social Group							
STs (=1)	0.082	0.271	0.066	0.135*	0.038	0.076	0.084
	(0.209)	(0.120)	(0.229)	(0.057)	(0.507)	(0.233)	(0.110)
SCs (=1)	-0.005	-0.147	-0.231	-0.143	-0.167	0.245	0.078
	(0.248)	(0.226)	(0.375)	(0.104)	(0.426)	(0.159)	(0.177)
OBCs(=1)	-0.031	-0.050	-0.185	-0.148*	-0.107	0.247***	-0.094
	(0.059)	(0.235)	(0.254)	(0.060)	(0.298)	(0.014)	(0.078)
Non-Hindu (=1)	0.035	0.106**	-0.195*	-0.012	-0.434	0.115	-0.191
	(0.083)	(0.026)	(0.082)	(0.207)	(0.282)	(0.107)	(0.246)
Constant	9.481***	10.000***	10.588***	10.531**	7.034**	8.763***	8.832**
	(0.904)	(0.922)	(0.951)	(2.296)	(1.974)	(0.720)	(1.996)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.264	0.060	0.228	0.221	0.258	0.233	0.125

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

See Table 4 for [a] and [b]

Table 2.7. Estimation results of migrants' use of formal institutions in finding jobs

	Use of formal institute
Age	-0.010*
	(0.006)
Age-squared	0.006
	(0.007)
Male (=1)	-0.067
	(0.113)
Single (=1)	-0.072*
	(0.037)
Head (=1)	-0.003
	(0.046)
Year of schooling	-0.022***
	(0.005)
English skill (=1)	-0.024
	(0.041)
Log of Asset Value[b]	-0.037**
	(0.017)
Land size (ha)[a]	-0.015
	(0.010)
Number of workers	0.006
	(0.007)
Distance to market	1.214
	(0.777)
STs/SCs (=1)	-0.098***
	(0.031)
Constant	1.876***
	(0.237)
Observations	1,311
R-squared	0.303

Note: Robust standard errors in brackets.
* significant at 10%; ** significant at 5%; *** significant at 1% level.
See table 4 for other notes.

Table A2.1 Questions in the survey regarding the off-farm activities

Questions reagarding local off-farm jobs

1 Did this person work off-farm in the past 12 months?

About the activity from which the person earned largest income:

2 What non-farm activity was this peson engaged?

3 What was the frequency of the activity?

4 What was amount of earnings (it is for wage work and production minus paid-out cost for self-employment) in rupee?

5 In the last 12 months, how many times did you receive wage in the unit specified

Questions regarding migratory off-farm jobs

1 Has this person migrated (continuously away from home for more than one month searching for jobs or working)?

About the latest migration:

2 In which year, did this person migrate out?

3 In which month, did this person migrate out?

4 How many monthes did it continue?

5 Is this person currently a migrant?

6 What non-farm activity was this peson engaged?

7 What was amount of monthly earnings (it is wage for wage work and production minus paid-out cost for self-employment) in rupee ?

Table A2.2 Estimation results of the multinomial logit model on occupational choice at the individual level (male only)

	Marginal effects on the probability of choosing *j*						
	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment /service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employment /service
Individual characteristics							
Age	-0.0016*	0.0041	0.0016	0.0200***	0.0007	0.0046**	0.0046***
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)
Age-squared	0.0008	-0.0072	-0.0023*	-0.0244***	-0.0015	-0.0073***	-0.0057***
	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.002)	(0.002)
Single (=1)	-0.0252	-0.0203*	0.0014	-0.0292	-0.0068	-0.0160	-0.0109
	(0.016)	(0.008)	(0.004)	(0.015)	(0.007)	(0.013)	(0.012)
Head (=1)	0.0207	-0.0084*	-0.0021	-0.0137	-0.0013	-0.0343***	-0.0288***
	(0.016)	(0.004)	(0.001)	(0.023)	(0.005)	(0.009)	(0.004)
Year of schooling	-0.0045*	-0.0042***	0.0000	0.0095***	-0.0007	-0.0011	0.0025*
	(0.002)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
English skill (=1)	-0.0252	0.0029	-0.0013	0.0217**	-0.0049	0.0125	0.0155***
	(0.018)	(0.009)	(0.002)	(0.007)	(0.006)	(0.010)	(0.005)
Household characteristics							
Land size (ha)[a]	-0.1000	-0.0265*	-0.0067***	0.0076	-0.0008	-0.0076*	0.0029
	(0.057)	(0.012)	(0.001)	(0.010)	(0.002)	(0.003)	(0.003)
Log of Asset Value[b]	0.0010	-0.0001	0.0013	0.0222***	0.0010	-0.0005	-0.0008
	(0.002)	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Number of workers	-0.0057***	-0.0056*	-0.0009	-0.0037	-0.0005	0.0038***	0.0022**
	(0.002)	(0.002)	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Distance to market	0.0530	-0.1145	-0.0021	0.0155	-0.0211	0.0508	0.0437
	(0.097)	(0.082)	(0.026)	(0.278)	(0.039)	(0.056)	(0.059)
Irrigated ratio (=1)	-0.1447**	0.014	(0.002)	0.0193**	-0.0066*	-0.0119**	-0.0179**
	(0.049)	(0.012)	(0.003)	(0.007)	(0.003)	(0.004)	(0.007)
Submerge percent	0.000	0.000	0.000	0.000	0.000	0.0002**	(0.000)
	0.000	(0.000)	0.000	(0.000)	(0.000)	(0.000)	(0.000)
Drought duration (10 days	-0.0012	0.0004*	0.0002*	0.0008	0.0003	-0.0003	0.0000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Social Group							
STs (=1)	0.0364***	0.0086	-0.0024	-0.0124	-0.0071	-0.0114***	-0.0066
	(0.010)	(0.014)	(0.002)	(0.007)	(0.005)	(0.003)	(0.007)
SCs (=1)	0.0493**	0.0176***	0.0045	0.0090	0.0010	-0.0053	-0.0067
	(0.016)	(0.003)	(0.003)	(0.016)	(0.004)	(0.007)	(0.005)
Non-Muslim (=1)	0.0344	-0.0044	0.0018	0.0434***	0.0053	0.0113**	-0.0080
	(0.027)	(0.009)	(0.004)	(0.012)	(0.004)	(0.004)	(0.009)
Observations	14135	14135	14135	14135	14135	14135	14135

Note: Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.
Base occupational category is self-employed farming.
Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, elect and computers.

Table A2.3 Estimation results of off-farm monthly earning estimation at the individual (male only)

	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment /service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employment /service
Individual characteristi							
Age	0.015*	-0.007	0.032	-0.001	0.063	0.050	0.002
	(0.006)	(0.012)	(0.026)	(0.034)	(0.034)	(0.033)	(0.037)
Age-squared	-0.025*	0.013	-0.035	0.018	-0.075	-0.069	0.008
	(0.009)	(0.022)	(0.038)	(0.044)	(0.044)	(0.045)	(0.042)
Single (=1)	-0.188	-0.086	-0.058	-0.029	-0.018	-0.119	-0.072
	(0.102)	(0.044)	(0.253)	(0.048)	(0.102)	(0.059)	(0.111)
Head (=1)	-0.040	0.047	-0.217	-0.146**	-0.086	-0.154	-0.008
	(0.042)	(0.056)	(0.139)	(0.029)	(0.139)	(0.161)	(0.261)
Year of schooling	-0.010	0.010	0.015	0.031	0.012	-0.002	0.054
	(0.014)	(0.015)	(0.036)	(0.017)	(0.023)	(0.005)	(0.028)
English skill (=1)	-0.057	0.025	0.229	0.121**	0.080	0.174	0.236
	(0.150)	(0.077)	(0.218)	(0.034)	(0.119)	(0.092)	(0.114)
Household characteristics							
Land size (ha)[a]	-0.221	0.114	0.248*	0.035*	-0.083	-0.020	0.029
	(0.225)	(0.148)	(0.078)	(0.012)	(0.068)	(0.039)	(0.045)
Log of Asset Value[b]	0.010	0.040	-0.025	0.074	0.042	0.028*	0.036
	(0.005)	(0.020)	(0.053)	(0.035)	(0.049)	(0.010)	(0.026)
Number of workers	-0.025	0.012	-0.078*	-0.006	0.042	0.008	0.020
	(0.027)	(0.022)	(0.026)	(0.018)	(0.025)	(0.026)	(0.015)
Distance to market	-1.108	0.302	0.892	0.033	0.603	0.262	2.056
	(1.060)	(0.240)	(3.954)	(0.997)	(0.599)	(1.144)	(1.455)
Irrigated ratio (=1)	-0.523	-0.116	-0.296	-0.060	0.142	0.013	-0.152
	(0.396)	(0.100)	(0.189)	(0.094)	(0.125)	(0.038)	(0.189)
Social Group							
STs (=1)	0.109	0.002	-0.154	0.123***	-0.065	-0.016	0.114
	(0.070)	(0.085)	(0.131)	(0.018)	(0.246)	(0.089)	(0.164)
SCs (=1)	0.113	-0.081	0.016	-0.086	0.090	-0.035	0.072
	(0.124)	(0.047)	(0.310)	(0.080)	(0.250)	(0.104)	(0.123)
Non-Hindu (=1)	0.021	0.030	0.063	0.053	0.020	0.033	0.043
	(0.060)	(0.071)	(0.275)	(0.061)	(0.172)	(0.201)	(0.060)
Constant	8.028***	9.353***	9.664**	8.292**	7.623***	7.309**	8.028**
	(0.353)	(0.669)	(1.784)	(1.594)	(0.313)	(1.405)	(1.897)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	903	649	129	1,838	112	490	461
R-squared	0.081	0.119	0.229	0.254	0.131	0.163	0.304

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.4 Estimation results of off-farm annual income estimation at the individual level (male only)

	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment/ service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employment/ service
Individual characteri							
Age	0.005	0.011	0.054	-0.036	0.215*	0.060**	0.029
	(0.032)	(0.022)	(0.025)	(0.038)	(0.085)	(0.017)	(0.063)
Age-squared	-0.023	-0.028	-0.045	0.062	-0.259	-0.070**	-0.032
	(0.043)	(0.026)	(0.038)	(0.051)	(0.128)	(0.018)	(0.088)
Single (=1)	-0.436	-0.112	0.266	-0.097	0.426	0.178*	-0.021
	(0.191)	(0.130)	(0.185)	(0.134)	(0.405)	(0.058)	(0.195)
Head (=1)	0.034	0.097	-0.013	-0.203***	-0.485	-0.019	-0.226
	(0.115)	(0.053)	(0.181)	(0.019)	(0.381)	(0.094)	(0.274)
Year of schooling	-0.026	-0.005	0.073**	0.033	0.023	0.004	0.026
	(0.035)	(0.033)	(0.022)	(0.023)	(0.039)	(0.010)	(0.027)
English skill (=1)	0.290	-0.075	0.003	0.190*	-0.060	-0.107	0.315
	(0.198)	(0.217)	(0.145)	(0.076)	(0.140)	(0.050)	(0.146)
Household characteristics							
Land size (ha)[a]	-0.429	-0.167	0.065	0.047**	-0.045	0.031	0.062*
	(0.519)	(0.197)	(0.202)	(0.012)	(0.066)	(0.042)	(0.025)
Log of Asset Value[l]	0.010	0.056	0.043	0.082	0.008	0.068***	0.038
	(0.025)	(0.039)	(0.066)	(0.050)	(0.072)	(0.008)	(0.028)
Number of workers	-0.015	0.081	0.039	0.008	0.034	-0.025	0.004
	(0.040)	(0.073)	(0.045)	(0.041)	(0.047)	(0.031)	(0.009)
Distance to market	-3.439***	-2.011	-2.317	-1.591	0.368	-3.413**	-0.485
	(0.555)	(1.202)	(4.439)	(1.299)	(1.334)	(0.969)	(2.115)
Irrigated ratio (=1)	-0.342	0.215	-0.071	0.023	0.074	0.316***	-0.124
	(0.755)	(0.119)	(0.154)	(0.194)	(0.213)	(0.041)	(0.106)
Social Group							
STs (=1)	0.267	0.335*	0.232	0.258***	0.105	-0.122	0.176
	(0.264)	(0.140)	(0.100)	(0.027)	(0.391)	(0.215)	(0.124)
SCs (=1)	0.132	-0.090	-0.066	-0.079	-0.079	0.046	0.178
	(0.383)	(0.116)	(0.549)	(0.093)	(0.309)	(0.152)	(0.147)
Non-Hindu (=1)	0.179	0.128	-0.091	0.036	-0.362	-0.067	-0.158
	(0.172)	(0.129)	(0.238)	(0.166)	(0.197)	(0.111)	(0.200)
Constant	9.524***	9.498***	9.855***	11.158**	7.316**	9.312***	9.082**
	(1.504)	(0.495)	(1.089)	(1.949)	(1.305)	(0.670)	(1.767)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	903	649	129	1,838	112	490	461
R-squared	0.178	0.071	0.225	0.230	0.258	0.248	0.129

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.5 Estimation results of the multinomial logit model on occupational choice at the individual level (without English skill)

	Marginal effects on the probability of choosing *j*						
	j =1	*j* =2	*j* =3	*j* =4	*j* =5	*j* =6	*j* =7
	agricultural wage work	local construction	local factory	local self-employment/service	migrant construction	migrant factory	migrant self-employment/service
Individual characteristics							
Male (=1)	0.0054	0.0856***	0.0149***	0.1246***	0.0257***	0.0817***	0.0508***
	(0.007)	(0.016)	(0.002)	(0.006)	(0.004)	(0.017)	(0.006)
Age	-0.0006	0.0023	0.0010*	0.0121***	0.0005	0.0026**	0.0025***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age-squared	0.0001	-0.0041*	-0.0013**	-0.0149***	-0.0010	-0.0040***	-0.0031***
	(0.000)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Single (=1)	-0.0084	-0.0114*	0.0012	-0.0138*	-0.0037	-0.0092	-0.0061
	(0.010)	(0.005)	(0.002)	(0.007)	(0.004)	(0.007)	(0.007)
Head (=1)	0.0240	-0.0052*	-0.0010	-0.0051	-0.0007	-0.0194***	-0.0157***
	(0.019)	(0.002)	(0.001)	(0.013)	(0.003)	(0.005)	(0.003)
Year of schooling	-0.0040**	-0.0023***	-0.0001	0.0069***	-0.0005*	-0.0002	0.0021***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Household characteristics							
Land size (ha)[a]	-0.1525**	-0.0131	-0.0037***	0.0087	0.0000	-0.0030	0.0028
	(0.055)	(0.008)	(0.000)	(0.005)	(0.001)	(0.002)	(0.002)
Log of Asset Value[b]	0.0025	-0.0003	0.0008	0.0131***	0.0005	-0.0001	-0.0004
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Number of workers	-0.0032***	-0.0032**	-0.0005	-0.0023	-0.0002	0.0022***	0.0013***
	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)	(0.000)	(0.000)
Distance to market	0.0606	-0.0742	-0.0021	0.0177	-0.0104	0.0241	0.0089
	(0.117)	(0.049)	(0.015)	(0.154)	(0.024)	(0.034)	(0.040)
Irrigated ratio (=1)	-0.1639***	0.008	(0.001)	0.0126*	-0.0040*	-0.0070*	-0.0105*
	(0.038)	(0.007)	(0.002)	(0.006)	(0.002)	(0.003)	(0.005)
Submerge percent	0.000	0.000	0.000	0.000	0.000	0.0001*	(0.000)
	0.000	(0.000)	0.000	(0.000)	(0.000)	0.000	(0.000)
Drought duration (10 day	-0.0015*	0.0002	0.0001*	0.0004	0.0002	-0.0002	-0.0001
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Social Group							
STs (=1)	0.0293*	0.0053	-0.0014	-0.0043	-0.0039	-0.0063***	-0.0029
	(0.013)	(0.008)	(0.001)	(0.003)	(0.003)	(0.002)	(0.004)
SCs (=1)	0.0388***	0.0113***	0.0026	0.0091	0.0006	-0.0026	-0.0035
	(0.011)	(0.002)	(0.001)	(0.009)	(0.002)	(0.004)	(0.003)
Non-Hindu (=1)	0.0292	-0.0034	0.0010	0.0254***	0.0030	0.0064**	-0.0037
	(0.024)	(0.005)	(0.002)	(0.007)	(0.003)	(0.002)	(0.006)
Observations	25512	25512	25512	25512	25512	25512	25512

Note: Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.
Base occupational category is self-employed farming.
Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, and computers.

Table A2.6 Estimation results of off-farm monthly earning estimation at the individual level (without English skill)

	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment /service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employment /service
Individual characteri							
Male (=1)	0.362**	-0.353	-0.040	0.294	-0.053	0.643	-0.026
	(0.110)	(0.415)	(0.339)	(0.198)	(0.664)	(0.627)	(0.472)
Age	0.008	-0.004	0.041	0.004	0.061	0.045	0.003
	(0.008)	(0.011)	(0.021)	(0.024)	(0.040)	(0.031)	(0.032)
Age-squared	-0.014	0.009	-0.051	0.009	-0.073	-0.061	0.008
	(0.010)	(0.018)	(0.029)	(0.031)	(0.056)	(0.041)	(0.035)
Single (=1)	-0.120	-0.095*	-0.002	-0.037	-0.040	-0.103*	-0.056
	(0.094)	(0.038)	(0.230)	(0.031)	(0.093)	(0.040)	(0.056)
Head (=1)	-0.056	0.046	-0.116	-0.134*	-0.120	-0.118	0.013
	(0.048)	(0.058)	(0.145)	(0.043)	(0.147)	(0.130)	(0.194)
Year of schooling	0.000	0.009	0.030	0.047**	0.015	0.009**	0.071**
	(0.014)	(0.008)	(0.028)	(0.013)	(0.021)	(0.002)	(0.019)
Household characteristics							
Land size (ha)[a]	-0.078	0.104	0.244**	0.030	-0.076	-0.007	0.032
	(0.200)	(0.126)	(0.073)	(0.014)	(0.066)	(0.032)	(0.036)
Log of Asset Value	0.006	0.038	-0.027	0.085**	0.051	0.030*	0.033
	(0.004)	(0.017)	(0.050)	(0.027)	(0.055)	(0.012)	(0.026)
Number of worker	-0.019	0.010	-0.053	-0.007	0.031	0.006	0.022
	(0.020)	(0.019)	(0.031)	(0.017)	(0.027)	(0.023)	(0.012)
Distance to market	-1.523	0.300	0.522	-0.151	0.566	0.079	1.210
	(1.069)	(0.310)	(3.501)	(0.854)	(0.758)	(1.047)	(1.151)
Irrigated ratio (=1)	-0.236	-0.102	-0.279	-0.029	0.093	0.041	-0.132
	(0.313)	(0.104)	(0.177)	(0.094)	(0.126)	(0.050)	(0.174)
Social Group							
STs (=1)	0.025	-0.001	-0.179	0.138**	-0.058	-0.001	0.093
	(0.049)	(0.086)	(0.120)	(0.025)	(0.230)	(0.088)	(0.159)
SCs (=1)	0.013	-0.072	0.019	-0.064	0.067	-0.020	0.014
	(0.067)	(0.041)	(0.324)	(0.089)	(0.258)	(0.110)	(0.117)
Non-Hindu (=1)	-0.063	0.035	0.041	0.064	0.020	0.009	0.037
	(0.039)	(0.073)	(0.265)	(0.044)	(0.164)	(0.197)	(0.043)
Constant	8.129***	9.578***	9.345**	7.538***	7.711***	6.867**	8.079**
	(0.181)	(1.022)	(1.703)	(1.266)	(0.755)	(1.879)	(2.010)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,07v4	664	132	1,956	114	495	476
R-squared	0.133	0.127	0.244	0.266	0.128	0.163	0.308

