TECHNICAL AND VOCATIONAL EDUCATION AND TRAINING IN INDIA: A STUDY OF CHOICE AND RETURNS

Namrata Tognatta

A DISSERTATION

in

Education

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2014

Supervisor of Dissertation:

_____

Rebecca A. Maynard, University Trustee Professor of Education and Social Policy

Graduate Group Chairperson:

_____

Stanton E. F. Wortham, Professor of Education

Dissertation Committee:

Rebecca A. Maynard, University Trustee Professor of Education and Social Policy

Robert Boruch, University Trustee Chair Professor of Education and Statistics

Devesh Kapur, Associate Professor of Political Science

Subha Mani, Assistant Professor of Economics

TECHNICAL AND VOCATIONAL EDUCATION AND TRAINING IN INDIA:

A STUDY OF CHOICE AND RETURNS

DEDICATION

To my parents, whose hard work and commitment inspires me to do better.

To my sister and role model.

To Jonathan.

ACKNOWLEDGEMENTS

This dissertation was possible because of the guidance and support provided by many over the past 5 years and I am indebted to them all. My committee has been instrumental in my development as a research scholar through the process of this dissertation and my time as a Ph.D. student. My deepest gratitude to Becka for giving me the opportunity to pursue my doctoral ambitions and for being a role model to those of us interested in the pursuit of reliable evidence. You have always been an inspiration and I have treasured and valued your mentorship.

I had the incredible good fortune of working with Devesh and Subha during my doctoral training. Your insight, feedback and unwavering support has been invaluable. You have encouraged and challenged me to be a better scholar and a better methodologist. Devesh, my association with you and CASI has influenced my development as an India researcher and enriched my doctoral training experience. Subha, working with you has been a tremendous learning experience. I aspire to the kind of dedication and expertise you bring to your work and I deeply value our research partnership and friendship.

My sincerest gratitude to Bob for never letting me lose sight of the big picture, for always being available with advice and encouragement, and being a continuous source of confidence.

I would also like to thank Professor Jon Supovitz, Henry May and Philip Sirinides at CPRE for giving me various research opportunities that allowed me to hone my research skills and grow as an applied researcher. Henry and Phil, I am grateful for the time you invested in training me and responding to numerous methodological questions.

My friends and colleagues at Penn have been a tremendous source of support. I cherish our friendship and the positive energy you always bring in to my life.

Finally, this research would not have been possible without the financial support provided by Toby Linden and Michelle Neuman at the World Bank, Sunanda Mane at Lend-a-Hand-India, the TRIPS team in Bombay and the Education Policy Division at Penn GSE. Thank you for enabling me to see this research endeavor to fruition.

The research reported here was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant #R305B090015 to the University of Pennsylvania. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

ABSTRACT

TECHNICAL AND VOCATIONAL EDUCATION AND TRAINING IN INDIA: A STUDY OF CHOICE AND RETURNS

Namrata Tognatta

Rebecca A. Maynard

India has made remarkable progress and achieved near universal enrollment in primary school education. However, the quality of learning and progress beyond primary education are of concern; nearly 50 percent of fifth graders are unable to read second grade material and retention rates at the secondary level are quite low. The higher education sector has also shown impressive growth but faces several challenges around inequitable access and low quality. Low outcomes at the secondary and higher education levels have resulted in a significant deficit in employable and vocationally trained individuals in the workforce. Evidence shows that just 14 percent of new entrants to the workforce are likely to have a college or graduate degree. Research also shows that over the long-term low outcomes at the secondary and postsecondary levels are likely to translate into low lifetime earnings and well-being.

In light of low educational and employment outcomes, policy in India has focused on skill development through the technical and vocational education and training (TVET) sector. The primary objective of these policies is to significantly improve the rate at which youth and young adults participate in these programs. However, there is limited research evidence on TVET in India.

This dissertation addresses the need for empirical evidence on TVET to enable the policy dialogue on meeting the country's education and training challenges. Specifically, it examines the role of individual, household and macro-level factors in human capital investment decisions, especially as those might relate to participation in vocational education and training. Since the expected returns to education and training are a key determinant of investment decisions, the dissertation examines the economic returns to vocational education and training in India. Finally, the dissertation examines the impact of secondary-level vocational education on high school completion rates and postsecondary enrollment among participants.

Large-scale secondary and primary data are used in empirical models to address the questions posed above. The findings thus generated present reliable, generalizable estimates that have the potential to inform the future direction of policy in vocational education and training in India. The findings also identify groups differentially affected by current policies and can thereby be used to address inequitable access to and stratification in education and training programs in India.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Technical and vocational education and training (TVET) issues have received much attention this past decade and TVET topics have been the focus at global forums organized by the United Nations Educational, Scientific, and Cultural Organization (UNESCO), the Organisation for Economic and Cultural Development (OECD), and the International Labour Organization (ILO)[1]. Major world reports related to TVET have been released to document these discussions on the future direction of the vocational education sector.[2] While TVET discussions in OECD countries have covered various topics ranging from shortages of skilled workers (Australia, Portugal, Spain), retention and completion rates at the secondary level (U.S., England, Denmark), to regional imbalances in development (Germany and Korea) (Grubb, 2006), in emerging and less-developed countries TVET discussions have focused on improving economic growth and competitiveness, and addressing issues around social exclusion and equity (Psacharopoulos, 1997).

In developing countries specifically, the recent rounds of debate around TVET are driven by concerns around the supply and demand of labor (World Bank, 2013). The imbalance in the supply and demand of labor has been attributed to massive demographic shifts ("youth bulges") (World Bank, 2013), the changing nature of work and technological innovations (Grubb, 2006), low secondary education outcomes, especially

[1] *Third International Congress* on TVET organized by UNESCO in 2012, *Global Dialogue Forum on Vocational Education and Training* organized by ILO in 2010.
[2] The *World Development Report on Jobs* (2013); *EFA Global Monitoring Report* (2012) on 'Youth and Skills'; *OECD Reviews of Vocational Education and Training – Learning for Jobs* (2010); *Technical and Vocational Education and Training for the Twenty-First Century – UNESCO and ILO Recommendations* (2001).

among females (World Bank, 2012), poor flow of information between employers and job seekers, and a mismatch between skills, aspirations and labor market needs (Aggarwal et al., 2011; World Bank, 2012). While reforms in the TVET sector are not the only identified solutions to correct labor market imbalances[3], they have been in the spotlight in several developing countries (World Bank, 2012; 2013) and guide the focus of this dissertation.

India currently faces several of the education and labor challenges described above. Nearly 50 percent of fifth graders in India are unable to read second grade material, and the dropout rate at the secondary school level is nearly 30 percent (Kingdon, 2007). Further, only a small proportion of labor force entrants (14 percent) are likely to have a college degree or some vocational training (Confederation of Indian Industries, 2009). In response, policymakers have focused on expanding skills training opportunities at the secondary and postsecondary level[4]. Even though TVET at the secondary school level has not been popular in India (Tilak, 2002), one of the aims of a recent secondary school-level reform, the *Rashtriya Madhyamik Shiksha Abhiyan* (Government of India, 2009a; RMSA)[5], is to attract and retain students in secondary school by introducing vocational content at the secondary level. Similarly, the recent *National Skill Development Policy* (Government of India, 2009b), targets expanding TVET opportunities through public-private partnerships and aims to train 500 million people over the next 10 years.