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.7 Estimation results of off-farm annual income estimation at the individual level (without English skill)

	j =1	*j* =2	*j* =3	*j* =4	*j* =5	*j* =6	*j* =7
	agricultural wage work	local construction	local factory	local self-employment /service	migrant construction	migrant factory	migrant self-employment /service
Individual characteristi							
Male (=1)	0.719**	-0.111	-0.333	0.293	0.497	0.225	0.339
	(0.158)	(0.597)	(0.293)	(0.421)	(1.301)	(0.351)	(0.414)
Age	-0.001	0.009	0.053	-0.016	0.220*	0.058**	0.033
	(0.020)	(0.026)	(0.027)	(0.034)	(0.088)	(0.018)	(0.062)
Age-squared	-0.012	-0.022	-0.049	0.037	-0.268	-0.067**	-0.036
	(0.025)	(0.024)	(0.039)	(0.047)	(0.133)	(0.018)	(0.087)
Single (=1)	-0.308	-0.095	0.243	-0.110	0.417	0.185*	-0.021
	(0.162)	(0.095)	(0.196)	(0.104)	(0.411)	(0.061)	(0.131)
Head (=1)	0.016	0.082	0.085	-0.204***	-0.504	-0.006	-0.258
	(0.112)	(0.057)	(0.179)	(0.008)	(0.352)	(0.067)	(0.238)
Year of schooling	-0.005	-0.007	0.075**	0.061*	0.020	-0.002	0.055
	(0.023)	(0.020)	(0.017)	(0.021)	(0.040)	(0.009)	(0.027)
Household characteristics							
Land size (ha)[a]	-0.248	-0.136	0.104	0.035	-0.037	0.034	0.060*
	(0.443)	(0.142)	(0.229)	(0.016)	(0.062)	(0.044)	(0.023)
Log of Asset Value[b]	0.005	0.054	0.031	0.108	0.009	0.068***	0.037
	(0.024)	(0.037)	(0.068)	(0.048)	(0.079)	(0.009)	(0.021)
Number of workers	-0.003	0.084	0.044	0.004	0.024	-0.025	0.006
	(0.032)	(0.065)	(0.048)	(0.035)	(0.042)	(0.030)	(0.008)
Distance to market	-4.922***	-1.871	-2.007	-1.734	0.434	-3.508**	-1.368
	(0.815)	(1.149)	(4.461)	(1.080)	(1.260)	(0.797)	(2.405)
Irrigated ratio (=1)	0.067	0.196	-0.088	0.058	0.042	0.291***	-0.117
	(0.502)	(0.133)	(0.134)	(0.183)	(0.210)	(0.037)	(0.100)
Social Group							
STs (=1)	0.081	0.320	0.205*	0.244***	0.089	-0.112	0.149
	(0.214)	(0.147)	(0.080)	(0.008)	(0.370)	(0.222)	(0.139)
SCs (=1)	-0.006	-0.103	-0.099	-0.042	-0.077	0.055	0.091
	(0.252)	(0.113)	(0.553)	(0.095)	(0.340)	(0.155)	(0.136)
Non-Hindu (=1)	-0.015	0.133	-0.085	0.079	-0.362	-0.053	-0.137
	(0.084)	(0.137)	(0.241)	(0.154)	(0.210)	(0.102)	(0.186)
Constant	9.495***	9.787***	10.344***	9.724**	6.733**	9.164***	8.460**
	(0.925)	(0.838)	(1.287)	(2.132)	(1.964)	(0.864)	(2.210)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.244	0.068	0.232	0.237	0.257	0.248	0.117

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.8 Estimation results of the multinomial logit model on occupational choice at the individual level (without years of school)

	Marginal effects on the probability of choosing *j*						
	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment /service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employment /service
Individual characteristics							
Male (=1)	-0.0001	0.0807***	0.0149***	0.1356***	0.0248***	0.0800***	0.0536***
	(0.006)	(0.015)	(0.001)	(0.006)	(0.005)	(0.017)	(0.008)
Age	-0.0002	0.0022	0.0010*	0.0127***	0.0004	0.0025**	0.0028***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age-squared	0.0000	-0.0037	-0.0013**	-0.0161***	-0.0009	-0.0039***	-0.0036***
	(0.000)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Single (=1)	-0.0117	-0.0129**	0.0012	-0.0104	-0.0040	-0.0095	-0.0050
	(0.010)	(0.004)	(0.002)	(0.009)	(0.004)	(0.007)	(0.008)
Head (=1)	0.0241	-0.0031	-0.0010	-0.0088	-0.0004	-0.0187***	-0.0164***
	(0.020)	(0.003)	(0.001)	(0.014)	(0.002)	(0.005)	(0.002)
English skill (=1)	-0.0248	-0.0105**	-0.0008	0.0439***	-0.0044*	0.0051	0.0160***
	(0.016)	(0.004)	(0.001)	(0.003)	(0.002)	(0.003)	(0.002)
Household characteristics							
Land size (ha)[a]	-0.1563**	-0.0156*	-0.0036***	0.0118*	-0.0004	-0.0034**	0.0035*
	(0.054)	(0.008)	(0.000)	(0.005)	(0.001)	(0.001)	(0.002)
Log of Asset Value[b]	0.0022	-0.0006	0.0008	0.0143***	0.0005	-0.0004	-0.0003
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Number of workers	-0.0036***	-0.0033**	-0.0005	-0.0022	-0.0002	0.0023***	0.0013***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.000)	(0.000)
Distance to market	0.0486	-0.0757	-0.0029	0.0387	-0.0135	0.0302	0.0250
	(0.114)	(0.051)	(0.014)	(0.151)	(0.023)	(0.035)	(0.036)
Irrigated ratio (=1)	-0.1667***	0.008	(0.001)	0.0118**	-0.0041***	-0.0074**	-0.0101**
	(0.038)	(0.007)	(0.002)	(0.005)	(0.001)	(0.003)	(0.004)
Submerge percent	-0.0001***	0.000	0.000	0.000	0.000	0.0001**	(0.000)
	0.000	(0.000)	0.000	(0.000)	(0.000)	0.000	(0.000)
Drought duration (10 day	-0.0015*	0.0001	0.0001*	0.0006	0.0002	-0.0002	0.0000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Social Group							
STs (=1)	0.0357*	0.0095	-0.0015	-0.0140***	-0.0034	-0.0053*	-0.0049
	(0.014)	(0.008)	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)
SCs (=1)	0.0437***	0.0141***	0.0025	0.0022	0.0010	-0.0021	-0.0046
	(0.010)	(0.001)	(0.001)	(0.009)	(0.002)	(0.003)	(0.002)
Non-Hindu (=1)	0.0319	-0.0014	0.0010	0.0205***	0.0034	0.0072**	-0.0043
	(0.024)	(0.006)	(0.002)	(0.006)	(0.002)	(0.003)	(0.006)
Observations	25512	25512	25512	25512	25512	25512	25512

Note: Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.
Base occupational category is self-employed farming.
Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.9 Estimation results of off-farm monthly earning estimation at the individual level (without years of schooling)

	j =1	*j* =2	*j* =3	*j* =4	*j* =5	*j* =6	*j* =7
	agricultural wage work	local construction	local factory	local self-employment /service	migrant construction	migrant factory	migrant self-employment /service
Individual characteristi							
Male (=1)	0.370*	-0.353	-0.349	-0.290	-0.056	0.596	-0.371
	(0.121)	(0.371)	(0.504)	(0.469)	(0.590)	(0.674)	(0.510)
Age	0.008	-0.004	0.023	-0.045	0.062	0.044	0.008
	(0.009)	(0.010)	(0.039)	(0.038)	(0.032)	(0.030)	(0.038)
Age-squared	-0.014	0.008	-0.029	0.072	-0.074	-0.059	0.001
	(0.010)	(0.016)	(0.054)	(0.051)	(0.043)	(0.039)	(0.051)
Single (=1)	-0.120	-0.087	-0.068	0.024	-0.018	-0.099	0.034
	(0.086)	(0.039)	(0.266)	(0.034)	(0.112)	(0.056)	(0.049)
Head (=1)	-0.053	0.044	-0.147	-0.124**	-0.141	-0.110	0.088
	(0.046)	(0.055)	(0.218)	(0.028)	(0.068)	(0.151)	(0.117)
English skill (=1)	0.004	0.067	0.311	0.128	0.146	0.155	0.335
	(0.159)	(0.032)	(0.178)	(0.125)	(0.128)	(0.068)	(0.210)
Household characteristics							
Land size (ha)[a]	-0.111	0.115	0.362**	0.055**	-0.066	-0.006	0.059
	(0.204)	(0.126)	(0.100)	(0.015)	(0.077)	(0.036)	(0.029)
Log of Asset Value[b]	0.006	0.039	-0.040	0.040	0.046	0.029*	0.035
	(0.004)	(0.017)	(0.065)	(0.039)	(0.050)	(0.011)	(0.036)
Number of workers	-0.020	0.012	-0.068*	0.000	0.031	0.004	0.005
	(0.020)	(0.019)	(0.026)	(0.019)	(0.022)	(0.026)	(0.014)
Distance to market	-1.476	0.345	1.320	-0.244	0.619	0.198	2.341
	(1.183)	(0.325)	(4.012)	(0.840)	(0.604)	(1.140)	(1.295)
Irrigated ratio (=1)	-0.283	-0.103	-0.263	-0.088	0.115	0.029	-0.084
	(0.315)	(0.101)	(0.209)	(0.117)	(0.106)	(0.037)	(0.133)
Social Group							
STs (=1)	0.034	-0.021	-0.196	0.120**	-0.087	0.001	0.089
	(0.058)	(0.100)	(0.198)	(0.026)	(0.186)	(0.086)	(0.118)
SCs (=1)	0.026	-0.085	-0.101	-0.104	0.051	-0.022	0.061
	(0.070)	(0.049)	(0.429)	(0.103)	(0.208)	(0.106)	(0.090)
Non-Hindu (=1)	-0.050	0.026	0.029	-0.037	-0.009	0.025	0.040
	(0.027)	(0.072)	(0.310)	(0.075)	(0.146)	(0.207)	(0.099)
Constant	8.090***	9.648***	10.910**	10.788**	7.920***	6.983**	9.349**
	(0.198)	(1.032)	(2.825)	(2.292)	(1.030)	(1.978)	(1.951)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.133	0.125	0.241	0.237	0.126	0.160	0.259

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.10 Estimation results of off-farm annual income estimation at the individual level (without years of schooling)

	j =1	*j* =2	*j* =3	*j* =4	*j* =5	*j* =6	*j* =7
	agricultural wage work	local construction	local factory	local self-employment/service	migrant construction	migrant factory	migrant self-employment/service
Individual characteris							
Male (=1)	0.682**	-0.422	-0.956	-0.510	0.521	0.084	0.204
	(0.126)	(0.814)	(0.591)	(0.737)	(1.213)	(0.437)	(0.159)
Age	0.008	0.001	0.025	-0.084	0.214*	0.055*	0.038
	(0.022)	(0.032)	(0.040)	(0.050)	(0.083)	(0.020)	(0.062)
Age-squared	-0.019	-0.009	-0.014	0.121	-0.260	-0.062*	-0.042
	(0.027)	(0.034)	(0.054)	(0.069)	(0.126)	(0.023)	(0.089)
Single (=1)	-0.298	-0.049	0.139	-0.029	0.416	0.206*	0.027
	(0.146)	(0.104)	(0.303)	(0.101)	(0.406)	(0.066)	(0.147)
Head (=1)	0.045	0.102*	-0.019	-0.187***	-0.578	0.021	-0.226
	(0.118)	(0.041)	(0.129)	(0.015)	(0.274)	(0.095)	(0.201)
English skill (=1)	0.326	-0.042	0.413**	0.181	0.037	-0.102	0.410
	(0.187)	(0.130)	(0.129)	(0.199)	(0.217)	(0.051)	(0.175)
Household characteristics							
Land size (ha)[a]	-0.309	-0.073	0.391	0.067**	-0.024	0.052	0.074**
	(0.458)	(0.223)	(0.183)	(0.017)	(0.077)	(0.039)	(0.020)
Log of Asset Value[b]	-0.002	0.055	0.013	0.043	0.014	0.068***	0.037
	(0.024)	(0.038)	(0.065)	(0.065)	(0.071)	(0.006)	(0.024)
Number of workers	-0.005	0.095	0.022	0.014	0.019	-0.028	-0.002
	(0.030)	(0.075)	(0.055)	(0.043)	(0.039)	(0.032)	(0.010)
Distance to market	-4.554**	-1.601	-0.713	-1.865	0.332	-3.621**	-0.233
	(0.789)	(1.137)	(5.269)	(1.016)	(1.312)	(0.861)	(2.066)
Irrigated ratio (=1)	-0.049	0.174	0.057	-0.026	0.033	0.311***	-0.095
	(0.538)	(0.141)	(0.126)	(0.209)	(0.180)	(0.042)	(0.061)
Social Group							
STs (=1)	0.139	0.290**	0.075	0.228**	0.048	-0.104	0.134
	(0.234)	(0.089)	(0.112)	(0.058)	(0.312)	(0.208)	(0.083)
SCs (=1)	0.036	-0.147	-0.447	-0.090	-0.141	0.059	0.136
	(0.268)	(0.164)	(0.625)	(0.104)	(0.245)	(0.142)	(0.143)
Non-Hindu (=1)	0.041	0.130	-0.165	-0.054	-0.409	-0.088	-0.134
	(0.092)	(0.116)	(0.258)	(0.176)	(0.200)	(0.123)	(0.213)
Constant	9.264***	10.527***	13.772**	14.146**	7.071*	9.528***	9.062**
	(1.023)	(1.659)	(3.158)	(3.252)	(2.244)	(1.100)	(1.562)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.263	0.069	0.173	0.219	0.255	0.250	0.125

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers electric fan, and computers.

Table A2.11 Estimation results of the multinomial logit model on occupational choice at the individual level (without land size)

	Marginal effects on the probability of choosing *j*						
	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment/service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employment/service
Individual characteristics							
Male (=1)	0.0078	0.0862***	0.0152***	0.1242***	0.0260***	0.0828***	0.0502***
	(0.006)	(0.015)	(0.002)	(0.006)	(0.004)	(0.018)	(0.007)
Age	-0.0015	0.0021	0.0009	0.0122***	0.0004	0.0025**	0.0026***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age-squared	0.0006	-0.0038	-0.0013*	-0.0150***	-0.0009	-0.0040***	-0.0032***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Single (=1)	-0.0073	-0.0115*	0.0011	-0.0137*	-0.0037	-0.0095	-0.0060
	(0.010)	(0.005)	(0.002)	(0.007)	(0.004)	(0.007)	(0.008)
Head (=1)	0.0274	-0.0045*	-0.0009	-0.0049	-0.0007	-0.0192***	-0.0154***
	(0.022)	(0.002)	(0.001)	(0.013)	(0.003)	(0.005)	(0.002)
Year of schooling	-0.0055***	-0.0026***	-0.0001	0.0058***	-0.0004	-0.0007	0.0015**
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
English skill (=1)	-0.0078	0.0012	-0.0011	0.0157**	-0.0026	0.0065	0.0094***
	(0.015)	(0.005)	(0.001)	(0.006)	(0.003)	(0.006)	(0.003)
Household characteristics							
Log of Asset Value[b]	-0.0006	-0.0008	0.0005	0.0125***	0.0005	-0.0005	-0.0006
	(0.002)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
Number of workers	-0.0065***	-0.0038***	-0.0007	-0.0023	-0.0003	0.0018***	0.0013***
	(0.002)	(0.001)	(0.001)	(0.003)	(0.000)	(0.000)	(0.000)
Distance to market	0.0051	-0.0800	-0.0080	0.0328	-0.0131	0.0257	0.0222
	(0.175)	(0.043)	(0.012)	(0.157)	(0.022)	(0.038)	(0.033)
Irrigated ratio (=1)	-0.2151***	0.007	(0.002)	0.0143***	-0.0036*	-0.0072***	-0.0094**
	(0.048)	(0.006)	(0.002)	(0.004)	(0.002)	(0.002)	(0.004)
Submerge percent	0.000	0.000	0.0000*	0.000	0.000	0.0001***	(0.000)
	(0.000)	(0.000)	0.000	(0.000)	(0.000)	0.000	(0.000)
Drought duration (10 days	-0.0017*	0.0002*	0.0001*	0.0005	0.0002	-0.0002	0.0000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Social Group							
STs (=1)	0.0397***	0.0067	-0.0010	-0.0046	-0.0039	-0.0053***	-0.0030
	(0.010)	(0.008)	(0.001)	(0.003)	(0.003)	(0.001)	(0.004)
SCs (=1)	0.0517***	0.0130***	0.0032**	0.0090	0.0008	-0.0013	-0.0036
	(0.009)	(0.002)	(0.001)	(0.009)	(0.002)	(0.004)	(0.003)
Non-Hindu (=1)	0.0290	-0.0047	0.0011	0.0254***	0.0030	0.0068***	-0.0035
	(0.025)	(0.005)	(0.002)	(0.007)	(0.003)	(0.002)	(0.006)
Observations	25512	25512	25512	25512	25512	25512	25512

Note: Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Base occupational category is self-employed farming.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.12 Estimation results of off-farm monthly earning estimation at the individual level (without land size)

	j =1	*j* =2	*j* =3	*j* =4	*j* =5	*j* =6	*j* =7
	agricultural wage work	local construction	local factory	local self-employment /service	migrant construction	migrant factory	migrant self-employment /service
Individual characteristic							
Male (=1)	0.416*	-0.215	0.078	0.300	0.168	0.984	0.178
	(0.132)	(0.391)	(0.484)	(0.367)	(0.920)	(0.796)	(0.665)
Age	0.008	0.001	0.049	0.007	0.057	0.053	0.015
	(0.009)	(0.009)	(0.029)	(0.033)	(0.039)	(0.033)	(0.040)
Age-squared	-0.016	0.002	-0.058	0.007	-0.071	-0.076	-0.007
	(0.009)	(0.016)	(0.039)	(0.044)	(0.050)	(0.045)	(0.048)
Single (=1)	-0.110	-0.114*	-0.031	-0.035	-0.088	-0.152	-0.076
	(0.082)	(0.039)	(0.280)	(0.037)	(0.092)	(0.071)	(0.055)
Head (=1)	0.004	0.033	-0.214	-0.137**	-0.116	-0.197	-0.049
	(0.014)	(0.053)	(0.107)	(0.035)	(0.152)	(0.182)	(0.224)
Year of schooling	-0.013	0.005	0.024	0.040*	0.007	-0.003	0.064**
	(0.011)	(0.013)	(0.038)	(0.017)	(0.028)	(0.006)	(0.019)
English skill (=1)	-0.002	0.027	0.167	0.120*	0.063	0.186	0.266
	(0.109)	(0.078)	(0.263)	(0.038)	(0.129)	(0.091)	(0.148)
Household characteristics							
Log of Asset Value[b]	0.005	0.041	-0.020	0.088*	0.053	0.029*	0.027
	(0.005)	(0.019)	(0.054)	(0.030)	(0.060)	(0.010)	(0.027)
Number of workers	-0.029	0.008	-0.070	-0.004	0.029	0.011	0.029
	(0.024)	(0.021)	(0.035)	(0.017)	(0.028)	(0.025)	(0.013)
Distance to market	-1.396	0.220	1.288	-0.059	0.190	0.284	2.219
	(1.041)	(0.329)	(4.295)	(0.900)	(0.661)	(1.143)	(1.460)
Irrigated ratio (=1)	-0.709	-0.079	-0.272	-0.038	0.033	-0.006	-0.156
	(0.508)	(0.090)	(0.167)	(0.105)	(0.182)	(0.032)	(0.160)
Social Group							
STs (=1)	0.107	0.001	-0.199	0.131***	-0.106	-0.028	0.066
	(0.096)	(0.093)	(0.131)	(0.021)	(0.274)	(0.082)	(0.136)
SCs (=1)	0.125	-0.064	0.012	-0.072	0.067	-0.028	0.027
	(0.086)	(0.046)	(0.298)	(0.090)	(0.245)	(0.104)	(0.118)
Non-Hindu (=1)	-0.009	0.039	0.014	0.063	0.028	0.054	0.043
	(0.019)	(0.070)	(0.258)	(0.078)	(0.146)	(0.208)	(0.040)
Constant	7.780***	9.237***	8.996**	7.468**	7.196**	6.055*	7.275*
	(0.183)	(0.945)	(2.332)	(1.981)	(1.431)	(2.248)	(2.640)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.143	0.124	0.223	0.264	0.118	0.165	0.301

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.13 Estimation results of off-farm annual income estimation at the individual level (without land size)

	j =1	*j* =2	*j* =3	*j* =4	*j* =5	*j* =6	*j* =7
	agricultural wage work	local construction	local factory	local self-employment /service	migrant construction	migrant factory	migrant self-employment /service
Individual characteris							
Male (=1)	0.784***	0.779	-0.041	0.329	0.653	0.451	0.474
	(0.130)	(0.411)	(0.433)	(0.529)	(1.447)	(0.460)	(0.567)
Age	-0.001	0.029	0.073**	-0.010	0.212*	0.064**	0.044
	(0.018)	(0.024)	(0.020)	(0.041)	(0.081)	(0.019)	(0.047)
Age-squared	-0.015	-0.060*	-0.075*	0.030	-0.260	-0.077**	-0.047
	(0.023)	(0.020)	(0.031)	(0.056)	(0.124)	(0.021)	(0.063)
Single (=1)	-0.280*	-0.199	0.269	-0.108	0.362	0.154*	-0.032
	(0.100)	(0.094)	(0.167)	(0.110)	(0.354)	(0.064)	(0.172)
Head (=1)	0.101	0.055	0.081	-0.209***	-0.512	-0.058	-0.309
	(0.116)	(0.061)	(0.277)	(0.006)	(0.383)	(0.101)	(0.312)
Year of schooling	-0.037	-0.029	0.078**	0.050	0.019	0.004	0.044
	(0.023)	(0.023)	(0.021)	(0.021)	(0.042)	(0.008)	(0.022)
English skill (=1)	0.411**	-0.074	-0.035	0.201*	-0.077	-0.091	0.355**
	(0.084)	(0.222)	(0.159)	(0.079)	(0.158)	(0.055)	(0.075)
Household characteristics							
Log of Asset Value[b]	-0.005	0.042	0.046	0.112	0.019	0.071***	0.029
	(0.021)	(0.038)	(0.082)	(0.051)	(0.076)	(0.012)	(0.020)
Number of workers	-0.024	0.046	0.040	0.008	0.021	-0.018	0.016
	(0.025)	(0.061)	(0.049)	(0.037)	(0.048)	(0.032)	(0.011)
Distance to market	-4.438***	-2.707	-1.963	-1.598	0.078	-3.509**	-0.008
	(0.717)	(1.160)	(4.269)	(1.129)	(1.386)	(0.906)	(2.382)
Irrigated ratio (=1)	-0.611	0.216	-0.121	0.043	-0.026	0.288**	-0.129
	(0.535)	(0.099)	(0.168)	(0.185)	(0.182)	(0.054)	(0.102)
Social Group							
STs (=1)	0.203	0.371*	0.173*	0.233***	0.068	-0.135	0.122
	(0.218)	(0.149)	(0.065)	(0.009)	(0.380)	(0.210)	(0.135)
SCs (=1)	0.155	0.022	-0.040	-0.051	-0.094	0.043	0.111
	(0.256)	(0.092)	(0.488)	(0.093)	(0.306)	(0.159)	(0.130)
Non-Hindu (=1)	0.080	0.102	-0.058	0.084	-0.352	-0.050	-0.101
	(0.045)	(0.126)	(0.231)	(0.172)	(0.184)	(0.128)	(0.210)
Constant	9.107***	7.765***	8.937***	9.497**	6.368*	8.597***	7.864**
	(0.836)	(0.216)	(1.161)	(2.599)	(2.587)	(1.117)	(1.865)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.270	0.066	0.231	0.242	0.256	0.249	0.114

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.14 Estimation results of the multinomial logit model on occupational choice at the individual level (without value of asset)

	Marginal effects on the probability of choosing *j*						
	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employmen t/service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employmen t/service
Individual characteristics							
Male (=1)	0.0056	0.0857***	0.0148***	0.1223***	0.0258***	0.0814***	0.0502***
	(0.007)	(0.016)	(0.002)	(0.007)	(0.004)	(0.017)	(0.006)
Age	-0.0006	0.0023	0.0010*	0.0126***	0.0005	0.0026**	0.0026***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age-squared	0.0000	-0.0041*	-0.0013**	-0.0152***	-0.0009	-0.0041***	-0.0032***
	(0.000)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Single (=1)	-0.0086	-0.0114*	0.0012	-0.0134*	-0.0036	-0.0091	-0.0062
	(0.010)	(0.005)	(0.002)	(0.006)	(0.004)	(0.007)	(0.008)
Head (=1)	0.0234	-0.0052*	-0.0010	-0.0057	-0.0009	-0.0192***	-0.0155***
	(0.020)	(0.002)	(0.001)	(0.013)	(0.003)	(0.005)	(0.002)
Year of schooling	-0.0037*	-0.0024***	0.0000	0.0065***	-0.0004	-0.0006	0.0015**
	(0.002)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
English skill (=1)	-0.0033	0.0018	-0.0006	0.0201***	-0.0023	0.0071	0.0092***
	(0.018)	(0.005)	(0.001)	(0.005)	(0.003)	(0.005)	(0.003)
Household characteristics							
Land size (ha)[a]	-0.1501**	-0.0133*	-0.0032***	0.0115*	0.0002	-0.0034*	0.0024
	(0.055)	(0.007)	(0.001)	(0.005)	(0.001)	(0.002)	(0.001)
Number of workers	-0.0027***	-0.0032*	-0.0004	-0.0007	-0.0002	0.0022***	0.0013**
	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.001)	(0.000)
Distance to market	0.0665	-0.0729	-0.0014	0.0667	-0.0094	0.0315	0.0199
	(0.117)	(0.045)	(0.014)	(0.153)	(0.022)	(0.036)	(0.037)
Irrigated ratio (=1)	-0.1631***	0.008	(0.001)	0.0148***	-0.0036*	-0.0073**	-0.0108**
	(0.038)	(0.007)	(0.002)	(0.003)	(0.001)	(0.003)	(0.004)
Submerge percent	0.000	0.000	0.000	0.000	0.000	0.0001**	(0.000)
	0.000	(0.000)	0.000	(0.000)	(0.000)	0.000	(0.000)
Drought duration (10 days	-0.0015*	0.0002*	0.0001*	0.0006	0.0002	-0.0002	0.0000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Social Group							
STs (=1)	0.0267*	0.0055	-0.0021	-0.0116***	-0.0046	-0.0060**	-0.0023
	(0.013)	(0.007)	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)
SCs (=1)	0.0382***	0.0113***	0.0023	0.0043	0.0004	-0.0024	-0.0029
	(0.011)	(0.002)	(0.001)	(0.009)	(0.002)	(0.004)	(0.003)
Non-Hindu (=1)	0.0243	-0.0033	0.0001	0.0071	0.0022	0.0067*	-0.0027
	(0.022)	(0.005)	(0.002)	(0.016)	(0.003)	(0.003)	(0.008)
Observations	25512	25512	25512	25512	25512	25512	25512

Note: Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.
Base occupational category is self-employed farming.
Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, and computers.

Table A2.15 Estimation results of off-farm monthly earning estimation at the individual level (without value of asset)

	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment/service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employme
Individual characterist							
Male (=1)	0.362*	-0.383	-0.070	0.072	0.039	0.587	-0.469
	(0.114)	(0.309)	(0.435)	(0.248)	(0.639)	(0.588)	(0.517)
Age	0.009	-0.006	0.037	-0.009	0.065	0.048	-0.018
	(0.007)	(0.007)	(0.025)	(0.030)	(0.031)	(0.028)	(0.034)
Age-squared	-0.014	0.012	-0.047	0.028	-0.078	-0.063	0.035
	(0.008)	(0.012)	(0.034)	(0.039)	(0.043)	(0.037)	(0.043)
Single (=1)	-0.119	-0.088	-0.033	-0.015	-0.030	-0.093	-0.005
	(0.086)	(0.054)	(0.261)	(0.029)	(0.116)	(0.043)	(0.039)
Head (=1)	-0.055	0.049	-0.129	-0.128*	-0.112	-0.109	0.147
	(0.041)	(0.062)	(0.157)	(0.043)	(0.133)	(0.115)	(0.168)
Year of schooling	-0.000	0.011	0.016	0.033	0.010	0.001	0.042*
	(0.010)	(0.009)	(0.036)	(0.014)	(0.022)	(0.004)	(0.017)
English skill (=1)	0.018	0.030	0.199	0.111	0.117	0.162	0.170
	(0.117)	(0.078)	(0.222)	(0.051)	(0.131)	(0.092)	(0.153)
Household characteristics							
Land size (ha)[a]	-0.080	0.143	0.257*	0.050**	-0.094	0.005	0.043
	(0.191)	(0.102)	(0.093)	(0.012)	(0.081)	(0.032)	(0.031)
Number of workers	-0.018	0.020	-0.069**	0.002	0.034	0.009	0.015
	(0.019)	(0.016)	(0.020)	(0.014)	(0.024)	(0.024)	(0.011)
Distance to market	-1.489	0.529	1.103	-0.392	0.873	0.290	1.962
	(1.123)	(0.272)	(4.008)	(0.706)	(0.799)	(1.126)	(1.357)
Irrigated ratio (=1)	-0.242	-0.093	-0.326	-0.039	0.121	0.056	-0.079
	(0.308)	(0.090)	(0.169)	(0.111)	(0.112)	(0.031)	(0.161)
Social Group							
STs (=1)	0.021	-0.036	-0.170	0.086**	-0.069	-0.023	0.095
	(0.053)	(0.085)	(0.157)	(0.015)	(0.252)	(0.105)	(0.132)
SCs (=1)	0.013	-0.089*	0.005	-0.101	0.090	-0.031	0.066
	(0.058)	(0.032)	(0.297)	(0.092)	(0.276)	(0.110)	(0.112)
Non-Hindu (=1)	-0.066**	0.034	0.074	0.016	0.013	0.006	-0.000
	(0.020)	(0.072)	(0.272)	(0.022)	(0.200)	(0.187)	(0.057)
Constant	8.155***	9.927***	9.375**	9.223***	7.793***	7.083**	10.031**
	(0.120)	(0.676)	(1.903)	(1.398)	(0.691)	(1.685)	(2.072)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.133	0.106	0.242	0.243	0.108	0.153	0.303

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, co electric fan, and computers.

Table A2.16 Estimation results of off-farm annual income estimation at the individual level (without value of asset)

	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment/service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employme
Individual characterist							
Male (=1)	0.715**	-0.231	-0.350	0.021	0.460	-0.158	0.036
	(0.151)	(0.663)	(0.297)	(0.477)	(1.211)	(0.337)	(0.385)
Age	0.004	0.005	0.055*	-0.033	0.216*	0.058**	0.020
	(0.022)	(0.025)	(0.019)	(0.042)	(0.085)	(0.016)	(0.070)
Age-squared	-0.017	-0.015	-0.051	0.058	-0.261	-0.062**	-0.018
	(0.027)	(0.022)	(0.029)	(0.056)	(0.131)	(0.018)	(0.099)
Single (=1)	-0.276	-0.072	0.242	-0.082	0.414	0.241**	0.013
	(0.140)	(0.113)	(0.199)	(0.099)	(0.370)	(0.071)	(0.125)
Head (=1)	0.024	0.095	0.093	-0.197***	-0.512	0.071	-0.161
	(0.113)	(0.053)	(0.179)	(0.005)	(0.352)	(0.066)	(0.188)
Year of schooling	-0.019	0.004	0.076**	0.040	0.022	0.009	0.025
	(0.024)	(0.030)	(0.024)	(0.023)	(0.036)	(0.009)	(0.022)
English skill (=1)	0.435**	-0.060	-0.004	0.188*	-0.038	-0.135**	0.291
	(0.096)	(0.219)	(0.151)	(0.076)	(0.160)	(0.042)	(0.178)
Household characteristics							
Land size (ha)[a]	-0.243	-0.068	0.088	0.059**	-0.048	0.094	0.067**
	(0.439)	(0.162)	(0.217)	(0.011)	(0.067)	(0.055)	(0.016)
Number of workers	-0.003	0.100	0.050	0.015	0.023	-0.023	0.003
	(0.023)	(0.072)	(0.053)	(0.033)	(0.043)	(0.031)	(0.006)
Distance to market	-4.486***	-1.507	-2.245	-2.029	0.485	-3.435**	-0.283
	(0.693)	(0.995)	(4.565)	(0.965)	(0.799)	(0.908)	(2.097)
Irrigated ratio (=1)	0.019	0.208	-0.061	0.041	0.037	0.386***	-0.085
	(0.534)	(0.124)	(0.181)	(0.214)	(0.140)	(0.057)	(0.090)
Social Group							
STs (=1)	0.092	0.262	0.215**	0.176***	0.115	-0.135	0.134
	(0.179)	(0.131)	(0.065)	(0.024)	(0.384)	(0.246)	(0.083)
SCs (=1)	0.001	-0.136	-0.091	-0.089	-0.088	0.055	0.140
	(0.239)	(0.124)	(0.521)	(0.102)	(0.315)	(0.150)	(0.128)
Non-Hindu (=1)	0.016	0.128	-0.079	0.023	-0.360	-0.141	-0.187
	(0.059)	(0.123)	(0.239)	(0.094)	(0.221)	(0.083)	(0.153)
Constant	9.439***	10.423***	10.427***	11.813**	6.961**	10.263***	9.904**
	(0.772)	(1.169)	(1.083)	(2.148)	(1.662)	(0.929)	(2.457)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.264	0.060	0.228	0.221	0.258	0.233	0.125

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% leve
Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, c electric fan, and computers.

Table A2.17 Estimation results of the multinomial logit model on occupational choice at the individual level (with interaction terms of social group dummies and land size, and social group and log of asset value)

	Relative risk ratio of choosing *j*						
	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment/ service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employment/service
Individual characteristics							
Male	2.550***	101.370***	80.360***	20.137***	187.385***	202.235***	48.478***
	(0.389)	(67.858)	(31.197)	(2.905)	(130.698)	(174.245)	(24.490)
Age	1.031**	1.171*	1.315***	1.274***	1.153*	1.227***	1.222***
	(0.014)	(0.095)	(0.134)	(0.037)	(0.096)	(0.055)	(0.048)
Age-squared	0.946***	0.779***	0.688***	0.736***	0.787**	0.739***	0.769***
	(0.012)	(0.074)	(0.078)	(0.026)	(0.074)	(0.033)	(0.041)
Single (=1)	0.765	0.538**	1.064	0.703**	0.488	0.538	0.621
	(0.143)	(0.145)	(0.490)	(0.124)	(0.271)	(0.240)	(0.288)
Head (=1)	1.338	0.759**	0.714	0.816	0.787	0.347***	0.412***
	(0.407)	(0.098)	(0.161)	(0.169)	(0.315)	(0.085)	(0.033)
Year of schooling	0.945**	0.914***	1.011	1.101***	0.942	0.988	1.084**
	(0.026)	(0.013)	(0.025)	(0.017)	(0.053)	(0.030)	(0.037)
English skill (=1)	0.980	1.170	0.959	1.355**	0.788	1.596	1.742***
	(0.314)	(0.285)	(0.236)	(0.184)	(0.409)	(0.509)	(0.354)
Household characteristics							
Land size (ha)[a]	0.082	0.399*	0.357**	0.936**	0.238***	0.648***	0.814***
	(0.167)	(0.213)	(0.185)	(0.030)	(0.059)	(0.057)	(0.056)
Land size (ha) * Caste group							
STs (=1)	0.596	0.815	1.310	0.538**	3.048**	0.613	0.704***
	(1.267)	(0.294)	(0.698)	(0.167)	(1.688)	(0.214)	(0.085)
SCs (=1)	0.859	1.608	0.853	1.302	2.122	1.193	2.158
	(1.802)	(1.152)	(0.943)	(0.332)	(2.008)	(0.315)	(1.188)
OBCs(=1)	0.645	0.994	0.908	0.912**	2.874***	0.984	1.071
	(0.566)	(0.403)	(0.356)	(0.039)	(0.616)	(0.160)	(0.111)
Non-Muslim (=1)	3.318	0.339***	0.110	0.525***	3.428***	1.217	1.598***
	(7.120)	(0.089)	(0.273)	(0.097)	(1.428)	(0.270)	(0.273)
Log of Asset Value[b]	1.074	0.894	0.992	1.242***	1.073	0.993	1.076
	(0.055)	(0.075)	(0.168)	(0.031)	(0.107)	(0.056)	(0.076)
Log of Asset Value * Caste group							
STs (=1)	0.888	0.930	1.313	0.926	1.081	1.053	1.062
	(0.101)	(0.117)	(0.290)	(0.063)	(0.291)	(0.074)	(0.092)
SCs (=1)	0.896***	1.091	1.411	0.864**	1.239	0.976	0.953
	(0.037)	(0.115)	(0.319)	(0.058)	(0.284)	(0.097)	(0.102)
OBCs(=1)	1.039	1.054	1.072*	1.018***	0.986	1.011	0.983
	(0.037)	(0.039)	(0.038)	(0.006)	(0.052)	(0.014)	(0.011)
Non-Muslim (=1)	1.128**	1.488***	1.774*	1.098	1.090	1.118	0.898
	(0.054)	(0.033)	(0.595)	(0.076)	(0.140)	(0.077)	(0.084)
Number of workers	0.940***	0.869**	0.889	0.963	0.948	1.103***	1.060***
	(0.003)	(0.049)	(0.102)	(0.047)	(0.064)	(0.021)	(0.003)
Distance to market	1.380	0.023	0.593	1.586	0.115	3.867	3.982
	(2.272)	(0.057)	(1.918)	(3.827)	(0.456)	(7.316)	(7.850)
Irrigated ratio (=1)	0.066***	0.851	0.493	0.867	0.352***	0.495***	0.441***
	(0.046)	(0.376)	(0.221)	(0.134)	(0.085)	(0.119)	(0.130)
Submerge percent	1.000	1.003	1.008	1.001	1.006	1.006**	0.998
	(0.001)	(0.003)	(0.005)	(0.003)	(0.009)	(0.003)	(0.003)
Drought duration (10 days)	0.975**	1.005	1.025*	1.006	1.023	0.992	0.998
	(0.010)	(0.004)	(0.014)	(0.009)	(0.024)	(0.009)	(0.011)
Social Group							
STs (=1)	4.149***	0.659	0.058*	3.070***	0.087*	0.839	0.811
	(1.252)	(0.340)	(0.086)	(1.023)	(0.119)	(0.726)	(0.810)
SCs (=1)	6.782***	4.740	0.221	3.725**	0.503	0.845	0.629
	(4.431)	(4.986)	(0.446)	(2.299)	(0.918)	(0.385)	(0.511)
Non-Muslim (=1)	0.565*	0.078***	0.025	1.216	0.644	0.630	1.404
	(0.170)	(0.003)	(0.063)	(0.832)	(0.386)	(0.404)	(1.036)
Observations	25,512	25,512	25,512	25,512	25,512	25,512	25,512

Note: Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Base occupational category is self-employed farming.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, a

Table A2.18 Estimation results of off-farm monthly earning estimation at the individual level (with interaction terms of social group dummies and land size, and social group and log of asset value)

	j =1	*j* =2	*j* =3	*j* =4	*j* =5	*j* =6	*j* =7
	agricultural wage work	local construction	local factory	local self-employment/ service	migrant construction	migrant factory	migrant self-employment/ service
Individual characteristic							
Male	0.354*	-0.503	-0.115	0.236*	-0.167	0.686	-0.130
	(0.122)	(0.394)	(0.544)	(0.100)	(0.409)	(0.601)	(0.484)
Age	0.008	-0.010	0.026	-0.002	0.045	0.048	0.001
	(0.007)	(0.009)	(0.030)	(0.012)	(0.024)	(0.029)	(0.034)
Age-squared	-0.014	0.018	-0.035	0.016	-0.047	-0.065	0.009
	(0.007)	(0.015)	(0.034)	(0.015)	(0.034)	(0.039)	(0.041)
Single (=1)	-0.115	-0.077	-0.096	-0.038	0.083	-0.107*	-0.027
	(0.082)	(0.043)	(0.325)	(0.024)	(0.122)	(0.043)	(0.080)
Head (=1)	-0.057	0.054	-0.139	-0.116	-0.073	-0.134	0.070
	(0.026)	(0.070)	(0.142)	(0.050)	(0.113)	(0.124)	(0.184)
Year of schooling	-0.000	0.009	0.006	0.036***	0.007	0.000	0.042
	(0.009)	(0.009)	(0.037)	(0.005)	(0.014)	(0.006)	(0.020)
English skill (=1)	0.019	0.015	0.227	0.098**	0.135	0.167	0.199
	(0.105)	(0.070)	(0.231)	(0.022)	(0.073)	(0.089)	(0.134)
Household characteristics							
Land size (ha)[a]	-0.003	0.225*	0.144**	0.028	0.598	-0.016	0.050
	(0.191)	(0.085)	(0.043)	(0.019)	(0.689)	(0.050)	(0.066)
Land size (ha) * Caste group							
STs (=1)	-0.032	-0.195	-0.016	-0.036	-0.560	0.033	0.100
	(0.183)	(0.086)	(0.072)	(0.177)	(0.736)	(0.117)	(0.077)
SCs (=1)	-0.270	-0.215*	0.702	0.070	-0.462	0.284***	0.117
	(0.136)	(0.075)	(0.464)	(0.128)	(0.561)	(0.048)	(0.380)
OBCs(=1)	-0.196	-0.096	0.260	-0.032*	-0.719	-0.033	-0.035
	(0.096)	(0.056)	(0.192)	(0.012)	(0.740)	(0.086)	(0.053)
Non-Muslim (=1)	-0.016	-0.254	0.997	0.083	-1.214	0.139	0.030
	(0.072)	(0.221)	(0.737)	(0.085)	(0.747)	(0.279)	(0.091)
Log of Asset Value[b]	0.029	0.099	0.011	0.100***	0.107*	0.021	0.059
	(0.015)	(0.058)	(0.055)	(0.015)	(0.043)	(0.017)	(0.038)
Log of Asset Value * Caste group							
STs (=1)	-0.023	-0.040	0.085	-0.060	-0.099**	-0.008	0.036
	(0.027)	(0.058)	(0.097)	(0.036)	(0.019)	(0.020)	(0.041)
SCs (=1)	-0.033	-0.079	-0.096	-0.051	-0.017	0.005	-0.032
	(0.017)	(0.055)	(0.258)	(0.024)	(0.054)	(0.028)	(0.036)
OBCs(=1)	0.011	-0.010	-0.046	-0.013**	0.037	0.005	-0.000
	(0.005)	(0.008)	(0.034)	(0.004)	(0.054)	(0.015)	(0.008)
Non-Muslim (=1)	-0.000	-0.079	-0.229**	-0.018	-0.160**	0.018	-0.080
	(0.029)	(0.157)	(0.059)	(0.015)	(0.041)	(0.036)	(0.069)
Number of workers	-0.020	0.017	-0.056	-0.004	0.028	0.009	0.011
	(0.019)	(0.017)	(0.049)	(0.016)	(0.035)	(0.023)	(0.007)
Distance to market	-1.367	0.642	1.984	-0.135	1.199	0.206	2.045
	(1.072)	(0.356)	(3.698)	(0.890)	(0.625)	(1.037)	(1.424)
Irrigated ratio (=1)	-0.216	-0.110	-0.280	-0.044	0.170	0.020	-0.126
	(0.371)	(0.086)	(0.155)	(0.083)	(0.171)	(0.035)	(0.186)
Social Group							
STs (=1)	0.387	0.660	0.157	0.455	0.328	-0.155	0.336
	(0.226)	(0.492)	(2.244)	(0.326)	(0.886)	(0.120)	(0.354)
SCs (=1)	0.227	0.253	-0.980	0.418	1.105	0.043	-0.315
	(0.239)	(0.467)	(0.907)	(0.396)	(0.531)	(0.091)	(0.408)
Non-Hindu (=1)	-0.031	0.705	1.679**	0.114*	1.946**	-0.186	0.576
	(0.235)	(1.299)	(0.449)	(0.044)	(0.481)	(0.272)	(0.502)
Constant	7.927***	9.537***	10.249**	7.817***	7.554***	6.735**	8.028**
	(0.201)	(0.737)	(2.131)	(0.481)	(0.479)	(1.742)	(1.897)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	461
R-squared	0.149	0.152	0.288	0.276	0.230	0.170	0.304