[3] See the World Bank (2013) report on *Jobs* for a detailed discussion on this topic.
[4] Other policy instruments, not discussed here, target growth in the manufacturing sector.
[5] The 2009 RMSA policy targets improvements in secondary education in India.

There has been relatively little academic debate and research on TVET policy and practice in developing countries. The bulk of available research pertains to OECD countries. The research from developing countries is scant and what is available tends to be narrowly focused on employability of TVET graduates. Moreover, existing studies do not articulate an explicit theory of action that explains how a vocational program should work and the impact it should have (Grubb, 2006). In developed and developing countries alike, the research has tended to ignore issues of who is served by TVET programs and whether reforms reach the target groups that they purport to serve.

In light of the current expansions envisioned for TVET in India, some critical questions must be raised. What factors motivate participation in TVET? What are the economic returns to TVET for the individual and the household? Does participation in TVET in secondary school improve future educational and labor market outcomes? There has been no published research from India that has adequately addressed these questions. Further, the evidence from other developing countries has been largely missing in the case of determinants of participation or ambiguous in the case of TVET returns and impact of secondary TVET[6].

There are several reasons to advance our understanding of how individuals make decisions regarding participation in vocational programs, including the types of programs they chose and the returns they expect from participating in these programs. First, a recurring topic in policy discussions concerns the types of education and training opportunities that must be provided to best meet the needs of society and individuals. Individuals make decisions regarding accessing education and training programs from the

[6] Evidence from extant research is discussed in Chapter 3.

secondary stage and beyond. Understanding this decision-making process around human-capital investments, and the kinds of information and resources that are used in order to make these decisions is valuable for effective policy and program formulation.

Second, it would also be useful to gain an understanding of the factors that mediate or moderate the human capital investment decision-making process of individuals and families. This would be especially helpful in identifying circumstances that lead to inequitable access or differentially affect certain groups.

Third, most discussions around vocational education are focused on whether the sector is *responsive* to the needs of stakeholders. The issues extend beyond those related to manpower forecasting, institutional policies and supply-side activities to how vocational education is perceived and used by the population (Psacharopoulos, 1988).

Fourth, the TVET sector in India is a complex system offering a wide array of educational and training options for individuals at different levels of educational attainment.[7] There is significant variation not only in the types of programs offered (broadly, TVET programs can be classified as "formal" or "informal"), but also in the proportion of participants and profiles of participants across types of TVET programs. While "formal" TVET programs in India have received some research attention, little is known about "informal" TVET and the participants who access these programs.

This dissertation begins to address some of the gaps in TVET research in India using multiple secondary data sources, including nationally representative surveys, as well as primary data collected from one state in India. This dissertation poses three broad questions –

[7] Chapter 2 provides a detailed description of the TVET sector in India.

1. What are the determinants of TVET participation in India?
2. What are the individual economic returns to TVET in India?
3. What is the impact of TVET in secondary schools on school completion and further enrollment?

The findings from this empirical analysis have the potential to provide evidence based on which future TVET policy can be formulated. The evidence also has the potential to inform the development of a more nuanced approach towards the evaluation of these policies in the future.

Chapter 2: Structure of TVET in India

This section presents a brief overview of the structure of TVET in India drawing from the descriptions provided in Agrawal (2012), Sharma (2010), and the World Bank (2006). The role of TVET in India is also briefly discussed.

The structure of TVET in India is complex, as is the case in most of the world. About 17 different ministries within the government provide and finance various TVET programs. Although the bulk of TVET provisions fall under the purview of the education and labor departments (Agrawal, 2012), since TVET is a "concurrent"[8] subject, the centre and states share responsibility for provision of TVET in the country (Sharma, 2010). The terms 'vocational education' and 'vocational training' refer to two distinct strands of TVET in India, but are often used interchangeably. Vocational education programs are offered as part of the formal education cycle whereas vocational training programs fall outside of the formal school cycle (Agrawal, 2012).

At the secondary school level, TVET is managed by the Ministry of Human Resource Development ([MoHRD] or, the Education Department) and governed by the scheme[9] on the 'Vocationalization of Secondary Education', which was introduced in 1987. As part of this scheme, students can opt for a vocational curriculum in grades 9 to 12 at any of 6,500 public secondary schools offering vocational options. The range of vocational courses offered as part of this scheme includes disciplines like agriculture, health and home sciences, education and technology, and business and commerce

[8] As per the Constitution of India, the concurrent list is concerned with relations between the union and the states, and includes items like education, criminal law, economic and social planning, and so on.
[9] Centrally sponsored schemes (CSS) or 'schemes' are special fiscal transfers from the central government to state or local governments.

(Sharma, 2010). Students going through the formal vocational education system at the secondary school level can continue their education in the general education system or access vocational training options available at the postsecondary level (like polytechnics, also managed by the education ministry, and offering diploma-level programs in engineering and technology trades) (Agrawal, 2012).

The TVET programs managed by the Ministry of Labor in India are classified as 'vocational training'. These options include the 'Craftsmen Training Scheme' (CTS) and the 'Apprenticeship Training Scheme' (ATS) and are outside of the formal schooling cycle (Sharma, 2010; World Bank, 2006).

The CTS was designed to equip youth with skills for productive employment and ensure the needs of the labor market were being met with a steady flow of skilled industrial workers (Sharma, 2010). The 'Industrial Training Institutes' (ITIs) were set up as part of this scheme and offer certificate-level courses in about 115 trades. The ITIs have relatively flexible entry requirements – students can enroll upon completion of 8 grades of schooling as well as after graduating high school. This flexibility makes ITIs accessible to secondary school leavers as well as completers. The duration of the programs offered ranges from three months to about three years. Similar programs are offered at private institutions called Industrial Training Centres (ITCs). In total, there are about 6000 ITIs and ITCs currently operating in India.

Through the ATS, industries or establishments offer apprenticeships in about 140 trades covering agriculture, engineering, health and paramedical, home science, and so on. Like the ITIs, these programs also have flexible entry criteria making them accessible to school leavers. The ATS is managed by both, the education and labor departments

(Sharma, 2010; World Bank, 2006). Depending on the trade and the level of prior education and training of the student, it can take between 4 months to 4 years to gain various levels of certification in a selected trade.