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electri

Table A2.19 Estimation results of off-farm annual income estimation at the individual level (raction terms of social group dummies and land size, and social group and log of asset value)

	j =1 agricultural wage work	*j* =2 local construction	*j* =3 local factory	*j* =4 local self-employment/service	*j* =5 migrant construction	*j* =6 migrant factory	*j* =7 migrant self-employment/service
Individual characteristi							
Male	0.702**	-0.374	-0.413	0.161	0.405	0.271	0.094
	(0.142)	(0.615)	(0.685)	(0.338)	(1.033)	(0.279)	(0.318)
Age	0.005	-0.005	0.056	-0.027	0.204*	0.064***	0.025
	(0.022)	(0.027)	(0.051)	(0.025)	(0.084)	(0.011)	(0.065)
Age-squared	-0.017	-0.002	-0.055	0.049	-0.244	-0.076***	-0.026
	(0.026)	(0.023)	(0.071)	(0.035)	(0.127)	(0.007)	(0.090)
Single (=1)	-0.271	-0.078	0.243	-0.100	0.573	0.179**	0.003
	(0.132)	(0.139)	(0.142)	(0.089)	(0.418)	(0.045)	(0.147)
Head (=1)	0.010	0.095**	0.083	-0.187***	-0.445	-0.012	-0.148
	(0.096)	(0.024)	(0.172)	(0.014)	(0.320)	(0.068)	(0.209)
Year of schooling	-0.018	0.002	0.066**	0.040*	0.023	0.007	0.019
	(0.021)	(0.021)	(0.014)	(0.016)	(0.023)	(0.010)	(0.021)
English skill (=1)	0.433**	-0.071	-0.017	0.173*	-0.048	-0.075	0.257
	(0.086)	(0.216)	(0.194)	(0.062)	(0.124)	(0.050)	(0.154)
Household characteristics							
Land size (ha)[a]	-0.192	0.224	0.543	0.036	-0.380	0.047	0.081
	(0.430)	(0.172)	(0.310)	(0.021)	(0.861)	(0.034)	(0.043)
	0.182	-0.670	-0.883***	0.044	0.600	0.032	0.221
	(0.198)	(0.411)	(0.149)	(0.200)	(0.982)	(0.183)	(0.180)
	0.051	-0.448	0.035	0.163	-0.175	-0.027	0.102
	(0.302)	(0.256)	(0.496)	(0.150)	(0.583)	(0.253)	(0.236)
	-0.093	-0.193	-0.180	-0.069**	0.438	-0.110**	-0.028
	(0.153)	(0.233)	(0.241)	(0.019)	(0.887)	(0.032)	(0.040)
	-0.013	-0.904	0.590*	0.111	-0.288	0.244	-0.138**
	(0.239)	(0.515)	(0.213)	(0.167)	(0.955)	(0.168)	(0.030)
Log of Asset Value[b]	0.001	0.096	0.007	0.134*	0.257	0.065**	0.071
	(0.042)	(0.086)	(0.048)	(0.051)	(0.174)	(0.019)	(0.042)
	-0.000	-0.036	0.233	-0.099	-0.347	-0.007	-0.115
	(0.015)	(0.034)	(0.228)	(0.057)	(0.195)	(0.042)	(0.102)
	-0.017	-0.038	-0.313	-0.078	-0.269	0.038	-0.062
	(0.045)	(0.075)	(0.193)	(0.050)	(0.289)	(0.045)	(0.041)
	0.001	-0.003	-0.020	-0.007	-0.046	0.035***	-0.006
	(0.014)	(0.008)	(0.012)	(0.007)	(0.071)	(0.004)	(0.006)
	0.005	-0.131	-0.180	-0.115	-0.352*	-0.046	-0.131
	(0.011)	(0.154)	(0.110)	(0.055)	(0.141)	(0.039)	(0.061)
Number of workers	-0.009	0.103	0.065	0.007	0.020	-0.017	-0.007
	(0.030)	(0.064)	(0.068)	(0.037)	(0.035)	(0.027)	(0.010)
Distance to market	-4.579**	-1.399	-1.399	-1.693	0.419	-3.122*	-0.468
	(0.862)	(1.188)	(5.056)	(1.069)	(1.550)	(1.054)	(2.156)
Irrigated ratio (=1)	0.063	0.188	0.032	0.032	0.037	0.302**	-0.073
	(0.474)	(0.085)	(0.138)	(0.172)	(0.310)	(0.056)	(0.080)
Social Group							
STs (=1)	0.181	0.739	2.658	0.784	2.502	-0.220	0.640
	(0.485)	(0.713)	(1.674)	(0.502)	(2.349)	(0.441)	(0.369)
SCs (=1)	-0.087	0.349	-2.008	0.773	2.473	0.290	1.037
	(0.325)	(0.243)	(2.418)	(0.650)	(1.727)	(0.515)	(0.998)
Non-Hindu (=1)	-0.004	1.483	1.289	0.988	2.747	0.308	0.942
	(0.165)	(1.252)	(0.855)	(0.615)	(1.382)	(0.295)	(0.685)
Constant	9.511***	10.116***	11.089**	10.200***	5.213*	8.649***	9.228**
	(0.954)	(0.442)	(1.983)	(1.721)	(1.758)	(0.525)	(1.970)
Correction Term	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,074	664	132	1,956	114	495	476
R-squared	0.271	0.092	0.304	0.252	0.349	0.269	0.141

Note: Robust standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% level.

Sate dummies were included to control for the state level heterogeneity

[a] Land size of owned land, excluding the rented land.

[b] Assets include: bicycle, radio, TV, DVD players, mobile phones, two wheelers, four wheelers, refrigrators, coolers, electric fan, a

Chapter 3

Dynamics of school progression in Andhra Pradesh, India:

The role of gender and job opportunities in a transforming economy

3.1 Introduction

It has become a global understanding that education is a major catalyst for human and economic development. One of the target of the Millennium Development Goals on education was to "ensure that by 2015, children everywhere boys and girls alike, will be able to complete a full course of primary education." In India, access to basic education has considerably improved in recent years.[23] Consequently, India's new target in education is to ensure that children graduate from primary school and transit to secondary school.

The large body of research on schooling in India inquiries into the determinants of schooling many of which find that individual, household, and community level characteristics such as household wealth, parental education level, gender of children, availability of schools, school quality, and work opportunities have significant association with school participation. However, these studies focus on enrollment or

[23] The gross enrollment rate at primary school in 2013-14 was 99.3 %, up from 83.8 % in 1990-91 (Government of India, 2014b)

drop-out (Singh 1992; Foster and Rosenzweig, 1996; Jacoby and Skoufias, 1997; Dostie and Jayaraman, 2006; Kajisa and Palanichamy 2010), or the highest grade completed (Birdsall, 1985; Psacharopoulos and Patrinos, 1997; Dreze and Kingdon, 2001). While these studies offer useful insights into what influence children's schooling, their results cannot be used to design measures to reduce high drop-out rates, because it is not clear at what grade and under what conditions drop-out frequently occurs. Similarly, some studies on school progression in India analyze single stage transition such as transition from primary to secondary school, which is not sufficient to understand how schooling decision differ by educational stage (Mare 1980; Sabates, Hossnain, and Lewin 2013; Siddhu 2011).

Since education attainment is the result of a series of complex schooling decisions made over time at the household level, it is desirable to analyze the dynamic aspects of school progression. In other words, it is important to analyze the entire history of the child's education and the changing impact of the household characteristics (Sawada and Lokshin, 2009). Such studies, however, are in general not feasible due to the limited availability of cohort data.

This chapter attempts to fill a gap in the existing literature on schooling by analyzing the decision making process of school progression in a dynamic framework. It adopts sequential logit model, making use of unique panel data which follows the

same age cohort for 11 years during the period 2003-2014 in Andhra Pradesh state in India. We analyze school progression at seven educational stages from lower primary school to upper secondary school. The state of Andhra Pradesh has a diverse geography and diverse economic structure. The proportion of gross state domestic product of non-agricultural sector was 81 per cent in 2013-14 (Government of India, 2016), which is likely to be a result of state's strong initiatives to develop industry and service sectors. The use of data from cohort study in Andhra Pradesh will allow us to examine the differential effects of various factors such as job opportunities on schooling at various stages of schooling.

The remaining sections of this article are organized as follows. Section 2 explains the nature of the dataset used in this study, and section 3 summarizes the key features of the educational system, policies and schooling attainment status in India and in the state of Andhra Pradesh. Section 4 reviews the relevant literature and proposes testable hypotheses, and sections 5 and 6 present the descriptive analysis and empirical methodology, respectively. Analytical results are summarized in section 7, and section 8 draws conclusions and policy implications from the results of the analyses.

3.2 Data

This study uses a panel data set collected by the Young Lives study, following 3,000 children in Andhra Pradesh, India. Two cohorts of children, approximately 1,000 children aged 8 years old and 2,000 children aged 1 year old in 2002, have been followed since 2002 and four rounds of data were collected at child, household and community level in 2002, 2005, 2009 and 2013.[24] The present study will use longitudinal survey data of the 951 older cohort children, born in 1994–95, because the younger cohort are not old enough to analyze completed school progression as of 2013. The survey covers complete educational history as well as information on household members of the child.

Andhra Pradesh (including Telangana)[25] is fourth largest state in India by area with a population of over 84 million in 2011. It is divided into three regions, Coastal Andhra, Rayalaseema, and Telangana, with distinct regional patterns in climate, soil characteristics, and livelihood. It is one of the states that have successfully promoted the

[24] The Young Lives sample is distributed across the three main regions and covers about 100 communities (villages or urban wards) across 20 sub-districts. A careful comparison with representative data for Andhra Pradesh shows that the data in the Young Lives sample contains similar variation across comparable measures: a detailed explanation of the sampling methodology and the comparison of the characteristics of the Young Lives sample with the DHS sample on a range of observed characteristics is reported in Kumra (2008).

[25] The state of Andhra Pradesh was bifurcated into two states, Andhra Pradesh and Telengana in 2014.

IT and manufacturing industries. The State Government strategy to promote these industries as prime growth engine since as early as the mid-1990s seems to have resulted in the establishment of many IT and electronic companies (such a Microsoft, Oracle, and Dell), which are concentrated in Hyderabad. The state's industrial development policy targets to increase the contribution of manufacturing to GSDP from 9.95% (2013-14) to 15% by 2020 by attracting investment and creating employment opportunities for an additional 1 million people.

The Young Lives survey covers Hyderabad as well as other urban and rural areas in Andhra Pradesh, which gives the large variation in job opportunities across communities. Using the community level information, some variables are constructed to proxy the labor market environment in a locality to examine its effect on school progression.

3.3 Key features of schooling in India and Andhra Pradesh

The state of Andhra Pradesh follows the standard education system in India which consists of 5 years of lower primary, 3 years of upper primary, 2 years of lower secondary, and 2 years of upper secondary schools, followed by tertiary education.

Students must pass the exam at the end of lower secondary and upper secondary schools in order to proceed to the next level.

India has realized rapid improvement in basic education attainment due at least in part to government efforts to improve access to primary education. The government program *Sarva Shiksha Abhiyan* (SSA)[26] was implemented in 2001 to promote universal primary education (grade 1 to 8), and a midday meal scheme was adopted to provide cooked meals for students in all primary and upper primary public and private-aided schools in 2003.[27] The Right of Children to Free and Compulsory Education (RTE) Act came into effect in 2009. These actions led to an increase in the net primary school enrollment ratio from 83% in 2000 to 99.9% in 2010-11 (Government of India, 2014b).[28]

---

[26] Sarva Shiksha Abhiyan, or SSA, is an Indian Government program aimed at the universalization of elementary education "in a time bound manner", as mandated by the 86th Amendment to the Constitution of India making free and compulsory education to children between the ages of 6 and 14 (estimated to be 205 million children in 2001) a fundamental right.

[27] Private-aided schools are run under private management but receive government funding and support, have access to government schemes like the midday meal scheme, and follow the same regulations, including those for pay and tenure, as government schools (Singh and Dercon, 2014). Private-aided schools are common only in some states, and in AP, it accounts for 4% according to Mehta (2007).

[28] Net enrollment ratio (NER) is the number of children enrolled in a school who belong to the age group that officially corresponds to schooling, divided by the total population of the same age group. On the other hand, gross enrollment ratio (GER) is the number of children enrolled in a school, regardless of age, divided by the population of the age group that officially corresponds to the same level. In countries where many children enter school late or repeat a grade the GER can exceed 100%.

Now facing low attendance and low completion rates in secondary school (grade 9 to 12), the Indian government launched *Rashtriya Madhyamik Shiksha Abhiyan* (RMSA) in 2009 with the objectives of improving access to the descent quality of secondary education. The goal was set to raise the enrollment ratio from 52.3% (lower secondary) and 28.5% (higher secondary) in 2005-2006 to 90% and 75%, respectively by 2017. Despite this government effort, according to World Development Indicator, increase in secondary school gross enrollment ratio in India was modest, reaching only 68.9% for lower secondary in 2013-2014. Moreover, the lower secondary school completion rate in the same year was 80.9%, which means that even if children successfully transit from primary to lower secondary school, as many as 20% of them do not complete it.

One factor that must be considered in studies of schooling in India is gender disparity. Although the gender gap in primary school participation seems to have closed, girls in India have always been disadvantaged, not only in schooling but also in many other social and economic spheres. Girls marry much earlier than boys on average and receive lower wages than boys for the same work (Kingdon, 2007). In that light, and given the consequent lower expected returns on education for girls than for boys, parents might tend to invest less in girls' education. Another factor discouraging investment in girls' schooling is the high opportunity cost of sending girls to school: it

reduces their availability to help with domestic work. However, the recent increase in primary and secondary school enrollment has coincided with a narrowing of the gender gap in schooling, which can be attributed in part to an increase in returns on education of girls (World Bank, 2011). In fact, schooling decisions have been found to depend less on gender in recent decades (Kajisa and Palanichamy, 2010).

3.4 Literature review and Hypotheses

Vast literature on investment in schooling offer evidence that each additional year of schooling raises future earnings significantly (Psacharopoulos and Patrinos, 2004; Fulford, 2014). Expected returns to schooling depend on the local labor market environment such as average wage rate and availability of well-paid jobs. Descent work opportunity is one of the determinants consistently found to have a significant influence on investment in schooling. In India, new job opportunities brought about by recent globalization have created new demand for skilled workers, thereby stimulating increase in educational attainment (Oster and Steinberg, 2013; Shastry, 2012).

Oster and Steinberg (2013) find that establishment of IT service centers in neighborhoods in Karnataka, Andhra Pradesh, and Tamil Nadu states increased the primary school enrollment rate there. Munshi and Rosenzweig (2006) analyze the

impact of caste on choice of school and occupation in the globalizing economy of Bombay, and find that lower-caste girls, who have been kept out of the labor market historically, can take advantage of newly available white-collar job opportunities by attending English medium schools; it was found to be easier for lower-caste girls than boys to benefit from globalization because they are not tied to traditional networks, which direct boys to traditional jobs. These findings suggest that job opportunities are likely to drive households to increase investment in schooling, but labor market environment influences schooling decisions takes various forms and often differs according to gender.

The state government of Andhra Pradesh has been actively promoting development of the manufacturing sector, which has led to an increase in the number of factories. The presence of factories in a given locality leads to an increase in the number of high-return job opportunities for households there, which in turn is expected to increase family motivation to invest in schooling. Factory jobs may not be skill-intensive, but in most cases the level of education required is higher than that for casual jobs. This prompts our first hypothesis:

Hypothesis 1: The presence of factories in a locality has a positive effect on school progression among girls beyond secondary level.

Job opportunities might not always exert a positive influence on schooling; in many cases young children could be engaged in those jobs demanding little education. Parents might not reject the use of child labor if they consider it to be more advantageous than schooling. In fact, higher wages and better job opportunities for children have been found to lead to lower levels of schooling and increased incidence of child labor (Duryea and Arends-Kuenning, 2003; Shafiq, 2007). Such findings suggest that a child's wages represent the opportunity cost of sending that child to school and that high wages for children in unskilled occupations could in fact have a negative impact on school participation. Young children work in the casual job market for the most part. Casual jobs are often physical work, such as agricultural wage work and construction work, which favors boys over girls in terms of physical strength. In order to distinguish the positive and negative effects of job opportunities, we include wages for unskilled jobs in the analysis and propose the second hypothesis.

Hypothesis 2: Wages for unskilled jobs in a community have a negative impact on school progression, particularly for boys, not only in stages beyond lower secondary completion, at which point children are allowed to work, but also in the years before lower secondary entry, when child labor is illegal.

The labor market in India is distinctly different for girls and for boys because of the social norm that assigns specific roles to women. In an experiment by Jensen (2012),

recruiting services were provided for young women in rural Indian villages to support their job seeking in the business process outsourcing industry. This intervention increased the expected returns to girls' schooling and led to an increase in both investment in girls' human capital but also in marriage age.

The institution of early marriage for girls in India has a significant effect on girls' schooling and working status. Early marriage inhibits girls' schooling since they are generally compelled to discontinue their education when they marry. Parents may have lower educational aspirations for their daughters than for their sons, partly due to the social norm which encourages girls to marry young (Maertens, 2013). Jensen and Thornton (2003) examine data from Demographic and Health Surveys in India and find that average years of schooling received is only 0.8 years for females who marry before age 15, 1.8 years for women who marry between age 16 and 20, but more than 3 years for women who marry at age 21 or later after age 21. Desai and Andrist (2010) also find a positive association between marrying age and years of schooling, although the causal direction is unclear.

These findings from the literature demonstrate that choice of marriage and of schooling are made conjointly for girls. Sawada and Lokshin (2009) find that parents are more likely to let their first daughter marry early due to resource competition, and Kajisa and Palanichamy (2010) find that first born girls generally complete fewer years

of schooling. These findings suggest that parents regard their daughters' marriage as a means of reducing resource competition, which is a critical issue for households with little wealth and high dependency ratio.

In our work to gain insights into how girls' marriage and schooling choices are made, we classify girls into four groups: (1) in school and not married, (2) in school and married, (3) not in school and not married, and (4) not in school and married, and test the following hypotheses:

Hypothesis 3: Girls from households with less wealth and higher dependency ratio have a higher probability of marrying and being out of school.

Hypothesis 4: Eldest daughters have a higher probability of marrying and being out of school.

3.5.Explanatory variables and descriptive analysis

3.5.1 Framework of school progression

Before estimating the determinants of school progression, the basic framework of the analysis and the characteristics of the data set are discussed here. Figure 3.1 illustrates the sequence of educational decision making and how it is framed into our study. School progression in our analysis consists of eight stages: (S_0) is lower primary

school entry which is a given condition, (S_1) whether to complete lower primary school; (S_2) whether to enter upper primary school; (S_3) whether to complete upper primary school; (S_4) whether to enter lower secondary school; (S_5) whether to complete lower secondary school; (S_6) whether to enter upper secondary school; and (S_7) whether to complete upper secondary school. These 7 choices are analyzed in three separate ages, viz., age 12, age 15, and age 19, as depicted in Figure 3.1, using the data of the survey conducted in the year closest to each stage.[29] Age 12 covers choice 0 to choice 2, using the data from survey in 2006 together with history data from the first survey in 2003. Age 15 covers choice 2 to choice 4 with choice 2 (upper primary school entry) as given condition, using the data from survey in 2009 together with past rounds data. Lastly, age 19 covers choice 4 to choice 7 with choice 4 (lower secondary school entry) as a given condition, using the data from survey in 2013 together with past rounds data. Using time-variant and time-invariant background information in this approach makes it possible to conduct stage-specific analyses.

Table 3.1 shows the numbers of children who completed each of the eight educational stages from S_0 to S_7, and the progression rates from stage to stage. The figures in panel A and panel B represent the number of children who at least achieved

[29] The first round of survey was conducted in 2003 when the children were 8 years old, the second round in 2006 at age 12, the third round in 2009 at age 15, and the fourth round in 2013 at age 19.

each stage by age 15 and 19 respectively. For example, panel A shows that out of all 464 boys in our sample who entered lower primary school, 434 boys completed ; thus the progression rate is 0.98. The top row, proper age, is the predicted age of children other than those who repeated a grade or entered primary school later than age 6. In panel A it can be seen that few of the children enrolled in S_5 or higher started schooling before age 6.

Important observations can be made from Table 3.1 regarding critical stages and repetition. Critical stages are the educational stages where the progression rate is markedly lower than average. Panel B in Table 3.1 reveals that the progression rate from S_5 to S_6 is particularly low: only 89% of boys and 78% of girls who had completed lower secondary proceeded to upper secondary school. In contrast, once students enter upper secondary, most of them succeed in completing it, with a progression rate of 0.93 for boys and 0.92 for girls. For girls, transit from S_2 to S_3 seems to be another critical stage in terms of dropping out.

Another issue is repetition, as can be seen by comparing the figures in panel A and panel B. If all children entered lower secondary school at age 14, the historical data in Panels A and B should be identical. However, the figures in panel A indicate that only 349 out of 464 boys and 353 out of 487 girls actually entered lower secondary

school, with a subsequent increase in panel B to 399 for boys and 391 for girls. These gaps indicate that many children entered lower secondary after reaching age 15.

3.5.2 Descriptive statistics of school progression

Table 3.2 summarizes the descriptive statistics of individual, household, and community characteristics of sample children, which are used as explanatory variables in the regression analyses.[30] In order to examine the changes in those characteristics by children's educational level, the entire sample is divided into eight educational stages, as described in the previous section, according to the highest level of education completed when the children are age 19.

The first row shows the number of observations per schooling stage. That categorization is made according to the highest educational level achieved at age 19, which is the age at which upper secondary school is completed.

At individual level, there are two dummy variables, one for boys and the other for eldest boy or girl among siblings. The accumulated evidence suggests that girls in India are disadvantaged in schooling (Drèze and Kingdon, 2001; Duraisamy, 2002; Siddhu, 2011). Thus, it is expected that higher proportion of boys can proceed to higher level, but no clear trend emerges in Table 3.2. There are two variables which proxy

[30] The dataset covers 98 communities.

children's ability: Raven test score taken at the age 8, and a dummy variable for reading and writing ability.[31] Reading and writing skills increase with educational stages, but the Raven test results do not show such a tendency.

Household characteristics include dependency ratio, parental educational level, two indicators for household's wealth: size of owned agricultural land and household asset index.[32] This index, an average of three indices, housing quality index, consumer durables index, and housing services index, ranges in value between zero and one. Parental education level and household asset are expected to have a positive association with child educational level. A caste dummy is also included to indicate whether the family belongs to disadvantaged caste group such as Scheduled Caste (SC) and Scheduled Tribe (ST). We investigate whether and at what educational stages these disadvantaged caste groups have negative effects on child educational attainment.

[31] Raven test, a method of testing IQ, was taken at age 8 by all children in the sample took. Reading and writing skill is based on self-report and takes a value of one if a child can read sentences and can write without difficulty.

[32] Dependency ratio is defined as ratio of household members under the age of 15 to those between the age of 16 and 65. Owned land is not included in age 12 stage since it is not available in the round 2 survey. Wealth index is an average of three indices: Housing quality index (average of indicators: crowding, main material of walls, main material of roof, main material of floor), consumer durable index (average of set of dummy variables which takes the value 1 if they own radio, television, bicycle, motorbike, automobile, landline phone, mobile phone, refrigerator, fan), and housing services index (average of indicators: access to electricity, access to safe drinking water, access to sanitation, access to adequate fuel for cooking).

There are three community level variables.[33] The first is average daily wages for unskilled workers in a community, who are not only adults but also children. This is expected to reflect the opportunity costs of child being enrolled at school. Another variable related to labor market is a dummy variable which equals one if a community has a factory that employs local workers. Jobs at factories are mostly low-skilled or semi-skilled which may not require high human capital, even though it usually requires basic education and reading/writing skills. Therefore, it is expected that the existence of a factory in locality has a positive impact on finishing at least primary school, possibly higher education. Lastly, we control for availability of secondary school in a community. Primary school availability is not included in the regression analysis because all the communities have at least one primary school.

3.5.3 Descriptive statistics of girl's marital and schooling status

Table 3.3 reports the descriptive statics of variables used in an analysis of girl's choice between marriage and schooling The number of observations shows how many girls are in each status at age 15 and at age 19. The figures are presented for rural and urban areas respectively because there is a sharp contrast in the proportion of each status

[33] Communities are municipal wards and revenue villages in urban areas, and villages and their associated hamlets in rural areas.

between rural and urban areas. There is only a small number of girls who are married by age 15. At age 19, the number of married girls increases, but still less than those who are out of school without being married. It is also clear that more girls are married by age 19 in rural areas than in urban areas.

We use the same explanatory variables as school progression analysis. For all ages and areas, proportion of eldest girl and dependency ratio are highest for “married and not in school” girls. For the girls “in school”, reading/writing skills, parental education level, house wealth asset index are highest regardless of areas or ages. These observations are consistent with our third and fourth hypotheses.

3.6 School progression

3.6.1 Empirical issues

In school progression analysis, the empirical challenge is how to take account of the fact that a schooling decision made at one point is conditional on earlier decisions regarding the attainment of the preceding levels of education. In this study, the unique cohort panel data set is analyzed using a sequential logit model originally proposed by Mare (1980) to fully account for the conditionality of sequential decisions made in the course of progression through the stages of school.

There are two major criticisms of school progression analysis: identification and selection bias. “Identification” refers to the difficulty in school progression analysis of identifying different effects of background variables for different schooling stages because of the error variance assumption: if the regressors do no vary across stages within individual, which is usually the case for variables such as parental education level, social status, test scores at specific time, the different effect cannot be identified. It is demonstrated by Cameron and Heckman (1998) that unless one variable varies across transitions within individual, the error term is the only factor that varies across transition, in which case the assumption is violated. Some studies incorporate time-varying covariates in the model to identify the different effects of background by stages (Lucas, 2001; Lucas et al., 2011). In our study, the rich panel dataset provides various time-varying covariates, such as household level dependency ratio, asset level, owned land size, and community level labor market environment. Including them allows the model to identify the differences between the effects for each of seven stages.

The other issue, selection bias, occurs when unobserved heterogeneity is correlated with independent variables. It is a serious concern, particularly for the school progression analysis, because the population after the first stage is a nonrandom subgroup of the initial population. In a two transition analysis, bias can be reduced by introducing a variable that affects the first transition but not the second one. Lucas et al.

(2011) use proportion of the economy devoted to manufacturing and/or agriculture and argue that it has no effect on college entry given that the student has completed high school. However, this is not a viable solution for the current study, because it attempts to analyze seven progression stages simultaneously.

Mare (2011) proposes an alternative interpretation of bias and an approach to correction of the bias: the nonrandom selection after the first stage yields a population with unmeasured traits different from the initial population, which amounts to reweighting of the subgroup. This reweighting adjusts the measured and unmeasured traits so as to make them the same as for the initial group. However, as Mare (2011) points out, one needs to be aware that this approach assumes the selection to be random with respect to children's individual and family background, which "is considering a different institutional setting than the actual one we wish to study" (Mare, 2011). The purpose of this study is to obtain a detailed understanding of the mechanism of students staying in or dropping out of school at different stages in the current institutional setting in India. Therefore, instead of reweighting the group of students to match their attributes to those of the group in the first stage, a more practical approach is to analyze the behavior of subgroups who transit through a selective process as it happens in reality. We conduct a repeated logit analysis as a robustness check and present the results at the end of this section.

3.6.2 Analytical framework

Among the sample children, there are only 9 children out of the 951 older cohort children who were never enrolled in primary school. Therefore, entering primary school is a given condition in the model, which yields seven schooling decisions from lower primary completion to upper secondary completion. The sequential logit estimation models the probabilities of passing each of these seven stages indexed by *s*. After completing any level of education *s*, an individual *i* has an option to continue to the next level of education with probability p_k or to stop with probability (1–p_k).

Following Buis (2010), the probabilities of passing these transitions are estimated by logistic regression for each transition on the sub-sample. Probability of choice p_k is expressed as below.

$$\hat{p}_{kit} = \frac{exp(X_{it}\beta_k)}{1+exp(X_{it}\beta_k)} \quad \text{if} \quad p_{k-1,it} = 1 \qquad k\text{=1,2, t=12}$$

$$\hat{p}_{kit} = \frac{exp(X_{it}\beta_k)}{1+exp(X_{it}\beta_k)} \quad \text{if} \quad p_{k-1,it} = 1 \qquad k\text{=3, 4, t=15}$$

$$\hat{p}_{kit} = \frac{exp(X_{it}\beta_k)}{1+exp(X_{it}\beta_k)} \quad \text{if} \quad p_{k-1,it} = 1 \qquad k\text{=5, 6, 7, t=19}$$

where X_{it} are child, household, and community characteristics for individual *i at age t*, summarized in Table 3.2. In addition, children's ability is represented by Raven test

score taken at the age 8 as well as by a dummy variable for reading and writing ability at age 12.

3.6.3 Estimation results

The estimation results of the sequential progression model are summarized in Tables 3.4 and 5. Ages 12, 15, and 19 correspond to three stages of entire school progression as depicted in Figure 3.1. Children are categorized according to their educational status at age 12, at age 15, and at age 19, respectively. The figures presented are odds-ratio, and in addition to variables listed in the Table 3.2, region dummies are included to control for region level heterogeneity. The number of observation reduces as the stage proceeds since those who drop out of school by age 12 are excluded in the age 15 analysis, and those who drop out by age 15 are excluded in the age 19 analysis.

Rural areas vs urban areas

Table 3.4 reports the results of analysis by rural and urban areas. Striking is that the variables that represent children's ability, particularly reading and writing skills, present long-lasting positive effects on school progression.[34] Being able to read and

[34] Due to possible correlation between Raven test score and reading and writing skills, the regression is run without Raven test score and without reading and writing skills separately. The qualitative results were the same as the specification including both variables.

write at the age 12, which is the age children are supposed to have completed lower primary school, is a very important factor in continuing schooling decisions later on. This is a finding that can only be drawn by analyzing long-term cohort data and to our knowledge, this is the first study to demonstrate such long-lasting and positive effect of children‘s ability at early age on school progression.

Many household level attributes seem to have influence on school progression more strongly in rural areas than in urban areas. Parental education level, notably mother’s educational level, promotes the school progression at critical stage (upper secondary entry). ST or SC members are at a disadvantage in entering upper primary as well as lower secondary school only in rural areas. The positive effect on upper secondary entry in both areas is likely to be the results of the government support called post matric scholarship for Scheduled Caste students.

Two variables which proxy wealth level of the household, size of land and household asset index, are found to improve school progression, but only in rural areas. As Kajisa & Palanichamy (2010) discuss, the free education lowers the cost of schooling, and thus the cost of sending children equals the opportunity cost of time. Yet in rural areas where majority of households are farmer whose income is not stable, owning larger land and more wealth could contribute to income smoothing, which in turn ensures the children’s education attainment.

At community level, having a factory in vicinity increases the probability of staying in school beyond secondary level, which is consistent with our first hypothesis. Its impact is more evident in rural areas probably because in urban areas, there are many employment opportunities other than factories, while access to nonfarm jobs is limited in rural areas. On the contrary, average wages for unskilled workers lowers the probability of school progression. The effect is significant for younger children in rural areas than in urban areas. Since 80% of child workers are found in rural areas, it is possible that parents in rural areas consider practicing child labor from younger ages than parents in urban areas.

Availability of secondary school in a community is included in the analysis, but not reported because it does not have any significant impact on school progression in any specifications. The community level survey, conducted together with household survey, reveals that among all the secondary schools that the children attend, 92 are within a community and 70 are outside a community. This suggests that children of secondary school age manage to commute outside their own community. Moreover, the average time taken to go to secondary school is 20 minutes (from round 3 survey), which implies that even if the schools are outside their community, the school they attend is not far from their home. Therefore, the supply of secondary school is not likely to be a binding factor that keeps children out of school in the study area.

Gender difference is hardly observed except in entering upper secondary school in rural area where boys are at an advantage. This is in line with the findings of Kajisa & Palanichamy (2010) that the effect of gender has weakened over time. However, it is important to investigate why the boys have much higher probability in entering upper secondary school which is the stage a large number of children drop out. We will discuss this issue further in the next subsection where gender-wise analysis is conducted.

Boys vs girls

Table 3.5 summarizes the results of the same sequential model run separately for boys and girls. Comparing the boys and girls in the critical stage, i.e., upper secondary entry, it becomes obvious that reading and writing skill is the most significant factor which increases boy's probability of entering upper secondary school. By contrast, mother's education, presence of factory, and child's ability are the major significant contributing factors for girl's entering upper secondary school.

Eldest boys are more likely to complete upper primary and upper secondary school, but this effect is not observed for eldest girls. Since eldest girls are expected to reduce the resource competition either by marrying or working, it is unlikely that they

have higher probability of staying in school at any stage. This will be analyzed further in the next section. Persistent positive effects of Raven test score and reading and writing skills are evident for both boys and girls.

Household wealth seems to be an important factor for boys' entering and completing upper primary school. This may have to do with the fact that parent in general prefer to invest more on boy's education including sending them to private school.

Community level daily wages for unskilled worker negatively affects boys' completion of upper primary and lower secondary school, and entry of upper secondary school. However, it does not have any significant impact on girls' school progression, which confirms our second hypothesis.

Presence of factory in a community shows positive and differential impacts on school progression by child's gender. For boys, it increases the probability of upper primary completion and lower secondary entry. For girls, it has positive effects on the critical stage, upper secondary entry, and the magnitude of this effect is large. One possible reason is that parents understand the importance of education in getting a regular job and increase the investment in both boys' and girls' education, but they perceive the required education level to be higher for girls than for boys.

3.6.4 Repeated probit analysis

We conduct a repeated probit analysis using the same covariates as the sequential logit model conducted in the previous section. At each stage, probit analysis is run repeatedly using all the samples. For example, the dependent variable of analysis of stage 7 take the value one if a child complete upper secondary school and zero if they drop out before completing upper secondary regardless of the stage they stop schooling. The results are in Table 3.6 for rural and urban areas, and Table 3.7 for boys and girls.

The results of the analyses, whether by areas or by gender, show that repeated probit analysis using all the sample find households level attributes to have significant influence on school progression in general. This is coherent considering that the sequential logit compares a group of students who passed stage 5 against a group of students who passed stage 3 but did not pass stage 4, therefore these two groups are likely to have similar background. In contrast, individual and community level seem to maintain their influence over school progression though their magnitude might be smaller. Notably, the Raven test score and reading and writing skills show significant and persistent effect not only in sequential logit model but also in probit model.

3.7 Girl's marital and schooling status

3.7.1 Analytical framework

Although marriage is one of the factors that affect girl's educational attainment, it was not included in the above model due to its endogeneity; in other words, higher education could contribute to delaying marrying age, while marriage could discourage schooling. In this section, we conduct an analysis to investigate determinants of girl's marriage and schooling. Marriage does not necessary interfere with schooling if the marriage age is high. Therefore, analyses are conducted separately for the marriage at age 15 and at age 19 and we examine the differences in determinants.

Examining marrying age by gender, 8.6% of girls and 0% of boys are married by age 15, and 36.8% of girls and 1.9% of boys are married by age 19. From this, it is clear that marriage is unlikely to be the cause for boy's dropping out of school. Therefore, we analyze the determinants of schooling and marital status only for girls.

As explained in the previous section, girl's marital and schooling statuses consists of four types of mutually exclusive categories: (1) in school and not married, (2) in school and married, (3) not in school and not married, and (4) not in school and married. To examine the causes of dropping out of school, only the girls who were in school in previous rounds are included in the analysis. Out of 487 girls in our sample

children, proportion of each status are respectively 72%, 1%, 20%, and 7% at age 15, and 37%, 1%, 26%, and 36% at age 19.

In our sample, 20 per cent of the girls are in status (3), not in school and not married, at age 15, and 26 per cent at age 19, suggesting that there are other factors than marriage that pulls girls out of school. Parents may stop sending their eldest daughters to school and have them engage in wage work. This saves the cost of sending one child to school and at the same time increase the earning from wage work and, hence, reduces resource competition. Another possible cause is insufficient reading and writing skills. As shown in the previous analysis, girls who fail to acquire sufficient reading and writing school may not succeed in transiting to next level of education. Among the girls who are out of school and not married 69 per cent engage in paid work at the age 15, and 74 per cent at age 19. These observations support our assumption that marriage is not the only reason for girls' drop-outs and confirm that high proportion of who are out of school engage in some kind of wage work.

We estimate the girl's schooling and marital status using a multinomial logit model with the same explanatory variables as the sequential logit model. Since the status (2), in school and married, only accounts for 1%, it is excluded from the analysis. Therefore, the dependent variable includes three statuses and "in school and not married" ("in school" hereafter) is the base category.

Their probabilities are estimated by multinomial logit model as,

$$\Pr(z_{it} = j) = \frac{\exp(X_{it}\gamma_j)}{\sum_j \exp(X_{it}\gamma_j)} \ , \quad j=1,3,4,\ t=15,\ 19$$

where z_{it} is an indicator variable denoting the schooling and marital status of individual *i at age t* with respect to status *j*.

3.7.2 Estimation results

The results of multinomial logit are presented in Table 3.8. The results of statuses at age 15 does not include "married and not in school" in urban area since there are only one girl who fall into the group. This suggests that in urban areas, girls' drop-outs rarely accompany marriage.

As discussed in previous section, there are only 15 girls who are married by age 15. Therefore, marriage is unlikely the cause for the drop-outs of 15-year-old girls. Rather, it is lower reading and writing skills as well as higher dependency ratio and lower mother's educational level that increase the risk of drop-outs in rural areas. In urban areas, wealth increases the probability of staying in school.

At age 19, marriage plays a more important role. As we hypothesized, girls from households with higher dependency ratio and fewer assets are more likely to be married and out of school. We also find that eldest girls are more likely to be married,

which confirms our fourth hypothesis. Again, the reading and writing skills are vital factors and lowers the probability of being out of school without marriage.

Another finding is that mother's educational level lowers the probability of dropping out of school in urban areas. This is consistent with our findings from the previous analysis on school progression.

3.8 Conclusion

This chapter investigated the determinants of school progression using a unique panel data of children and their households in Andhra Pradesh, India. The dynamic framework adopted a sequential logit model accounting for the fact that schooling decisions made at one point are conditional on decisions to attain the previous level of education. The use of such unique cohort data and the adoption of dynamic framework allowed this study to offer two major contributions: it identified stage-specific factors influencing school progression and it investigated long-term effects of child's attributes at primary level.

It was found that the child's basic ability, characterized by the Raven test score and reading and writing skills, is a crucial factor for children to continue schooling up to

upper secondary level. Child's gender and school availability did not present strong effects on school progression, even though they had long been identified as important predictors of school participation. This is likely due to changes that took place in a transforming economy: increase in the number of schools, decline in schooling cost resulting from the free education policy and the midday meal scheme as well as other support from the government, and improvements in women's status. These are likely the salient factors influencing decision making on child's schooling.

Local labor market environment appeared to have significant effects on school progression as well. Availability of factory jobs was found to promote the secondary level of schooling by raising the expected returns to schooling. On the other hand, wages for unskilled workers, which represents the opportunity cost of staying at school, was found to negatively affect the schooling for boys who engage in child labor much more than girls do. Although children under age 14 are prohibited by law from working, children's schooling may be compromised when the opportunity cost is high. Enforcement of The Right of Children to Free and Compulsory Education Act and The Child Labor (Prohibition and Regulation) Act is one countermeasure against the high drop-out rate.

The dynamic framework also offered better understanding regarding the factors affecting schooling at critical stage where the progression rate is low: upper secondary

entry. In addition to child's ability which has positive influence throughout the stages, mother's education was found to increase the probability of girls' entering secondary school.