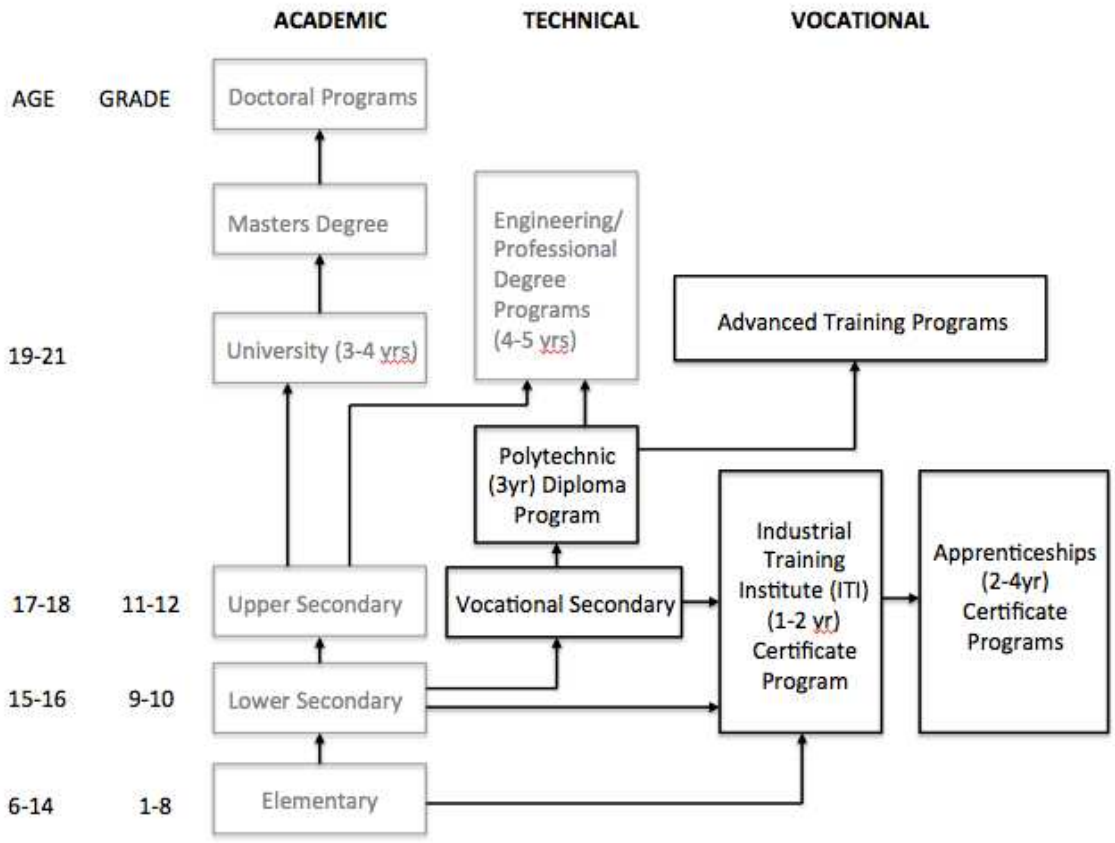

Figure 1. The TVET system in India (Adapted from World Bank, 2006)

Besides the formal structure of TVET described above, India also has a large private and informal network through which TVET is provided. The private, informal providers include non-government organizations (NGOs), community polytechnics, adult education centers, and establishments providing informal apprenticeships. These programs primarily offer relatively short-term training opportunities to informal sector workers (Sharma, 2010). The absence of any systematic documentation or research on

TVET provisions outside of the formal offerings makes the informal network somewhat of a black box.

2.1 Challenges facing TVET in India

The expansion of the TVET sector in India is a response to various educational and employment challenges facing the country. The context within which TVET operates is described below. Some of the challenges facing TVET that come in the way of fulfilling its objectives are also discussed.

While elementary education in India is nearly universal, the country faces major challenges at the secondary level (Planning Commission, 2013). Low participation rates and high dropout rates at this level result in high proportions of youth and young adults lacking the skills to successfully compete in the labor market. The universalization of elementary education has contributed to the expansion of the secondary and tertiary education systems to accommodate larger numbers of students continuing their education beyond the primary grades. The lack of education and skills required for gainful employment in formal sectors of the economy, coupled with declining employment opportunities in rural areas, has contributed to high levels of urban migration and rising numbers of youth seeking jobs in the unorganized or 'informal' sector of the economy, which currently employs nearly 90% of all workers.

The TVET system is considered a policy lever designed to improve equity and reduce unemployment rates especially among youth, balance the demand for higher education, provide skills to keep up with changes in technology, and build a knowledge economy. But the TVET system faces several challenges and is failing on many of these

counts (King, 2012). The literature cites several social, economic and political factors that create challenges for the TVET sector. These are related to perception and status issues, a mismatch between demand and supply, low quality of TVET programs and employability of TVET graduates, and mismanagement of the sector (ILO, 2003; World Bank, 2006).

That TVET is associated with low-status manual work and low-paying jobs in India is often cited as a reason for low participation rates in TVET (Tilak, 2002). In a survey of high school students in three districts of India, Aggarwal, Kapur & Tognatta (2011) found that students, irrespective of their academic achievement, aspire to careers in technology, medicine, finance and education, and are less interested in occupations traditionally targeted by TVET programs. Students and youth are interested in disciplines that are traditionally viewed as high status.

Reports examining the effectiveness and efficiency of TVET programs conclude that most programs offered at TVET institutions are irrelevant to the current needs of the economy. Further, the lack of financing, resources, and networks with industries and employers translate into outdated curricula and training programs, that produce unemployable graduates (ILO, 2003; World Bank, 2008).

Finally, the fragmented management system adopted for the TVET sector and lack of coordination between national-level and state-level bodies, leads to duplication of functions, diverse accountability, and a narrowing of roles and responsibilities. As a result, there is a preoccupation with all aspects of financing while more substantive functions related to upgradation and monitoring and evaluation of programs have been ignored (World Bank, 2006).

While TVET programs in India and other countries are viewed as a “second class” option for education and training, the lack of structural and financial resources for the sector has prevented any change in this perception through the improvement of TVET outcomes. But, the tendency of policymakers to use TVET as a catchall solution to educational and labor market problems has kept it alive as a policy tool.

Chapter 3: Review of Literature

This chapter discusses the major theories grounding the research in education and training decision-making. It also reviews the evidence from research on TVET and education, in general, highlighting key indicators identified to influence TVET participation and returns. The chapter begins with a brief discussion on what is meant by 'technical and vocational education and training' for the purposes of this dissertation.

3.1 Definition: Vocational Education/Training

Vocational education and training goes by various names, such as career and technical education, technical education, vocational education/training, skill development, and technical and vocational education and training. Across advanced and developing economies, vocational education and/or training programs are offered at various types of institutions, including schools, colleges, public and private vocational institutions, on the job, and at informal settings like the home or community (Grubb & Sweet, 2004; Karmel, 2011; Chappell 2003). Moreover, they are offered at various levels within the education system. The United Nations Institute of Statistics ([UN-UIS]; 2006) has identified students at four different levels of the International Standard Classification of Education – from level 2, which corresponds to lower secondary education, up to level 5, which corresponds to the first cycle of higher education.

In its 'Revised Recommendations for Technical and Vocational Education and Training', UNESCO (2001) provides a definition for vocational education and training

that reflects the shifts over time in thinking about what constitutes vocational activities. The shift has been from a view of vocational education quite narrowly in terms of preparing individuals for a particular job or occupation to a vision of it as a strategy for addressing various educational, economic, and social objectives. 'Technical and Vocational Education and Training' (TVET)[10] is defined as *"a comprehensive term referring to those aspects of the educational process involving, in addition to general education, the study of technologies and related sciences, and the acquisition of practical skills, attitudes, understanding and knowledge relating to occupations in various sectors of economic and social life"* (UNESCO, 2001). As such, TVET includes all activities undertaken at various stages, from secondary to postsecondary and on-the-job training. This dissertation focuses on TVET activities at the secondary and postsecondary level, regardless of the type of institution providing the training.

3.2 Theoretical Frameworks

Most theoretical models of investments in education and training have been conceptualized within an economic or sociological framework or a combination of the two. Economic models, and the human capital model in particular (Becker, 1962; Schultz, 1961), have been applied to research on educational decision-making since the human capital theory was first proposed in the 1960s. The human capital model posits that individuals (or households) make *rational* choices regarding investments in education and training with the ultimate goal of balancing direct costs and foregone earnings against the benefits that will be accrued from the education/training. These

[10] I follow the UNESCO convention and use 'TVET' to refer to vocational education and/or training.

models assume that information regarding (perceived) wages is especially important, but that nonmonetary factors are also important (Becker, 1993). This suggests that, other things equal, the demand for education will be stronger when benefits are expected to accrue over a longer period, and when the discount rate is relatively low. The economic model also recognizes the role of individual ability and individual/family preferences in investment decisions (Becker, 1993).

Human capital theory has three weaknesses. One is that it overlooks the fact that individuals often have imperfect or incomplete information about the value of education and training. Second, human capital investment decisions are often based on information other than monetary rewards, such as information on the health of the labor market and prospects for different types of education (Borghans et al., 1996). Finally, the human capital model fails to explain how students gather information regarding wages, the prospects associated with different types of education and training options, and how they develop different preferences.

While economists have addressed the first two concerns regarding imperfect information and the exclusion of labor market forecasts by including measures of wage or enrollment elasticity in their models (Borghans et al., 1996), the third concern has been largely ignored.

The sociological literature fills in some of these gaps in the human capital model and conceptualizes education decisions within a status attainment framework (Perna, 2006). Educational aspirations (based on demographic characteristics and academic achievement) are seen as influencing human capital investments (Hossler et al., 1999). More recent literature, such as that reviewed by Dika and Singh (2002), draws heavily on

the work of Bourdieu (1986) and Coleman (1988) to explain differences in educational attainment. Dika and Singh (2002) posit that social capital enables individuals to access resources through social networks and relationships and "build capital", while cultural capital is more indicative of class status and attributes such as cultural knowledge, language skills, artistic and literary pursuits, and so on. These forms of capital are hypothesized to create norms and standards that encourage educational attainment, engagement and achievement and are instrumental in developing human capital (Coleman, 1988). Further, the social context along with *habitus*, an internalized set of dispositions and preferences, contributes to an individual's attitudes, expectations and aspirations (McDonough, 1997) and, together, the social context and habitus determine an individual's options (Horvat, 2001).

Researchers have used a variety of measures of social and cultural capital to study education and training decisions. For example, these have included, measures of family structure, parent-child interactions, parents' involvement in schools, parents' expectations, parents' education, and intergenerational closure (Dika & Singh, 2002). In addition, school and community characteristics have been found to influence enrollment decisions and are included as indicators of structural context (McDonough, 1997; Perna & Titus, 2005).

Perna's (2006) criticism of the sociological models is that, while they clarify how students and families gather information (and explain group differences in information accumulation), they fail to clarify how this information influences decisions. Perna (2006) combines elements of the economic and sociological tradition in her theoretical framework of college access. This model assumes that economic utility maximization is

influenced by several layers of context within which it nested. In the context of TVET, the model would posit that individual demand for training is influenced by perceived costs and benefits, which in turn is affected by the individual's real and perceived ability, preferences, and degree of risk-aversion. These in turn are influenced by four contextual layers; (1) habitus or internalized mores, (2) the school and community context, (3) the higher education context, and (4) the general social, economic and policy context. Thus, the variation in enrollment decisions is examined as a function of the resources used or available to students during the decision-making process.

3.3 Determinants of Participation in TVET

The variables found to be important in explaining individual demand for TVET are classified as demand-side, or supply-side factors. The demand-side variables include those related to characteristics of the individual and household, and the supply-side variables are those that measure costs, benefits, institutional characteristics, and labor market indicators hypothesized to influence demand for TVET. A discussion of how the influence of these factors varies by demographic dimensions (age, gender, ethnicity, and urbanicity) is also included.

3.3.1 Academic achievement

That students who tend enroll in TVET are lower achieving, on average, has popular consensus and has been used to describe TVET participants in developing and developed countries (Agodini et al., 2004; Agrawal, 2012; Rothman, 2008). This is a

logical inference given the relatively low eligibility requirements and status accorded to TVET options. However, there is limited empirical evidence showing that academic achievement or ability *influences* TVET participation. Findings from studies that do examine the influence of academic achievement on decisions to enroll in TVET are ambiguous and vary by context and type of TVET.

In a study conducted by Mathematica (Agodini et al., 2004) in the U.S., findings showed that students with lower academic achievement (and low educational aspirations) were more likely to enroll in high school TVET than otherwise identical students. The study also found that controlling for academic achievement, participation rates were similar for African American and White students, while Hispanics were less likely to participate.

But in studies outside the U.S., contrary findings have been reported. Aypay (2003) compared the determinants of enrollment in secondary academic schools versus secondary vocational schools amongst a convenience sample of 873 students[11] and found that students with higher academic achievement (measured as prior GPA) were more likely to enroll in vocational schools than in general academic schools. Although the bias in the sample due to nonrandom selection and a high nonresponse rate raise some questions about the trustworthiness of his findings, similar results were reported in the case of Thailand (Moenjak & Worswick, 2003). This study used nationally representative data to examine factors related to participation within an econometric framework. Using a probit choice model, the authors found that academic achievement was positively and

[11] Surveys were distributed to 2100 students, yielding a response rate of about 41%.

significantly related to upper secondary TVET enrollments for males (but not for females), controlling for other household and regional characteristics.

While the association between achievement and postsecondary enrollment is by and large positive and significant at the postsecondary level, it varies by type of TVET. In Australia, TVET options at the postsecondary level include traineeships, apprenticeships, and TVET programs offered by public and private institutions. The latter offer a wide range of TVET options corresponding to various levels of certification from lower level certificates to advanced diplomas (Curtis, 2008). A study of these programs shows that students of lower academic ability (measured by skills in literacy and numeracy) are more likely to enroll in apprenticeships, traineeships and programs offering lower level certificates. But entry into TVET programs offering higher level certificates is associated with students of higher ability and aspirations (Ainley, 2005; Curtis, 2008).

These findings suggest that the role of educational attainment as a determinant of TVET is more complex at the secondary level than at the postsecondary level, and should be examined in relation to other contextual and economic indicators.

3.3.2 Household income

Most studies looking at the relationship between household income, educational pursuits, and labor market outcomes have found household income to exert a positive, although small, influence on enrollment decisions (Behrman & Knowles, 1997; 1999; Behrman et al., 1994; Duraisamy, 2002; Psacharopoulos, 1989). However, the true effect of household income on TVET enrollments has been difficult to isolate and studies show

ambiguous results (Foley, 2007; Perna & Titus, 2005; Sandefur et al., 2005; Teese & Walstab, 2008). Thus, although household income is an important demand-side determinant, it must be examined carefully.