Furthermore, this study examined the role of marriage in girl's schooling. At age 15, there was only one married girl. Therefore, marriage was an unlikely cause for girl's drop-outs at age 15. Rather, it was the lack of either wealth or reading and writing skills which put the girls at the risk of dropping out. We also found that the resource competition, which arises from low level of wealth and high dependency ratio, drove the parents towards guiding their daughter to marriage by age 19. Eldest daughters, who are expected to reduce resource competition, were also found to have higher probability of marrying by age 19 as well.

Two important policy implications can be derived. One is that enhancement of education services supply is not a sufficient measure to encourage investment in children's education. It was made clear that parents adjust their schooling decisions in response to changes in local labor market conditions. In other words, ample availability of well-paid jobs in a locality could raise parental incentive to keep the children in school until they develop adequate human capital. The other is the importance of improvement of school quality. Despite the increase in school participation, the level of children's learning stays low (Pritchett and Beatty, 2012). The results of sequential logit

model provided evidence that reading and writing skills at age 12 have strong and persistent positive effects on keeping the children in school until completion of upper secondary. This reiterates the importance of enhancing school quality and ensuring that children attending school acquire the basic primary level skills.

Table 3.1 Progression rate by educational stage as of age 15 and 19

			S_0 Lower primary entry	S_1 Lower primary completion	S_2 Upper primary entry	S_3 Upper primary completion	S_4 Lower secondary entry	S_5 Lower secondary completion	S_6 Upper secondary entry	S_7 Upper secodary completion
		Proper age	6	11	11	14	14	16	16	18
Panel A: As of age 15	Boy	Number of children	464	454	439	397	349	255	63	4
		Progression rate		0.98	0.97	0.90	0.88	0.73	0.25	0.06
	Girl	Number of children	487	474	452	394	353	260	80	9
		Progression rate		0.97	0.95	0.87	0.90	0.74	0.31	0.11
Panel B: As of age 19	Boy	Number of children	464	454	440	409	399	381	339	315
		Progression rate		0.98	0.97	0.93	0.98	0.95	0.89	0.93
	Girl	Number of children	487	474	453	405	391	368	286	262
		Progression rate		0.97	0.96	0.89	0.97	0.94	0.78	0.92

Table 3.2 Individual, household, and community characteristics of children by highest education level achieved by age 19

	Age 12			Age 15		Age 19			All
	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	
	Lower primary entry	Lower primary completion	Upper primary entry	Upper primary completion	Lower secondary entry	Lower secondary completion	Upper secondary entry	Upper secodary completion	
Individual level									
Proportion of boys	0.43	0.40	0.39	0.42	0.44	0.34	0.50	0.55	0.43
Proportion of eldest boys/girls	0.39	0.49	0.49	0.46	0.59	0.52	0.58	0.63	0.52
Raven test score (age 8)	23.36	22.29	21.63	22.17	21.76	21.89	23.52	23.51	22.52
Can read and write (=1) (age 12)	0.00	0.17	0.35	0.54	0.37	0.48	0.63	0.81	0.42
Household level									
Dependency ratio	0.37	0.20	0.25	0.25	0.23	0.19	0.22	0.11	0.23
Father's years of schooling	0.96	1.03	1.63	2.42	2.95	3.16	4.42	5.59	2.77
Mother's years of schooling	0.43	0.71	0.77	0.92	1.51	1.27	1.73	3.71	1.38
Owned land (ha)[a]	2.79	1.23	1.74	1.25	1.30	2.19	2.10	1.90	1.81
House asset wealth index[ab]	0.33	0.37	0.37	0.50	0.48	0.60	0.58	0.63	0.48
ST/SC (=1)	0.35	0.46	0.42	0.42	0.32	0.26	0.35	0.29	0.36
Community level									
Daily wage of unskilled worker (rupee)[a]	46.90	49.77	52.86	60.06	59.41	104.57	101.24	107.02	72.73
Factory within the community(=1)[a]	0.26	0.34	0.42	0.21	0.46	0.30	0.48	0.49	0.37
Secondary school in the community (=1)	n.a.	n.a.	n.a.	0.71	0.61	0.70	0.56	0.64	0.65
Number of observations	23	35	79	24	41	124	48	577	951
Boy	10	14	31	10	18	42	24	315	464
Girl	13	21	48	14	23	82	24	262	487

[a] These variables use data from the survey conducted in the year closest to each stage: :age 12 uses round 2 (2006), age 15 uses round 3 (2009), andage 19 uses round 4 (2013).

[b] This is an average of three index: Housing quality index, consumer durable index, and housing services index

Table 3.3 Individual, household, and community characteristics of children by marital and schooling status (girls only)

	Age 15 (in school at age 12)[c]						Age 19 (in school at age 15)[d]					
	Urban			Rural			Urban			Rural		
	In school	Not married and not in school	Married and not in school [e]	In school	Not married and not in school	Married and not in school [c]	In school	Not married and not in school	Married and not in school [c]	In school	Not married and not in school	Married and not in school [c]
Individual level												
Proportion of eldest girl	0.61	0.45	–	0.64	0.48	0.60	0.59	0.52	0.70	0.67	0.54	0.79
Raven test score (age 8)	23.89	24.73	–	22.61	23.10	22.53	24.46	23.89	22.15	23.28	21.04	22.86
Can read and write (=1) (age 12)	0.87	0.64	–	0.69	0.35	0.53	0.90	0.67	0.70	0.79	0.60	0.64
Household level												
Dependecy ratio	0.10	0.32	–	0.15	0.27	0.46	0.10	0.03	0.26	0.13	0.12	0.21
Father's years of schooling	8.20	4.09	–	4.53	1.70	0.73	9.81	5.59	4.70	5.31	3.44	3.24
Mother's years of schooling	5.75	2.36	–	2.44	0.30	1.33	6.82	3.41	1.30	3.37	1.69	1.29
Owned land (ha)[a]	0.29	0.39	–	2.24	2.06	2.67	0.84	0.11	1.85	1.96	1.86	2.76
House asset wealth index[ab]	0.69	0.58	–	0.48	0.45	0.48	0.74	0.71	0.69	0.59	0.55	0.58
ST/SC (=1)	0.12	0.18	–	0.37	0.40	0.27	0.16	0.11	0.20	0.37	0.43	0.29
Community level												
Daily wage of unskilled worker (rupee)[a]	131.7	130.8	–	97.3	99.8	95.2	149.9	138.4	167.4	180.7	185.6	173.9
Factory within the community(=1)[a]	0.67	0.55	–	0.23	0.18	0.13	0.69	0.52	0.40	0.29	0.17	0.19
Secondary school in the community (=1)	0.89	0.82	–	0.67	0.75	0.80	0.90	0.89	0.95	0.62	0.57	0.78
Number of observations	97	11	1	252	40	15	68	27	20	107	70	58

[a] These variables use data from the survey conducted in the year closest to each stage: :age 12 uses round 2 (2006), age 15 uses round 3 (2009), andage 19 uses round 4 (2013).

[b] This is an average of three index: Housing quality index, consumer durable index, and housing services index

[c] The sample only include those who were in school at the age of 12

[d] The sample only include those who were in school at the age of 15

[e] There is only one girl who was married and not in school in urban area at the age of 15, therefore this category excluded in the analysis.

Table 3.4 Estimation results of sequential logit model on school progression by urban and rural areas

	Rural							*Urban*						
	Age 12		Age 15		Age 19			Age 12		Age 15		Age 19		
	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_1	s_2	s_3	s_4	s_5	s_6	s_7
	Lower primary completion	Upper primary entry	Upper primary completion	Lower secondary entry	Lower secondary completion	Upper secondary entry	Upper secodary completion	Lower primary completion	Upper primary entry	Upper primary completion	Lower secondary entry	Lower secondary completio	Upper secondary entry	Upper secodary completion
Individual level														
Boy(=1)	1.157	1.086	1.346	0.896	1.154	2.879***	1.151	1.463	0.685	1.408	0.538	3.701	1.067	2.007
	(0.463)	(0.283)	(0.367)	(0.279)	(0.503)	(0.810)	(0.448)	(1.220)	(0.359)	(0.992)	(0.266)	(3.120)	(0.462)	(1.434)
Eldest boy/girl (=1)	0.742	0.946	1.361	1.017	0.817	1.174	1.214	0.901	1.143	1.599	1.388	1.149	2.355*	0.890
	(0.313)	(0.267)	(0.391)	(0.362)	(0.348)	(0.335)	(0.448)	(0.550)	(0.542)	(0.765)	(0.668)	(0.749)	(1.076)	(0.737)
Raven test score (age 8)	1.012	1.066***	1.021	1.089***	1.047	1.066**	0.964	1.098	1.205***	1.080	1.130***	1.109	1.052	1.021
	(0.038)	(0.025)	(0.025)	(0.030)	(0.042)	(0.028)	(0.039)	(0.107)	(0.062)	(0.074)	(0.053)	(0.081)	(0.043)	(0.058)
Can read and write (=1) (age 12)	11.142***	2.916***	4.515***	1.536	3.951***	4.298***	2.436**	7.159**	2.028	4.955**	1.469	10.912***	5.912***	1.203
	(6.850)	(0.786)	(1.289)	(0.467)	(1.786)	(1.175)	(0.887)	(5.523)	(1.210)	(3.083)	(0.836)	(8.585)	(2.763)	(1.244)
Household level														
Dependency ratio[a]	1.177	1.210	1.307	0.708	0.543	0.470	0.108***	1.038	1.725	1.201	0.748	0.355	0.311	2.143
	(0.475)	(0.308)	(0.590)	(0.272)	(0.559)	(0.284)	(0.079)	(0.514)	(0.840)	(0.468)	(0.357)	(0.473)	(0.303)	(2.707)
Father's years of schooling	1.044	0.992	1.155***	0.953	1.056	1.064	0.963	0.995	1.002	1.032	1.042	1.060	0.982	0.959
	(0.075)	(0.037)	(0.062)	(0.034)	(0.077)	(0.044)	(0.051)	(0.104)	(0.057)	(0.076)	(0.052)	(0.068)	(0.056)	(0.079)
Mother's years of schooling	1.815**	0.981	1.147	1.021	0.999	1.216***	1.161	1.034	1.089	1.093	1.082	1.037	1.134*	1.127
	(0.483)	(0.058)	(0.099)	(0.063)	(0.106)	(0.090)	(0.107)	(0.131)	(0.068)	(0.097)	(0.076)	(0.101)	(0.077)	(0.138)
ST/SC (=1)	0.733	0.439***	0.879	0.593*	2.285	1.829*	1.134	1.148	1.997	0.668	0.898	0.362	3.142*	0.367
	(0.280)	(0.118)	(0.255)	(0.186)	(1.276)	(0.571)	(0.486)	(1.234)	(1.378)	(0.557)	(0.531)	(0.246)	(2.167)	(0.355)
Owned land (ha)[a]			0.979	1.156**	1.270*	1.043	0.990			1.995	0.935	1.751	0.971	1.347
			(0.047)	(0.066)	(0.159)	(0.061)	(0.051)			(1.141)	(0.045)	(1.034)	(0.032)	(0.469)
House asset wealth index[ab]	0.979	6.129*	2.508	4.026	18.522*	0.635	1.241	5.779	1.358	5.047	0.906	0.144	1.098	1.294
	(1.512)	(5.819)	(2.422)	(4.321)	(31.698)	(0.652)	(1.630)	(18.674)	(2.368)	(17.578)	(2.457)	(0.520)	(3.255)	(5.404)
Community level														
Daily wage of unskilled worker[a]	1.039	1.000	0.974***	0.998	0.996*	1.002	1.004	1.019	0.968	1.015	0.985	1.005	0.995**	1.003
	(0.036)	(0.016)	(0.008)	(0.010)	(0.003)	(0.003)	(0.004)	(0.027)	(0.023)	(0.015)	(0.013)	(0.007)	(0.002)	(0.006)
Factory within the community(=1)[a]	2.993	0.598	2.270**	1.252	3.742*	3.194***	0.808	0.233	1.360	1.277	2.220	0.738	1.412	0.000***
	(3.124)	(0.260)	(0.866)	(0.513)	(2.720)	(1.295)	(0.391)	(0.496)	(1.069)	(0.855)	(1.466)	(0.847)	(0.951)	(0.000)
Observations	697	697	645	645	534	534	534	233	233	229	229	242	242	242

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1; Coefficients are odds-ratio; Region dummies are included to control for region level heterogeneity.
Regional dummy variables and availability of secondary in the community are not reported but included.

[a] These variables uses data from the survey round close to the time each choice is made:age 12 uses round 2 (2006), age 15 uses round 3 (2009), andage 19 uses round 4 (2013).

[b] This is an average of three index: Housing quality index, consumer durable index, and housing services index; [c] Data is not available; [d] Not applicable.

Table 3.5 Estimation results of sequential logit model on school progression by gender

	Boys							*Girls*						
	Age 12		Age 15		Age 19			Age 12		Age 15		Age 19		
	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_1	s_2	s_3	s_4	s_5	s_6	s_7
	Lower primary completion	Upper primary entry	Upper primary completion	Lower secondary entry	Lower secondary completion	Upper secondary entry	Upper secodary completion	Lower primary completion	Upper primary entry	Upper primary completion	Lower secondary entry	Lower secondary completion	Upper secondary entry	Upper secodary completion
Individual level														
Eldest boy/girl	0.633	1.044	2.415**	1.032	0.801	1.920	3.507**	0.886	0.884	1.081	1.141	0.954	1.221	0.451
	(0.360)	(0.376)	(1.031)	(0.421)	(0.429)	(0.843)	(1.875)	(0.419)	(0.308)	(0.343)	(0.435)	(0.492)	(0.365)	(0.252)
Raven test score (age 8)	1.031	1.108***	1.032	1.120***	0.973	1.020	0.970	1.015	1.077**	1.024	1.076**	1.135***	1.094***	0.986
	(0.039)	(0.031)	(0.041)	(0.035)	(0.048)	(0.032)	(0.039)	(0.048)	(0.037)	(0.029)	(0.037)	(0.055)	(0.032)	(0.041)
Can read and write (=1) (age 12)	13.104***	2.515***	6.773***	1.165	3.815**	8.740***	3.652***	7.470***	2.530***	2.572***	1.738	6.021***	2.707***	1.551
	(9.612)	(0.899)	(2.642)	(0.453)	(2.131)	(3.234)	(1.692)	(3.858)	(0.868)	(0.841)	(0.627)	(3.180)	(0.792)	(0.752)
Household level														
Dependency ratio[a]	0.555	1.139	0.582	0.560	1.572	0.408	0.068***	2.249	1.482	2.434	0.997	0.125**	0.462	0.434
	(0.243)	(0.362)	(0.246)	(0.256)	(1.725)	(0.378)	(0.064)	(1.267)	(0.468)	(1.357)	(0.438)	(0.131)	(0.304)	(0.452)
Father's years of schooling	1.027	0.938	1.057	0.959	1.098	1.018	0.909	1.017	1.031	1.121**	0.997	1.012	1.057	0.976
	(0.104)	(0.043)	(0.062)	(0.041)	(0.100)	(0.047)	(0.069)	(0.085)	(0.042)	(0.055)	(0.047)	(0.065)	(0.042)	(0.060)
Mother's years of schooling	1.129	1.130*	1.164	1.131*	1.009	1.045	1.277*	1.163	0.960	1.097	0.977	1.026	1.216***	1.131
	(0.165)	(0.080)	(0.116)	(0.081)	(0.120)	(0.074)	(0.188)	(0.158)	(0.059)	(0.072)	(0.063)	(0.090)	(0.073)	(0.097)
ST/SC (=1)	1.454	0.654	1.477	0.706	1.812	1.891	0.917	0.542	0.456**	0.534*	0.567	0.952	1.763	0.859
	(0.774)	(0.240)	(0.636)	(0.301)	(1.091)	(0.906)	(0.466)	(0.261)	(0.144)	(0.193)	(0.213)	(0.619)	(0.619)	(0.527)
Owned land (ha)[a]			0.988	1.009	1.277*	1.084	1.022			1.023	1.215	1.393	1.004	0.991
			(0.059)	(0.046)	(0.168)	(0.080)	(0.070)			(0.075)	(0.147)	(0.321)	(0.041)	(0.076)
House asset wealth index[ab]	4.247	8.601*	33.000**	1.524	20.787	2.470	11.022	1.529	1.493	0.296	7.073	1.164	0.440	0.293
	(11.240)	(10.130)	(52.385)	(2.350)	(49.009)	(4.098)	(21.178)	(2.813)	(1.729)	(0.349)	(9.632)	(2.976)	(0.531)	(0.494)
Urban(=1)	0.232	0.601	0.742	0.333	0.736	0.322**	0.709	0.263	1.061	0.844	1.940	0.884	0.768	1.248
	(0.221)	(0.351)	(0.581)	(0.227)	(0.508)	(0.180)	(0.589)	(0.282)	(0.783)	(0.494)	(1.821)	(0.550)	(0.328)	(0.716)
Community level														
Daily wage of unskilled worker[a]	1.016	0.988	0.977**	0.995	0.996*	0.995**	1.001	1.009	0.996	0.998	0.988	0.999	0.999	1.001
	(0.025)	(0.016)	(0.010)	(0.009)	(0.002)	(0.002)	(0.005)	(0.021)	(0.019)	(0.007)	(0.009)	(0.004)	(0.004)	(0.003)
Factory within the community(=1)[a]	0.945	0.753	2.440*	3.012**	2.158	1.604	1.190	0.870	1.097	1.027	0.645	0.825	3.329***	0.404
	(0.819)	(0.403)	(1.220)	(1.372)	(1.568)	(0.840)	(0.705)	(0.686)	(0.589)	(0.432)	(0.340)	(0.544)	(1.419)	(0.280)
Observations	455	455	433	433	393	393	393	475	475	441	441	383	383	383

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1; Coefficients are odds-ratio; Region dummies are included to control for region level heterogeneity.
Regional dummy variables and availability of secondary in the community are not reported but included.

[a] These variables uses data from the survey round close to the time each choice is made:age 12 uses round 2 (2006), age 15 uses round 3 (2009), andage 19 uses round 4 (2013).

[b] This is an average of three index: Housing quality index, consumer durable index, and housing services index; [c] Data is not available; [d] Not applicable.

Table 3.6 Estimation results of repeated logit model on school progression by urban and rural areas

	Rural							*Urban*						
	Age 12		Age 15		Age 19			Age 12		Age 15		Age 19		
	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_1	s_2	s_3	s_4	s_5	s_6	s_7
	Lower primary completion	Upper primary entry	Upper primary completion	Lower secondary entry	Lower secondary completion	Upper secondary entry	Upper secodary completion	Lower primary completion	Upper primary entry	Upper primary completion	Lower secondary entry	Lower secondary completion	Upper secondary entry	Upper secodary completion
Individual level														
Boy(=1)	1.157	1.105	1.458	1.177	1.376	2.173***	1.950***	1.463	0.794	1.498	0.815	1.600	1.364	1.408
	(0.463)	(0.255)	(0.341)	(0.237)	(0.303)	(0.454)	(0.394)	(1.220)	(0.351)	(0.923)	(0.309)	(0.775)	(0.469)	(0.456)
Eldest boy/girl (=1)	0.742	0.887	1.397	1.185	1.076	1.203	1.218	0.901	1.139	1.742	1.642	3.682***	3.020***	2.364***
	(0.313)	(0.220)	(0.350)	(0.255)	(0.243)	(0.255)	(0.244)	(0.550)	(0.486)	(0.767)	(0.613)	(1.574)	(1.023)	(0.756)
Raven test score (age 8)	1.012	1.056**	1.010	1.051**	1.029	1.055**	1.035*	1.098	1.178***	1.116*	1.123***	1.080*	1.054*	1.053*
	(0.038)	(0.023)	(0.023)	(0.021)	(0.025)	(0.023)	(0.021)	(0.107)	(0.057)	(0.067)	(0.043)	(0.047)	(0.034)	(0.032)
Can read and write (=1) (age 12)	11.142***	3.925***	6.649***	4.110***	5.777***	6.684***	5.901***	7.159**	3.172**	5.801***	2.873***	3.636***	5.116***	4.587***
	(6.850)	(0.936)	(1.650)	(0.819)	(1.336)	(1.438)	(1.205)	(5.523)	(1.588)	(3.225)	(1.172)	(1.593)	(1.817)	(1.557)
Household level														
Dependency ratio[a]	1.177	1.234	1.070	0.915	0.356***	0.380**	0.237***	1.038	1.384	1.505	1.049	0.291	0.281	0.412
	(0.475)	(0.269)	(0.370)	(0.252)	(0.136)	(0.154)	(0.101)	(0.514)	(0.464)	(0.602)	(0.418)	(0.255)	(0.245)	(0.315)
Father's years of schooling	1.044	1.003	1.163***	1.056*	1.106***	1.102***	1.063**	0.995	1.001	1.062	1.058	1.131***	1.065*	1.039
	(0.075)	(0.034)	(0.052)	(0.030)	(0.042)	(0.033)	(0.030)	(0.104)	(0.052)	(0.066)	(0.042)	(0.049)	(0.040)	(0.037)
Mother's years of schooling	1.815**	1.024	1.156*	1.058	1.159**	1.214***	1.209***	1.034	1.076	1.069	1.068	1.011	1.063	1.084*
	(0.482)	(0.060)	(0.088)	(0.051)	(0.077)	(0.067)	(0.058)	(0.131)	(0.062)	(0.076)	(0.056)	(0.062)	(0.053)	(0.050)
ST/SC (=1)	0.733	0.482***	0.871	0.711	1.207	1.526*	1.468*	1.148	1.765	0.754	0.806	0.571	1.359	0.915
	(0.280)	(0.113)	(0.216)	(0.152)	(0.288)	(0.339)	(0.311)	(1.234)	(1.066)	(0.572)	(0.395)	(0.271)	(0.596)	(0.396)
Owned land (ha)[a]			0.989	1.038	1.077*	1.067	1.049			2.163	0.950	1.394	1.008	1.024
			(0.039)	(0.036)	(0.048)	(0.050)	(0.040)			(1.403)	(0.038)	(0.313)	(0.040)	(0.041)
House asset wealth index[ab]	0.979	4.090*	4.517*	4.683**	6.833**	2.745	2.615	5.779	3.111	10.203	2.258	2.815	1.134	1.815
	(1.512)	(3.500)	(4.009)	(3.396)	(5.680)	(2.155)	(1.976)	(18.674)	(5.667)	(30.517)	(4.246)	(5.452)	(2.306)	(3.505)
Community level														
Daily wage of unskilled worker[a]	1.039	1.007	0.978***	0.983***	0.997*	0.999	0.999	1.019	0.981	1.013	0.994	0.997	0.995***	0.996**
	(0.036)	(0.015)	(0.007)	(0.006)	(0.002)	(0.002)	(0.002)	(0.027)	(0.018)	(0.014)	(0.009)	(0.002)	(0.002)	(0.002)
Factory within the community(=1)[a]	2.993	0.784	1.891**	1.697*	2.251***	3.176***	2.369***	0.233	0.887	0.897	1.451	0.757	0.805	0.672
	(3.124)	(0.317)	(0.582)	(0.463)	(0.669)	(0.932)	(0.654)	(0.496)	(0.646)	(0.568)	(0.680)	(0.405)	(0.423)	(0.319)
Secondary school in the community (=1)[a]			1.439	1.238	1.015	0.956	1.048			0.153*	0.317*	1.244	1.090	0.525
			(0.354)	(0.254)	(0.235)	(0.209)	(0.231)			(0.165)	(0.211)	(0.770)	(0.621)	(0.292)
Observations	697	697	696	696	664	664	664	233	233	235	235	268	268	268

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1; Coefficients are odds-ratio; Region dummies are included to control for region level heterogeneity.
Regional dummy variables and availability of secondary in the community are not reported but included.