There are several challenges in establishing causal relationships between family income and various educational outcomes including enrollment. In their review of over 40 studies, Behrman & Knowles (1999) noted that the main issues are endogeneity and multicollinearity. Because household income is correlated with unobservables such as parents' preferences towards human capital investments, OLS estimates of household income are likely to be biased (Mani et al., 2009). Behrman and Knowles (1999) find that most studies examining the effect of household income on human capital investments also include other household characteristics (parents' education, school characteristics, and so on) in the model. Since these variables are likely to be correlated with household income, the estimates on income could again be biased downward. As a result, some studies have used instrumental variables in an effort to address the endogeneity of the income variable. In most cases, these studies confirm that the OLS estimates for income are downward biased (Glewwe & Jacoby, 1995; Pal, 2004; Chaudhury et al., 2006).

Amongst the TVET studies reviewed, Sandefur et al. (2005) used a sociological framework to examine the influence of family resources, specifically parental education and family income, and aspects of social capital as determinants of enrollment in certificate courses, 2-year college, and 4-year college in the United States. The social capital indicators included family structure, number of siblings, parent expectations, parent-child discussions regarding school activities, intergenerational closure, and Catholic school attendance. Results showed that students from high-income households

have a higher probability of enrolling in 4-year college and a lower probability (although positive and significant) of enrolling in certificate programs and 2-year colleges. The study also found that the effect of household income diminishes when social capital indicators are included in the model. Similar evidence was found by Perna & Titus (2005) in their examination of 2-year and 4-year college enrollment. The coefficient on household income was positive and significant for 2-year college enrollments.

3.3.3 Parents' education

Parents' education is consistently identified in the literature as an important predictor of human capital investment decisions (Behrman & Wolfe, 1987; Birdsall, 1982; Lillard & Willis, 1994; Tansel, 2002). Further, maternal and paternal education appears to have slightly different effects on the education and training decisions for boys and girls (Behrman, 1999; Birdsall, 1982; Dostie & Jayaraman, 2006). The findings from these studies are mostly consistent with each other and show that father's education positively influences enrollment decisions of both, boys and girls, while the education of the mother has a stronger positive influence on educational attainment of girls in the household. These differences have been explained on the basis of bargaining models (Kambhampati & Pal, 2001) that argue that male and female heads have different utilities, and budget constraints, and thus make different decisions (Hoddinott, 1992).

The role of parents' education specifically with regard to TVET enrollments at the secondary level has not received much attention. One reason may be that the role of parents or household factors diminishes at the postsecondary level in general. Nonetheless, the few studies that have examined the relationship have reported positive

linear relationships between parents' education and TVET participation (Curtis, 2008; Fullarton, 2001; Moenjak & Worswick, 2003). However, Fullarton's (2001) examination of TVET demand in Australia found that as parents' education increases, students are less likely to enroll in secondary-level TVET.

3.3.4 Social and Cultural capital

Social capital indicators are commonly included in models of educational outcomes (Dika & Singh's 2002) but not specifically in TVET research. The former studies typically show that social capital indicators are positively linked to enrollment in education and training (Aypay, 2003; Sandefur et al., 2005; Perna & Titus, 2005). However, Dika and Singh (2002) also raise conceptual and methodological issues that are important to consider when interpreting these findings.

Of the TVET studies that examined the impact of social capital on TVET decisions, Aypay (2003) found that parent-child discussions about school were positively related to enrollment in academic schools and negatively related to enrollment in vocational schools; and parent guidance was negatively related to enrollment in both types of schools.

Sandefur et al. (2005) found slightly different results. They modeled social capital indicators *inside* the family (family structure, number of siblings and parental expectations) and those *outside* the family (school changes, intergenerational closure, parental involvement in school activities and parent-school contact about academic matters). Results showed that after controlling for parents' education and income and students' prior achievement, parent expectations, parent-child discussions, and parent-

school involvement improved the probability of TVET (as well as 4-year college) enrollments.

Finally, Perna & Titus's (2005) study examined differential access to social networks across ethnic and income groups. Their results suggest that social capital indicators are not only positively associated with either 2- or 4-year college enrollment, but that the relationship between social capital indicators and enrollment is different for African American and other youth. Measures of parent-student discussions were less predictive of college enrollment among African-Americans than non African American students, but measures of parent-school relationships were more predictive for African-Americans than non African Americans. The study also found a strong significant relationship between the volume of resources accessed via social networks at the school.

3.3.5 Costs and benefits

According to the human capital theory (Becker, 1993; Schultz, 1961), perceived marginal costs and marginal benefits are vital determinants of investments in education and training. Costs, in this context, include the direct costs of education and the opportunity costs associated with attending education or training. Benefits encompass a range of things such as increases in productivity and cognitive skills, better economic and health outcomes, and improved social status (Drèze & Kingdon, 1999). Although limited in volume and challenged by data and study design, the research generally reveals findings that are consistent with theory—namely that costs are negatively associated with decisions to enroll in TVET and benefits are positively associated with enrollment decisions (Chandrashekhar & Mukhopadhyay, 2006; Grubb, 1988; Kremer et al., 2004).

Grubb (1988) offers the most detailed examination of the economic model of decisions to enroll in TVET, or specifically, community colleges. He uses state-level data between 1970 and 1980 to accomplish two goals: (1) examine the role of economic conditions and labor markets on state-level community college enrollment rates, and (2) estimate the legislative demand for community college enrollment examining the political conditions that drive this decision. Student demand (operationalized as rate of enrollment in community college) is estimated as a function of tuition costs, opportunity costs (operationalized as average annual income for males and females between 18-24 years with 12 grades of schooling), returns (separately estimated for males and females), unemployment rate in the state, growth of professional occupations in the state, lagged enrollment rates, and a set of dummy variables for various ethnic groups. The results of this analysis show that tuition is significantly negatively associated with enrollment decisions and the effect of opportunity costs is not significant. Other economic studies (Corman & Davidson, 1984; Perna & Titus, 2005; Sulock, 1982) show similar results.[12]

Challenges in computing good measures of expected returns to education have contributed to a dearth of research that relate rates of return to enrollment decisions (Behrman, 2010). However, there have been several studies that use data on earnings instead of using information on *expected* or *perceived* returns (Jensen, 2010). For example, Grubb (1988) examined the relationship between *expected* returns (operationalized as the ratio of earnings of those with 1-3 years of college to those with

[12] The unit of analysis in all of these studies, save the one by Perna & Titus (2005), is the state or other geographic unit. The estimates therefore, might suffer from some aggregation bias.

high school degrees) and enrollment decisions and found a positive relationship. However, he also reported that the relationship was limited to females.

Two empirical investigations used experimental data to establish the link between perceived benefits and enrollment decisions (Jensen, 2010; Nguyen, 2008). As part of a cluster-randomized trial in the Dominican Republic, students at randomly selected treatment schools were provided information on the returns to different levels of schooling in the Dominican Republic. Using data from surveys administered before the intervention and a year following the intervention, the study found that treatment students' perceptions of returns were more accurate and that the rate of enrollment in secondary education had gone up compared to that of the control group (Jensen, 2010). Similar results were reported from an experimental study conducted in Madagascar (Nguyen, 2008).

Although the findings described above do not provide clear validation for the significance of costs and benefits on enrollment decisions in all contexts and at all levels of education, there is a strong theoretical basis for their inclusion in demand models.

3.3.6 Quality

The quality of education and training is considered an important supply-side factor expected to affect the demand for education and training (Hansushek, 1995; Kremer, 1995). Again, there is limited literature on this issue specific to TVET as opposed to education in general. However, overall, the literature generally supports the theory of positive associations between educational quality and enrollment (Birdsall, 1985; Glewwe & Jacoby, 1994; Tansel, 2002).

Higher quality is associated with higher enrollments and early and timely enrollments. However, a major methodological challenge in this research is the fact that the quality measures may themselves be biased. The reason is that students of higher ability are more likely, than their lower ability counterparts who apply, to be selected into schools/institutions with more and better resources—a factor that can introduce bias in the coefficient estimate of the quality measure (Mani et al., 2009). Researchers have used Heckman's selection correction method to account for school choice and address this issue (Glewwe & Jacoby, 1994).

Only Grubb (1988) has examined the influence of quality within a TVET framework. His measure of quality is the proportion of community college graduates receiving vocational degrees rather than degrees in general academic subjects. This measure is meant to capture the vocational differentiation available in the community college curriculum. The results of his study show that there is a negative relationship between the two variables. In the context of Grubb's study, the results imply that as the vocational content offered by a community college increases, students are less likely to enroll.

3.3.7 Labor market indicators

The unemployment rate, profile of industries or occupations in a region, and growth of different types of occupations have been used as labor market indicators in demand studies (Grubb, 1988; Walstab, 2008). Grubb (1988) argues that the role of unemployment (and other labor market indicators) as a determinant of school enrollment is ambiguous and difficult to interpret because these variables may indicate the future

economic benefits of getting an advanced degree, or the opportunity costs of attending school, or current labor market opportunities available to part-time students. He finds no relationship between unemployment rate and community college enrollment decisions but he does find a small positive relationship between the growth rate of professional occupations and community college enrollment (Grubb, 1988).

Contrary results are reported in a more recent Australian study (Walstab, 2008). The study uses regression methods to estimate the relative importance of demographic and economic factors on TVET participation rates and finds that regional labor market conditions and the industrial profile of a region explain up to 40 percent of the variation in regional participation rates. Low unemployment rates and a large proportion of workers employed in hospitality, manufacturing, and retail are positively associated with participation in all types of TVET. Further, comparing participation rates across public and private providers, the study finds that economic factors are stronger predictors of enrollments at private institutions than public institutions.

3.4 Returns to TVET

The literature on the returns to education is vast and has received significant attention within the field of education economics (Bennell, 1995; 1996; Kingdon et al., 2008; Psacharopoulos & Patrinos, 2004; Patrinos et al., 2006; Schultz, 2004). Several studies have discussed the methodological issues associated with estimating market (Behrman & Deolalikar, 1995; Card, 1999; 2001; Maluccio, 2003; Schultz, 2004) and non-market returns (McMahon, 2001) to education in developed and developing countries. Research on the returns to TVET (Grubb, 1992; Long & Shah, 2008; Meer,

2006), however, is relatively sparse and more so in the case of developing countries (Duraisamy, 2002; Grootaert; 1990; Moenjak & Worswick, 2003; Psacharopoulos & Patrinos, 1993).

Historically, studies estimating the rate of return to education found larger returns for lower levels of schooling (Psacharopoulos, 1981). Subsequent studies however, have found the returns function to be U-shaped, with the returns increasing with each level of education up to the secondary or higher secondary stage and then gradually declining at or beyond the college level (Colclough et al., 2009).

Studies examining the returns to TVET in developing countries have estimated returns to TVET in general (Duraisamy, 2002), to secondary-level TVET (Moenjak & Worswick, 2003; Psacharopoulos & Patrinos, 1993), and compared returns to formal and informal training (Grootaert, 1999). As noted by Griliches (1977), OLS estimates of returns often suffer from self-selection bias and omitted variable bias that must be accounted for in wage equations. The studies identified, each control for self-selection using Heckman's (1979) two-stage procedure, which allows for estimating participation in wage work and estimating wages in a simultaneous equation framework.

Duraisamy (2002) uses nationally representative survey data at two time points (1983 and 1993) to estimate the returns to academic education and TVET in India. The model is estimated separately for males and females and urban and rural residents but does not control for any household or context level factors. The findings indicate that, controlling for years of education the returns to "technical diploma/certificate" programs (Duraisamy, 2002; p 620) are higher than the returns to college education. Further, the

returns are highest for those in the youngest age cohort (15 to 29 year olds) and returns to TVET for rural residents are higher than for TVET participants in urban areas.

Moenjak and Worswick's (2003) study estimates returns to TVET at the higher secondary level in Thailand, controlling for several individual and family characteristics including marital and migration status, parent's education, parent's occupation, location, and household size. They also find statistically higher returns to secondary TVET than general education at the same level. Psacharapoulos and Patrinos (1993) found similar results for secondary TVET in seven out of 11 Latin American countries.

Grootaert's (1990) examination of formal and informal TVET in Cote d'Ivoire takes a more nuanced approach and estimates wage returns conditional upon the sector of employment. His study uses a large-scale survey of 1600 households in Cote d'Ivoire. Controlling for several demographic and household characteristics, as well as for costs of TVET, the results indicate that in contrast to formal TVET, the private returns to informal TVET are significantly lower. Further, his examination by the sector of employment finds that schooling, and postsecondary formal TVET are significantly associated with employment in the public sector. He also finds that degree attainment is more strongly associated with public sector employment than years of education. In contrast, the private sector values the type of TVET for employment decisions. Thus, those receiving informal TVET are more likely to obtain work in the informal sector. In general, the study estimates that the returns for both types of TVET (formal and informal) are about 10 percent for each year of TVET.

The studies reviewed show positive significant returns to TVET programs. But the lack of research in this area limits the generalizability of these findings. Further, data

constraints in several developing countries imply that reported estimates perhaps suffer from some degree of bias and must be interpreted with caution.

3.5 Impact of TVET

Studies examining the impact of TVET programs are generally context-specific (Agodini & Deke, 2004; Plank, 2001; Kemple et al., 2008) owing to the varied nature of TVET and variations in delivery across contexts. Nonetheless, researchers have conducted cross-national examinations of the outcomes of TVET programs (Hanushek et al., 2011; Psacharopoulos, 1993). The outcomes measured by these studies have focused on dropout prevention (Agodini & Deke, 2004), high school completion (Plank, 2001), and labor market outcomes. Recently, research has also looked at the impact of TVET participation over the lifecycle (Hanushek et al., 2011). The methodological problems encountered in evaluating the outcomes of TVET (Ryan, 2001) and the mixed results from studies make it difficult to generalize findings across settings.