[a] These variables uses data from the survey round close to the time each choice is made:age 12 uses round 2 (2006), age 15 uses round 3 (2009), andage 19 uses round 4 (2013).

[b] This is an average of three index: Housing quality index, consumer durable index, and housing services index; [c] Data is not available; [d] Not applicable.

Table 3.7 Estimation results of repeated logit model on school progression by gender

	Boys							*Girls*						
	Age 12		Age 15			Age 19		Age 12		Age 15			Age 19	
	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_1	s_2	s_3	s_4	s_5	s_6	s_7
	Lower primary completion	Upper primary entry	Upper primary completion	Lower secondary entry	Lower secondary completion	Upper secondary entry	Upper secodary completion	Lower primary completion	Upper primary entry	Upper primary completion	Lower secondary entry	Lower secondary completion	Upper secondary entry	Upper secodary completion
Individual level														
Eldest boy/girl (=1)	0.633	0.966	2.679**	1.614	2.393***	2.301***	2.702***	0.886	0.903	1.105	1.107	1.053	1.236	1.007
	(0.360)	(0.309)	(1.025)	(0.505)	(0.750)	(0.661)	(0.742)	(0.419)	(0.266)	(0.300)	(0.268)	(0.274)	(0.293)	(0.228)
Raven test score (age 8)	1.031	1.099***	1.030	1.090***	1.009	1.015	1.008	1.015	1.068**	1.021	1.047*	1.055**	1.087***	1.066***
	(0.039)	(0.031)	(0.039)	(0.030)	(0.031)	(0.027)	(0.025)	(0.048)	(0.033)	(0.027)	(0.025)	(0.028)	(0.027)	(0.025)
Can read and write (=1) (age 12)	13.104***	3.658***	11.370***	3.986***	7.020***	9.386***	8.803***	7.470***	3.466***	3.636***	3.129***	3.861***	3.826***	3.448***
	(9.612)	(1.134)	(4.146)	(1.019)	(2.184)	(2.517)	(2.263)	(3.858)	(1.048)	(1.001)	(0.751)	(0.974)	(0.887)	(0.794)
Household level														
Dependency ratio[a]	0.555	0.869	0.617	0.559	0.836	0.608	0.256**	2.249	1.687*	1.915	1.459	0.214***	0.262***	0.274***
	(0.243)	(0.247)	(0.256)	(0.227)	(0.525)	(0.394)	(0.159)	(1.267)	(0.473)	(0.804)	(0.495)	(0.088)	(0.127)	(0.129)
Father's years of schooling	1.027	0.954	1.119*	1.020	1.138***	1.079**	1.035	1.017	1.033	1.116***	1.076**	1.086**	1.085***	1.065**
	(0.104)	(0.041)	(0.066)	(0.036)	(0.049)	(0.036)	(0.035)	(0.085)	(0.039)	(0.046)	(0.033)	(0.038)	(0.032)	(0.030)
Mother's years of schooling	1.129	1.137**	1.112	1.117**	1.111	1.086	1.128**	1.163	0.990	1.083	1.025	1.036	1.136***	1.136***
	(0.165)	(0.073)	(0.091)	(0.060)	(0.077)	(0.055)	(0.056)	(0.158)	(0.057)	(0.061)	(0.047)	(0.057)	(0.053)	(0.047)
ST/SC (=1)	1.454	0.788	1.237	1.003	1.332	1.540	1.410	0.542	0.460***	0.621	0.568**	0.834	1.189	1.158
	(0.774)	(0.262)	(0.461)	(0.305)	(0.426)	(0.462)	(0.406)	(0.261)	(0.130)	(0.186)	(0.147)	(0.230)	(0.301)	(0.285)
Owned land (ha)[a]			1.031	1.016	1.158**	1.148**	1.113**			1.002	1.035	1.078	1.020	1.016
			(0.059)	(0.044)	(0.078)	(0.067)	(0.059)			(0.047)	(0.046)	(0.060)	(0.037)	(0.032)
House asset wealth index[ab]	4.247	6.806*	33.086**	10.302**	15.190**	9.522**	14.395**	1.529	1.775	0.882	1.832	3.341	1.391	1.036
	(11.240)	(7.574)	(47.854)	(11.223)	(18.567)	(10.377)	(14.962)	(2.813)	(1.917)	(0.886)	(1.613)	(3.270)	(1.304)	(0.918)
Urban(=1)	0.232	0.482	0.711	0.443	0.441*	0.330**	0.357***	0.263	0.701	0.796	0.992	0.918	0.703	0.769
	(0.221)	(0.261)	(0.488)	(0.223)	(0.216)	(0.144)	(0.142)	(0.282)	(0.469)	(0.394)	(0.460)	(0.323)	(0.249)	(0.260)
Community level														
Daily wage of unskilled worker[a]	1.016	0.995	0.983*	0.988*	0.997*	0.995***	0.996**	1.009	1.000	0.996	0.993	0.998	0.998	0.999
	(0.025)	(0.014)	(0.009)	(0.007)	(0.002)	(0.002)	(0.002)	(0.021)	(0.016)	(0.006)	(0.006)	(0.002)	(0.002)	(0.002)
Factory within the community(=1)[a]	0.945	0.787	1.394	1.976**	2.158*	2.053**	1.804*	0.870	1.031	1.248	0.999	1.096	1.995**	1.446
	(0.819)	(0.391)	(0.599)	(0.667)	(0.872)	(0.732)	(0.602)	(0.686)	(0.477)	(0.447)	(0.317)	(0.365)	(0.637)	(0.440)
Secondary school in the community (=1)[a]			1.018	0.997	1.045	1.085	1.011			1.037	0.957	0.750	0.586*	0.664
			(0.332)	(0.274)	(0.334)	(0.325)	(0.296)			(0.324)	(0.256)	(0.221)	(0.160)	(0.176)
Observations	455	455	456	456	456	456	456	475	475	475	475	476	476	476

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1; Coefficients are odds-ratio; Region dummies are included to control for region level heterogeneity.
Regional dummy variables and availability of secondary in the community are not reported but included.

[a] These variables uses data from the survey round close to the time each choice is made:age 12 uses round 2 (2006), age 15 uses round 3 (2009), andage 19 uses round 4 (2013).

[b] This is an average of three index: Housing quality index, consumer durable index, and housing services index; [c] Data is not available; [d] Not applicable.

Table 3.8 Multinomial logit estimation of girl's marital and schooling status at age 15 and at age 19 (Base category: In school)

	Age 15 (in school at age 12)				Age 19 (in school at age 15)			
	Urban		Rural		Urban		Rural	
	Not married and not in school	Married and not in school [c]	Not married and not in school	Married and not in school	Not married and not in school	Married and not in school	Not married and not in school	Married and not in school
Individual level								
Proportion of eldest girl	-0.109*	-	-0.06	0.003	-0.087*	0.101*	-0.074	0.119*
	[0.043]	-	[0.033]	[0.008]	[0.035]	[0.049]	[0.048]	[0.051]
Raven test score (age 8)	0.011	-	0.004	0.000	-0.002	0.003	-0.014***	0.001
	[0.006]	-	[0.005]	[0.001]	[0.003]	[0.003]	[0.003]	[0.006]
Can read and write (=1) (age 12)	-0.02	-	-0.118***	-0.015	-0.174***	0.003	-0.151***	-0.008
	[0.030]	-	[0.033]	[0.018]	[0.039]	[0.025]	[0.038]	[0.031]
Household level								
Dependency ratio[a]	0.069	-	0.052*	-0.074***	-0.754**	0.276***	-0.196	0.162
	[0.053]	-	[0.026]	[0.008]	[0.233]	[0.083]	[0.150]	[0.135]
Father's years of schooling	-0.004	-	-0.006	-0.013	-0.018	0.003	-0.002	-0.002
	[0.006]	-	[0.008]	[0.007]	[0.012]	[0.009]	[0.005]	[0.008]
Mother's years of schooling	-0.004	-	-0.029**	0.009	-0.011	-0.016**	-0.011	-0.003
	[0.007]	-	[0.011]	[0.005]	[0.014]	[0.006]	[0.008]	[0.006]
ST/SC (=1)	-0.077	-	0.013	-0.043	-0.081	-0.054	0.06	-0.147
	[0.046]	-	[0.074]	[0.043]	[0.062]	[0.060]	[0.044]	[0.089]
Owned land (ha)[a]	0.025	-	0.000	0.004***	-0.055	0.013	-0.004	0.01
	[0.021]	-	[0.004]	[0.001]	[0.029]	[0.007]	[0.010]	[0.006]
House asset wealth index[ab]	-1.115***	-	0.029	-0.274***	0.438	-0.657*	0.304***	-0.803***
	[0.283]	-	[0.117]	[0.054]	[0.422]	[0.327]	[0.072]	[0.159]
Community level								
Daily wage of unskilled worker[a]	0.000	-	0.001	0.000	0.000	0.000	0.000	0.000
	[0.001]	-	[0.001]	[0.001]	[0.000]	[0.000]	[0.001]	[0.001]
Factory within the community(=1)[a]	0.217	-	-0.021	-0.004	0.081	-0.091	-0.159***	-0.027
	[0.123]	-	[0.045]	[0.030]	[0.062]	[0.065]	[0.027]	[0.079]
N	108	-	297		112		231	

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1; Figures are marginal effects.

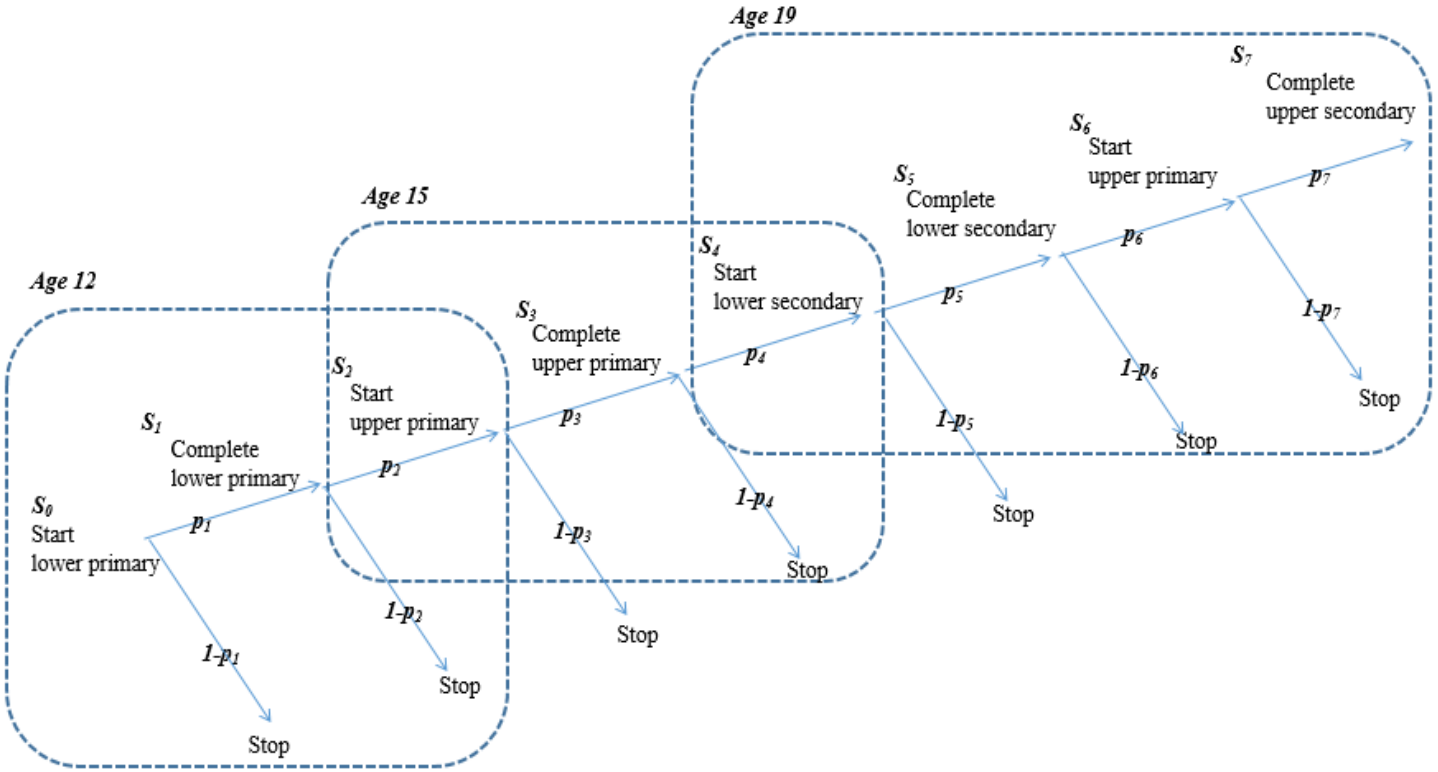
Regional dummy variables and availability of secondary in the community are not reported but included.

[a] These variables use data from the survey conducted in the year closest to each stage: :age 12 uses round 2 (2006), age 15 uses round 3 (2009), and age 19 uses round 4 (2013).

[b] This is an average of three index (Housing quality index, consumer durable index, and housing services index) of the girl's original household before marriage.

[c] There is only one girl who was married and not in school in urban area at the age of 15, therefore this category excluded in the analysis.

Figure 3.1 Decision making tree in school progression



Chapter 4

Conclusion

4.1 Summary of major findings

This dissertation was an attempt to identify the determinants of job choice of rural households in eastern India, and the determinants of school progression in Andhra Pradesh state. The job choice estimation in chapter 2 used a large scale survey data from 2015 with rich information about their economic activities as well as other household and individual attributes. This primary dataset allowed me to address the issues of nonfarm job classification that was not fully dealt with in many of past studies. The analysis in chapter 2 contributes to the literature on off-farm diversification by (1) classifying the nonfarm jobs in small groups and (2) including both local and migratory nonfarm jobs. This is a critical issue in accounting for the rural household's job choice in a dynamic labor market which has been rapidly changing in India.

Jobs were categorized into eight groups in this study: self-employed farming (including housewives and inactive workers), agricultural wage, local construction, local factory, local service and self-employment, migratory construction, migratory factory, and migratory service and self-employment.

I adopted two-stage analysis in which I estimated job choice in the first stage and then estimate monthly earnings and annual income in the second stage. This is to address the selection bias issue which arises due to lack of income information for the workers in “self-employed farming, housewives, and inactive workers” category.

Job choice estimation adopted a multinomial logit model. The results confirmed that workers with less human capital, in terms of years of schooling and English skill, and less wealth, in terms of asset and owned land, face entry barriers in high-return nonfarm jobs. It was also evident that disadvantaged caste group workers had higher probability of working in casual and low-return jobs. This could be due to the high dependency on the personal network when lower caste workers find jobs. In order to further investigate whether this is the case, I estimated the use of formal institution as opposed to personal network in finding migratory jobs. The results indicated that the lower caste workers depend more on the personal network than other workers.

The results of monthly earnings estimation were consistent with the hypothesis that there is no systematic difference in earnings across social groups. This implies that even if one does not have a high probability of getting a high-return job, there is no difference in wage once the job is acquired. It also showed that human capital has

positive impact on earnings. Therefore, human capital is an important factor not only for acquisition of high-return jobs but also for higher earnings.

One major contribution of the chapter 3 is that the results from all the analyses suggest that factory job can be a potential job to effectively support the disadvantaged households to improve their livelihood in rural areas in eastern India. Firstly, the wage of factory workers was much higher than casual nonfarm jobs. Secondly, the human capital levels of factory workers were not significantly higher than that of casual nonfarm jobs. Thirdly, the lower social group workers were able to engage in factory jobs for a long term because it is not seasonal or casual, therefore could earn at least the same level of annual income as other workers.

Chapter 3 focused on education which was found to help people get a high-return job and earn more. Having achieved remarkable improvement in primary education attainment, India now faces new issues: low completion rate of primary school and low participation of secondary school due to high drop-out rates. The objective of chapter 3 was to identify the determinants of school progression at primary and secondary level education in Andhra Pradesh state.

Making use of unique dataset of a survey which follows the children of the same age cohort from 2003 to 2013, I conducted sequential logit analysis which accounts for the dynamic nature of school decision making process in which the

decision made at one time is conditional on the past decisions. The analysis paid close attention to the effects of job opportunities and the different trends in school progression by gender.

The school progression estimation was designed in such way that it could track the long-term effects of determinants. Among all the individual and household attributes, child's ability (Raven's test results at age 8 and reading and writing ability at age 12) presented long-lasting positive effects on school progression.

The results also revealed that higher wage of casual jobs had negative effects on children's school progression. The presence of factory, on the other hand, showed a positive effect on schooling progression. This implies that parents respond to job opportunities in their local area either by pulling their children out of school to work, or by keeping their children in school for longer years. The wealth explained little of children's school progression presumably because the direct and indirect cost of sending the children has lowered considerably.

The analysis was run separately for boys and girls. The wage of casual work had negative effects on schooling progression of both boys and girls. However, the effect was present at lower level of education for boys while it was only observed at higher level of education for girls. This may be due to the higher demand for young boys for agricultural wage work or construction work.

With regard to the expected roles for the eldest child, the results were consistent with the previous finding that the eldest sons had higher probability of progressing to secondary school, but not the eldest daughters. In order to investigate further how the schooling decision is made for girls, I analyzed the schooling and marital status of girls. The results of the multinomial logit estimation of the status revealed that the eldest daughters drop out of school not to work, which is different from previous findings that the eldest daughters drop out of school to marry early to reduce the resource crowding of her family.

4.2 Policy Implications and Concluding Remarks

India is the world's fastest growing major economy with an average growth rate of about 7 % over the last two decades. Most of this remarkable growth has been driven by the service sector which accounted for 66 % of the country's GDP in 2012 (GGDC, 2014). The rapid population growth has resulted in a labor market with abundant unskilled labor. This calls for large number of job generation to absorb these workers, but the recently growing modern service sector is dominantly skill-requiring and labor-saving. Therefore, the unskilled workers often have no choice but to work as low-return casual workers. This is one cause of persistent poverty which is prevalent in rural areas in particular.