In the United States, research on TVET has comprised evaluations of traditional career and technical education programs offered in public high schools as well as the Career Academies programs. The latter are high school based learning communities organized around a vocational theme that integrate academic and TVET curricula and provide students work-based learning opportunities (Kemple et al., 2008). Career Academies have been well researched using randomized controlled designs. Findings from MDRC's (Kemple et al., 2004; 2008) rigorous eight-year follow-up of program participants indicates that while students at high-risk of dropping out were more likely to stay in school until the end of high school, the program had no impact on high school

completion rates per se. But high school completion was higher in Career Academies than the national average. For students who entered Career Academies at low or medium risk of dropping out were also more likely to finish high school, and during that time showed increased participation in career development activities. At the postsecondary level, Career Academies were seen to have no impact on postsecondary enrollment. Again, postsecondary outcomes were higher among students at Career Academies (and in the control group) than the national average. The impact on average monthly earnings was positive and persisted throughout the follow-up period. While this impact was more stable among young men, for women it was not statistically significant over time. Further, students who entered the Academies at high risk of dropping out were seen to have the strongest labor market outcomes.

Other U.S. studies have examined the impact of high school TVET on dropout behavior (Agodini & Deke, 2004; Plank, 2001). In their study Agodini and Deke (2004) compare the probability of dropping out among "vocational concentrators"[13] and those in general academic programs. They find no difference in dropout rates in the two groups. But their study finds that students who want to pursue the vocational track are less likely to dropout when enrolled as "vocational concentrators" rather than as "vocational explorers"[14]. Plank's (2001) study suggests slightly different findings. His study used transcript data to compute the ratio of career and technical credits to academic course credits of high school students. He concludes that the probability of dropping out of high school is significantly reduced with a ratio of three TVET courses to four academic

[13] "Vocational concentrators" are required to take three or more courses in a single occupational area and three fewer low-level academic courses (Agodini & Deke, 2004).
[14] The study defines "vocational explorers" as students in broader occupational training programs where they can take courses in a variety of occupational areas (Agodini & Deke, 2004).

credits. Plank's (2001) study does not control for any of the selection issues in comparing students who take a combined curriculum to other students in the sample and thus must be interpreted with caution.

Research on the impact of TVET on educational and labor outcomes outside of the United States has also had mixed results. Hanushek et al. (2011) recently used an international sample of labor market outcomes from 18 OECD countries to compare outcomes of individuals with general education to those with TVET. The study adopted a difference-in-differences approach to control for selection bias, as well as propensity score matching and included several controls for background characteristics and ability. While there was significant variation in estimates across countries, the overall results showed that individuals with general education have lower initial employment outcomes and wage patterns than those with TVET. Over the lifecycle (as early as age 50), however, those with general education experience higher probabilities of employment, while the initial advantages of TVET participants diminish.

The impact of TVET in developing and emerging economies has also received some attention. In the case of Latin America, Psacharopoulos (1993) examines the impact of secondary-level TVET on earnings in 11 Latin American countries. He finds that in seven countries, TVET graduates have significantly higher gross earnings than general secondary education students. In some cases the earnings of TVET graduates are up to 20 percent higher. The study finds that after controlling for costs of schooling and foregone earnings, the impact on individual earnings is only significantly positively higher in four countries.

A more rigorous study is conducted by Malamud & Pop-Eleches (2010) in Romania. They use a regression discontinuity design (RDD) to examine the shift from vocational education to general education and compare labor market outcomes of students affected by the shift in policy. Controlling for a range of background characteristics and omitted variable bias through the use of RDD, the study finds no difference in labor market outcomes, as measured by employment status and wages, between those in the TVET track and those in the general education track.

The study of impacts of TVET has largely focused on employment and wage outcomes. While some research from OECD countries explores the effect of TVET on educational outcomes, in most cases the results have been mixed. This is largely due to the lack of methodological rigor in study design. As Ryan (2001) notes, controlling for the effects of selection along with the varied nature of TVET delivery within and across countries, makes TVET evaluations a complex endeavor.

Chapter 4: The Predictors of Participation in Technical and Vocational Education and Training in India

A review of the literature on the predictors of participation in TVET programs reveals that there are various limitations to building a consensus on the factors associated with demand for TVET. Besides the paucity of research, the nature of TVET complicates research in this area. Yet, the TVET literature provides some direction on the factors that are most likely to influence TVET enrollment decisions. Student educational attainment and aspirations (Agodini et al., 2004; Aypay, 2003; Curtis, 2008; Moenjak & Worswick, 2003), perceived costs and benefits of TVET programs (Chandrashekhar & Mukhopadhyay, 2006; Grubb, 1988; Kremer et al., 2004), household income (Sandefur et al., 2005), parents' education (Curtis, 2008; Fullarton, 2001; Moenjak & Worswick, 2003), indicators of the quality of TVET options (Grubb, 1988) and the macroeconomic context (Grubb, 1988; Walstab, 2008) have been found to have an association with participation decisions. This direction is useful in building a conceptual model for studying demand in developing countries where the TVET sector is relatively nascent and undergoing massive restructuring and expansion.

In the case of India, changes in TVET policies have focused on expansion of programs, along with the development of a comprehensive qualification and certification framework to recognize skills acquired through informal apprenticeships. These policy measures are designed to meet the national target of "skilling" 500 million Indians by 2020 (King, 2012). Programs to improve the technical capability and quality of new and existing institutions have also been initiated (Planning Commission, 2013). One

motivation underlying this redesign is to make the TVET system more "demand-driven" (Federation of the Indian Chambers of Commerce and Industry [FICCI], 2012; Planning Commission, 2007; UNESCO-UNEVOC, 2013). For reasons discussed earlier, there is also a need to make the TVET system more focused on individual users. An understanding of user-related issues thus far has been based on descriptive information on participation rates. More recently, an attempt has been made to examine the aspirations and constraints faced by youth and young adults in accessing TVET opportunities, albeit through descriptive methods (Aggarwal et al., 2011; FICCI, 2012). Findings from these surveys indicate that limited awareness about TVET options and the perceived "low status" of TVET-related careers are correlated with TVET participation decisions (Aggarwal et al., 2011). Examining these questions through rigorous, empirical work is critical given the scale and cost of proposed reforms in the sector.

This chapter attempts to fill in some of the gaps in our understanding of the factors that predict TVET participation in India by addressing the following questions –

1. What are the predictors of TVET participation, controlling for district-level variation, in India?
    a. Do the predictive relationships vary by type of TVET – formal or informal?
2. Are the predictive relationships for TVET participation different for males and females?

4.1 Conceptual framework

Building on extant research, a conceptual model is proposed for the Indian context and tested empirically using data from a nationally representative large-scale survey. The proposed framework builds on human capital and sociological theories and models education and training investment decisions as influenced by various social, economic and political factors within the household, the community, and society. The proposed conceptual model (illustrated in figure 2) draws largely from previous work (Perna & Titus, 2005; Perna, 2006) on access and choice in postsecondary enrollment decisions.