This dissertation addressed this country-wide issue by examining how the rural households diversify their activities in four eastern states and how the households invest in schooling in Andhra Pradesh. The analysis in chapter 2 provided evidence of entry barriers in nonfarm labor market. The workers with higher level of human capital, wealth, and social status were found to have higher probability of getting a high-return job. One cause of the entry barriers was the high dependency of low caste workers on personal network in finding a job. Since there was no evidence of wage discrimination, it is the access to jobs that should be improved by interventions. Provision of alternative means of finding information on job opportunities such as recruitment firms and vocational training center could widen their access to high-return jobs. As it was the case in the experiment conducted by Jensen (2012), provision of recruiting services to newly emerging industry could support those who had no network in the past that could channel them into those jobs.

Chapter 2 also highlighted the factory job as a potential job that could effectively support the disadvantaged group that face barriers to enter into the high-return modern service sector. From the analysis results, it became evident that although the wage of factory job is much higher than other casual jobs, the human capital of factory workers is not significantly higher than that of casual workers, given other attributes. In addition, it allows the workers to engage in long-term job since factory job is not seasonal.

Chapter 3 contributes to literature on schooling by identifying the determinants of school progression in a dynamic framework. The findings that child's ability at younger age has long-term positive effects on school progression stress the importance of learning the basics in primary school. At primary school, the decline of learning level rather than participation is an issue today. Improving the school quality in terms of teacher's absenteeism, student-teacher ratio, learning materials could all contribute to improving the students' learning.

The findings on the effects of job opportunities in local areas offered important policy implications. Firstly, boys were more likely to drop out from lower primary school when the casual job wage is high, but the same effects were evident only above upper primary school for girls. The presence of factory in a locality, on the other hand, led to an increase in the probability of continuing the schooling to higher level, which implies that the higher expected returns to education contribute to keeping the children in school. Enforcement of The Right of Children to Free and Compulsory Education Act and The Child Labour (Prohibition and Regulation) is one action that can mitigate dropping-out. Also important is to create regular jobs which have higher expected returns to education than casual jobs, and at the same time, to ensure the parents are well aware of it.

The overall findings in this dissertation highlighted that jobs and education which are the core of people's living have influence over each other in various ways. Their effects are intertwined in such way that it cannot be independent of each other. Therefore, integrated approach combining the following two actions is proposed. First is the development of nonfarm sector, especially the manufacturing sector which can create large number of employment and is more accessible to the disadvantaged workers than modern service sector. Second is providing means to acquire more information about job opportunities. Both of these actions could lead to an increase the investment in education as well as access to high-return jobs, which are the foundation for human capital, and in turn, improve people's livelihood.

References

Agrawal, T. and Chandrasekhar, S. (2015). Labor Market Outcomes of Itinerant Workers in Rural India. *Journal of International Development,* Advance online publication.

Banerjee, B. and Knight, J.B. (1985). Caste Discrimination in the Indian Labor Market. *Journal of Development Economics,* 17(3), 277–307.

Barrett, C. B., Reardon, T., & Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications. Food policy, 26, 315-331.

Becker, G. (1957). *The economics of discrimination*. Chicago, IL: University of Chicago Press.

Birdsall, N. (1985). Public Inputs and Child Schooling in Brazil. *Journal of Development Economics,* 18(1), 67–86

Borooah, V. K. (2005). Caste, Inequality, and Poverty in India. *Review of Development Economics,* 9, 399–414.

Borooah, V. K., Dubey, A., and Iyer, S. (2007). The Effectiveness of Jobs Reservation: Caste, Religion and Economic Status in India. *Development and Change,* 38, 423–45.

Borooah, V. K., & Iyer, S. (2005). Vidya, Veda, and Varna: The influence of religion and caste on education in rural India. Journal of Development Studies, 41, 1369-1404.

Buis, Maarten L. (2011). The Consequences of Unobserved Heterogeneity in a Sequential Logit Model. *Research in Social Stratification and Mobility*, 29(3), 247-262.

Cain, J. S., Hasan, R., Magsombol, R., and Tandon, A. (2010). Accounting for Inequality in India: Evidence from Household Expenditures. *World Development,* 38(3), 282–297.

Cameron, S. V., & Heckman, J. J. (1998). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political economy*, 106(2), 262-333.

Chamarbagwala, R. (2008). Regional returns to education, child labor, and schooling in India. *Journal of Development Studies,* 44(2), 233–257.

Cherdchuchai, S. and Otsuka, K. (2006). Rural Income Dynamics and Poverty Reduction in Thai Villages from 1987 to 2004. *Agricultural Economics*, 35(Suppl), 409–423.

Corral, L. and Reardon, T. (2001). Rural Nonfarm Incomes in Nicaragua. *World Development*, 29(3), 427–442.

Davis, B., Winters, P., Reardon, T., and Stamoulis, K. (2009). Rural Nonfarm Employment and Farming: Household-Level Linkages. *Agricultural Economics*, 40, 119–123.

de Haas, H. (2010). The Internal Dynamics of Migration Processes: A Theoretical Inquiry. *Journal of Ethnic and Migration Studies*, 36(10), 1587–1617.

de Janvry, A. and Sadoulet, E. (2001). Income Strategies among Rural Households in Mexico: The Role of Off-Farm Activities. *World Development*, 29(3), 467–480.

Deininger, K., Jin, S., and Nagarajan, H. (2013). Wage Discrimination in India's Informal Labor Markets: Exploring the Impact of Caste and Gender. *Review of Development Economics*, 17(1), 130–147.

Desai, S. & Andrist, L. (2010) gender script and age at marriage in India, *Demography*, 47(3), 667-687.

Deshpande, A. (2000). A Look at Inequality in Kerala, India. *American Economic Review*. 90(2), 322–25.

Dostie, D., & Jayaraman, R. (2006). Determinants of school enrollment in Indian villages. *Economic Development and Cultural Change*, 54(2), 405–421.

Dre`ze, J., & Kingdon, G. G. (2001). School participation in rural India. *Review of Development Economics*, 5(1), 1–24.

Duraisamy, P., James, E., Lane, J., & Tan, J-P. (1998). Is there a quantity quality trade-off as enrollments increase? Evidence from Tamil Nadu, India. *International Journal of Educational Development*, 18(5), 367–383.

Duryea, S. & Arends-Kuenning, M. (2003). school attendance, child labor and local labor market fluctuations in urban Brazil. *World Development*, 31(7): 1165-1178.

Dutta, P. V. (2006). Returns to Education: New Evidence for India, 1983–1999. *Education Economics,* 14, 431–451.

Estudillo, J. P. and Otsuka, K. (1999). Green Revolution, Human Capital, and Off Farm Employment: Changing Sources of Income among Farm Households in Central

Luzon, 1966–94. *Economic Development and Cultural Change*, 47(3), 497–523.

Estudillo, J. P. and Otsuka, K. (2016). *Moving out of poverty: An inquiry into the inclusive growth in Asia*. Abington: Routledge.

Estudillo, J. P., Sawada, Y., Otsuka, K. (2008). Poverty and Income Dynamics in Philippine Villages, 1985–2004. *Review of Development Economics*, 12(4), 877–890.

Eswaran, M., Kotwal, A., Ramaswami, B., and Wadhwa, W. (2009). Sectoral Labor Flows and Agricultural Wages in India, 1983-2004: Has Growth Trickled Down? *Economic & Political Weekly*, 44(2), 46–55.

Foster, A. D., & Rosenzweig, M. R. (1996). Technical change and human capital returns and investments: Evidence from the Green Revolution. *American Economic Review*, 86(4), 931–953.

Fulford, S. (2014). Returns to Education in India. *World Development*, 59, 434-450.

GGDC (Groningen Growth and Development Centre). (2014). GGDC 10 sector database version 2014. Retrieved from http://www.rug.nl/research/ggdc/data/10-sector-database

Government of India. (2012). *An Anthology of Research Studies on the Mahatma Gandhi National Rural Employment Guarantee Act, 2005, 2006–2012*. New Delhi: Orient BlackSwan.

Government of India (2014a). *Report of the Expert Group to Review the Methodology for Measurement of Poverty*. New Delhi: Planning Commission.

Government of India (2014b). "Educational Statistics at a Glance." Ministry of Human Resource Development Bureau of Planning, Monitoring & Statistics. New Delhi.

Government of India (2016). *Percent of Population below Poverty Line – 1993-94 & 2004-05 and Poverty by Castes & Other Sub-Groups – 1983, 1993-94 & 2004-05*. Retrieved Jan 9, 2016, from http://planningcommission.gov.in/data/datatable/index.php?data=datatab

Government of India (2005). *Sources of Household Income in India 1999-2000*, National Sample Survey Organisation (NSSO), New Delhi: Ministry of Statistics and Programme Implementation.

Haggblade, S., Hazell, P., and Reardon, T. (Eds.). (2007). *Transforming the Rural Nonfarm Economy: Opportunities and Threats in the Developing World*. Baltimore, MD: The Johns Hopkins University Press.

Haggblade, S., Hazell, P., and Reardon, T. (2010). The Rural Nonfarm Economy: Prospects for Growth and Poverty Reduction. *World Development*, 38(10), 1429–1441.

Himanshu, Lanjouw, P., Murgai, R., and Stern, N. (2013). Nonfarm Diversification, Poverty, Economic Mobility, and Income Inequality: A Case Study in Village India. *Agricultural Economics*, 44,461–473.

Hoff, K. and Pandey, P. (2006). Discrimination, Social Identity, and Durable Inequalities. *American Economic Review*, 96, 206–211.

Huang, J., Wu, Y., and Rozelle, S. (2009). Moving off the Farm and Intensifying Agricultural Production in Shandong: A Case Study of Rural Labor Market Linkages in China. *Agricultural Economics* 40(2), 203–218.

Imai, K. S., Gaiha, R., and Thapa, G. (2015). Does Non-farm Sector Employment Reduce Rural Poverty and Vulnerability? Evidence from Vietnam and India. *Journal of Asian Economics*, 36(2), 47–61.

Indiastat (2016). Socio-Economic Statistical Information about India. Retrieved from http://www.indiastat.com/default.aspx

International Labor Organisation. (2013). *Global Employment Trends 2013: Recovering from a Second Jobs Dip*. Geneva: ILO.

Ito, T. (2009). Caste Discrimination and Transaction Costs in the Labor Market: Evidence from Rural North India. *Journal of Development Economics*, 88, 292–300.

Ito, T. and Kurosaki, T. (2007). Weather Risk, Wages in Kind, and the Off-Farm Labor Supply of Agricultural Households in a Developing Country. *American Journal of Agricultural Economics*, 91(3), 697-710.

Jacoby, H., & Skoufias, E. (1997). Risk, Financial Markets, and Human Capital in a Developing Country. *Review of Economics Studies*, 64(3), 311–335.

Jensen, R. (2012). Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India. *Quarterly Journal of Economics*, 127, 753-792.

Jensen, R. and Thornton, R. (2003). Early Female Marriage in the Developing World, *Gender & Development*, 11(2), 9-1

Jha, A. K. (2013). Climate Change and Internal Migration in India: Response of the State, Market, and the Civil Society. *Poverty & Public Policy*, 5(2), 133–145.

Kajisa, K. (2007). Personal Networks and Nonagricultural Employment: The Case of a Farming Village in the Philippines. *Economic Development and Cultural Change*, 55(4), 669–707.

Kajisa, K, and Planichamy, N. V. (2010). Schooling Investments over Three Decades in Rural Tamil Nadu, India: Changing Effects of Income, Gender, and Adult Family Members' Education. *World Development*, 38(3), 298-314.

Kijima, Y. (2006). Caste and tribe inequality: Evidence from India, 1983–1999. *Economic Development and Cultural Change*, *54*, 369–404.

Kijima, Y. and Lanjouw, P. (2005). Economic Diversification and Poverty in Rural India. *Indian Journal of Labor Economics*, 48(2), 349–374.

Kingdon, G. (2007). The progress of school education in India. *Oxford Review of Economic Policy*, 23(2), 168-195.

Kingdon, G. G. Kingdon, G. G. (1998). Does the labour market explain lower female schooling in India? *Journal of Development Studies*, 35(1), 39-65.

Kochar, A, (1999). Smoothing Consumption by Smoothing Income: Hours-of-Work Responses to Idiosyncratic Agricultural Shocks in Rural India. *Review of Economic Statistics*. 81, 50–61.

Kochar, A. (2004). Urban influences on rural schooling in India. *Journal of Development Economics*, 74: 113-136.

Kumar, G. (2013). Child labour: determinants, dimensions and policies in India. *Economic Affairs*, 58(4), 417-429.

Kumra, N. (2008). "An Assessment of the Young Lives Sampling Approach in Andhra Pradesh, India." Technical Note no. 2, Young Lives, Oxford.

Kung, J. and Lee, Y. (2001). So What if there is Income Inequality? The Distributive Consequence of Nonfarm Employment in Rural China. *Economic Development and Cultural Change*, 50(1), 19–46.

Kurosaki, T. and Khan, H. (2006). Human Capital, Productivity, and Stratification in Rural Pakistan. *Review of Development Economics*, 10(1), 116–134.

Lanjouw, J. and Lanjouw, P. (2001). The Rural Non-farm: Sector: Issues and Evidence from Developing Countries. *Agricultural Economics*, 26, 1–23.

Lanjouw, P. and Murgai, R. (2009). Poverty Decline, Agricultural Wages, and Nonfarm Employment in Rural India: 1983–2004. *Agricultural Economics*, 40, 243–263.

Lanjouw, P., Murgai, R., and Stern, N. (2013). Nonfarm diversification, poverty, economic mobility, and income inequality: A case study in village India. *Agricultural Economics, 44*, 461–473. doi: 10.1111/agec.12029

Lanjouw, P. and Shariff, A. (2004). Rural Non-Farm Employment in India: Access, Income and Poverty Impact. *Economic and Political Weekly*, 39(40), 4429–4446.

Lee, L. (1983). Generalized Econometric Models with Selectivity. *Econometrica*, 51(2), 507-512.

Liu, P.-W., Zhang, J., and Chong, S.-C. (2004). Occupational segregation and wage differentials between natives and immigrants: Evidence from Hong Kong. *Journal of Development Economics, 73*, 395-413. doi:10.1016/j.jdeveco.2002.12.003

Lucas, S. R. (2001). Effectively maintained inequality: Education transitions, track mobility, and social background effects. *American journal of sociology*, 106(6), 1642-1690.

Lucas, S. R., Fucella, P. N., & Berends, M. (2011). A neo-classical education transitions approach: A corrected tale for three cohorts. *Research in Social Stratification and Mobility*, 29(3), 263-285.

Madheswaran, S. and Attewell, P. (2007). Caste Discrimination in the Indian Urban Labor Market: Evidence from the National Sample Survey. *Economic and Political Weekly*, 42(41), 4146–4153.

Maertens, A. (2013). Social Norms and Aspirations: Age of Marriage and Education in Rural India. *World Development*, 47, 1-15.

Mano, Y., T. Yamano, A. Suzuki, and Matsumoto, T. (2011). Local and Personal Networks in Employment and the Development of Labor Markets: Evidence from the Cut Flower Industry in Ethiopia. *World Development*, 39 (10), 1760–1770.

Mare, R. D. (1980). Social background and school continuation decisions. *Journal of the American Statistical Association,* 75, 295-305.

Mare, R. D. (2011). Introduction to symposium on unmeasured heterogeneity in school transition models. *Research in social stratification and mobility*, 29(3), 239-245.

Matsumoto, T., Kijima, Y., and Yamano, T. (2006). The Role of Local Nonfarm Activities and Migration in Reducing Poverty: Evidence from Ethiopia, Kenya, and Uganda. *Agricultural Economics*, 35(suppl), 449–458.

Mehta, Arun C. (2007). *Elementary Education in India: Progress towards UEE; Analytical Report 2005–06.* National University of Educational Planning and Administration and Department of School Education and Literacy, Ministry of Human Resource Development, Government of India.

Micevska, M. and Rahut, D. B. (2008). Rural Nonfarm Employment and Incomes in the Himalayas. *Economic Development and Cultural Change*, 57(1), 163–193.

Ministry of Rural Development, Government of India. (2012). *An Anthology of Research Studies on the Mahatma Gandhi National Rural Employment Guarantee Act, 2005, 2006–2012*. New Delhi: Orient BlackSwan.

Mohanty, M. (2006). Social inequality, labour market dynamics and reservation. *Economic and Political Weekly*, *41*, 3777–3789. Retrieved from http://www.epw.in/journal/2006/35/special-articles/social-inequality-labour-market-dynamics-and-reservation.html

Mosse, D., Gupta, S., Mehta, M., Shah, V., Rees, J., and the KRIBP Project Team (2002). Brokered Livelihoods: Debt, Labor Migration and Development in Tribal Western India. *Journal of Development Studies*, 38(5), 59–88.

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

Oster, E. and Steinberg, B. M. (2013). "Do IT Service Centers Promote School Enrollment? Evidence from India." *Journal of Development Economics*, 104, 123-135.

Osborne, E. (2001). Culture, development, and government: Reservations in India. Economic Development and Cultural Change, 49, 659-685. doi:10.1086/452520

Otsuka, K., Estudillo, P. J., and Sawada, Y. (2008). *Rural poverty and income dynamics in Asia and Africa.* Abingdon: Routledge.

Pal, S. (2004). Child Schooling in Peru: Evidence from a Sequential Analysis of School Progression. *Journal of Population Economics*, 17(4), 657–680.

Patrinos H.A. and Psacharopoulos G (1997). Family Size, Schooling and Child Labor in Peru – An Empirical Analysis. *Journal of Population Economics,* 10(4):387–405.

Pritchett, L., & Beatty, A. (2012). The negative consequences of overambitious curricula in developing countries. Center for Global Development Working s, 293.

Psacharopoulos, C. and Patrinos, H.A. (2004) Returns to Investment in Education: A Further Update. *Education Economics*, 12(2), 111-13.

Psacharopoulos, C. & Patrinos, H.A. (2004). Returns to investment in education: A Further Update. *Education Economics*, 12:2, 111-13.

Rao, V. (1993). The rising price of husbands: A hedonic analysis of dowry increase in rural India. *Journal of Political Economy*, 101(4), 666-677.

Rangaraju, B., Tooley, J., & Dixon, P. (2012). The private school revolution in Bihar: Findings from a survey in Patna urban. New Delhi, India Institute.

Reardon, T., Taylor, J., Stamoulis, K., Lanjouw, P., and Balisacan, A. (2000). Effects of Non-Farm Employment on Rural Income Inequality in Developing Countries:

An Investment Perspective. *Journal of Agricultural Economics,* 51(2), 266–288.

Rosenzweig, M. & Stark, O. (1989). Consumption smoothing, migration and marriage: evidence from rural India. *Journal of Political Economy*, 97(4), 905-926.

Sabates, R., Hossain, A., & Lewin, K. M. (2013). School Drop-out in Bangladesh: Insights using Panel Data. *International Journal of Educational Development*, 33(3), 225-232.

Saha, B. and Bahal, R. (2015). Factors Leading to Success in Diversified Occupation: A Livelihood Analysis in India. *Journal of Agricultural Education and Extension*, 21(3), 249–266.

Sawada, Y. and Lokshin, M. (2009). "Obstacles to School Progression in Rural Pakistan: An Analysis of Gender a Sibling Rivalry using Field Survey Data." *Journal of Development Economics*, 88, 335-347.

Senaratna Sellamuttu, S., Aida, T., Kasahara, R., Sawada, Y., & Wijerathna, D. (2014). How access to irrigation influences poverty and livelihoods: A case study from Sri Lanka. *Journal of Development Studies, 50*, 748-768.

Shafiq, M. N. (2007). Household schooling and child labor decisions in rural Bangladesh. *Journal of Asian Economics*, 18: 946-966.

Shastry, Gauri Kartini, (2012). Human capital response to globalization: education and information technology in India. *Journal of Human Resources*, 47 (2), 287–330.

Siddhu, G. (2011). "Who Makes It To Secondary School? Determinants of Transition to Secondary Schools in Rural India." *International Journal of Educational Development*, 31(4), 394-401.

Singh, R., & Bangay, C. (2014). Low fee private schooling in India–More questions than answers? Observations from the Young Lives longitudinal research in Andhra Pradesh. International Journal of Educational Development, 39, 132-140.

Singh, A. and Park, A., and Dercon, S. (2014). School Meals as a Safety Net: An Evaluation of the Midday Meal Scheme in India. *Economic Development and Cultural Change*, 62(2), 275-306.

Takahashi, K. and Otsuka, K. (2009). The Increasing Importance of Nonfarm Income and the Changing Use of Labor and Capital in Rice Farming: The Case of Central Luzon, 1979–2003. *Agricultural Economics*, 40, 231–242.

World Bank (2012) *World Development Report 2013: Jobs*. Washington DC: World Bank.

World Bank (2011) *World Development Report 2012: Gender Equality and Development*. Washington DC: World Bank.

Yunez-Naude, A. and Taylor, J. (2001). The Determinants of Nonfarm Activities and Income of Rural Households in Mexico, with Emphasis on Education. *World Development*, 29(3), 561–572