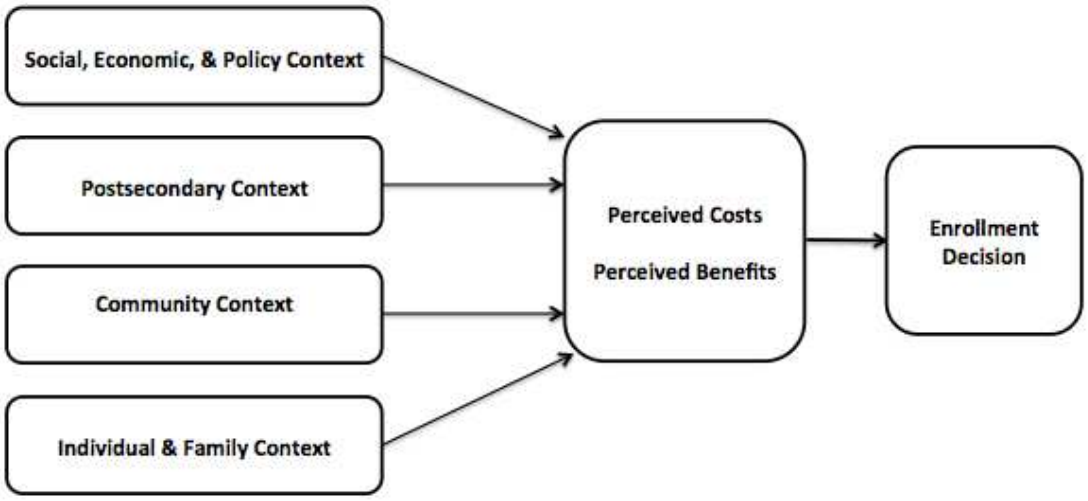

Figure 2. Proposed conceptual framework for studying individual demand for TVET in India

The proposed model posits that enrollment decisions reflect cost-benefit assessments that are impacted by a variety of contextual factors. In the Indian context,

these influences include those at the individual and family level, those operating in the community, in the postsecondary education space, and at macro levels.

At the individual and family level, educational attainment and prior achievement, household income and parents' education, along with demographic indicators (age, gender, urbanicity, and marital status) influence decision-making. The role of social capital, although seen to contribute significantly in explaining group differences (Perna, 2006), is excluded from the analytic model for two reasons. Firstly, as noted by Dika & Singh (2002), the use of social capital in estimating enrollment decisions is often governed by data limitations and leads to erroneous conceptualizations of social capital. Nationally representative datasets available in India have so far not collected any information on social and cultural capital indicators, and until recently, research examining educational outcomes in India have not used social capital indicators. Therefore, there is no evidence of how well these indicators perform in empirical models for India. Recently, Iyengar (2012) used qualitative methods to examine the role of social capital in school participation in one district of India. She found little evidence that social capital was linked to education discussions and decisions within the family or within the village/community.

At the individual/family level, the model has been adapted from Perna's (2006) model in two ways. First, marital status has been added to other demographic variables. In the Indian context, marital status is an important demographic dimension of interest but to date, it has not been discussed in the TVET literature. Studies on educational participation in India note that in the case of girls (exogenous) marriage practices and the gender division of labor in the household influence enrollment and participation decisions

(Drèze & Sen, 1995). Research confirms that ‘age at marriage’ variables are particularly important in explaining female participation in education and training in India (Kingdon, 2010).

Second, the model hypothesizes that occupational prestige or occupational status considerations influence individual and family decisions on TVET enrollments. While occupational status has not been studied in India, qualitative research indicates that TVET is often rejected based on its association with low-prestige occupations (Agrawal, 2012). So far, no occupational status index or similar measure has been developed for the Indian context.

At the level of the community, social and cultural norms have been shown to influence enrollment decisions. In the Indian context, socio-cultural norms related to patriarchy and perceptions around female education and employment have been found to significantly explain gender variations in enrollment (Boissiere, 2004; Kingdon, 2007; Pal, 2004).

Moving up to the postsecondary or higher education context, institutional characteristics and quality of education and training are predictive of TVET enrollment.

Finally, the social, economic and policy context is hypothesized to have both direct and indirect effects on TVET enrollment. This includes labor market indicators that describe the economic context (for e.g. unemployment rate, growth in certain types of occupations), demographic factors that describe the social context (for e.g., changes in the proportion of working age adults), and the extent of public-private partnerships representing the policy context (for e.g., expansion of TVET services through public-private partnerships). The role of macro context variables has not been studied in the

Indian context but is relevant to incorporate given recent changes in the TVET sector in the country.

Following the conceptual model depicted above, Table 1 summarizes the specific factors hypothesized to affect TVET enrollment in India. However, due to data constraints (discussed below), the variables in parentheses were not included in the empirical analysis.

Table 1

Factors hypothesized to predict participation in TVET in India

Demographic Controls	Individual /Family Level	Community Level	Postsecondary Level	Macro Context
Age	Schooling	*(Community wage rate)*	Size of TVET sector	Unemployment rate
Gender	Social Group	*(Norms)*	*(Quality)*	*(Job Growth)*
Urbanicity	Household Income	*(Occupational Prestige)*	*(Cost)*	
Marital Status	Parents' Education	*(Access to electricity)*	*(Access to TVET)*	
	Household Occupation	*(Access to roads)*	*(Access to college)*	
	Household Size			
	(Ability)			
	(Social Capital)			
	(Cultural Capital)			

Note. Factors in parentheses cannot be included in the analytic models for this study due to data limitations.

Indicators of individual ability, social and cultural capital, and social norms were not included in the analytic model as there is no available data on these measures. While some large-scale surveys have gathered information on these constructs those surveys lack detailed information on participation in TVET.

Occupational prestige or occupational status scales have been constructed using factor analysis of data on individuals' rankings of various occupations (Nakao & Treas, 1989). However, this information is unavailable for the Indian context and therefore not included in the analytic model proposed here.

Previous examinations of the quality of TVET institutions in India have focused on employability of TVET graduates, the teaching-learning methods at TVET institutions, networks with employers and industry, and their funding mechanisms. While these could serve as indicators of quality of TVET institutions information on these indicators has not been collected in any systematic, quantifiable way.

Although survey data do not include information on the cost incurred by an individual to participate in TVET, reports indicate that the cost of attending public TVET institutions is negligible. The cost of private TVET, on the other hand, is significantly higher and could present barriers to entry (Tilotia, n.d.). These data were not systematically collected or available for inclusion in the present analysis.

The effect of supply side factors like growth in the number of jobs and expansion of TVET services is best captured in a longitudinal framework. Longitudinal data capturing these indicators along with data on participation in TVET is not available.

4.2 Methods

This study is a departure from previous attempts to understand TVET participation in India in that it examines participation decisions through empirical analysis of large-scale survey data and examines factors hypothesized to affect TVET decisionmaking beyond those at the individual and household level.

4.2.1 Data

Data for this study were drawn from the National Sample Survey Organization's (NSSO) Employment and Unemployment Survey (Schedule 10). Specifically, the 61st and 66th rounds of the NSSO were used. The Employment and Unemployment Survey has been conducted by the NSSO every five years since 1972. The 61st round, conducted in 2004-2005 was the first time that information on participation in TVET was collected as part of this survey. A second round on participation in TVET was collected in 2009-2010 as part of the 66th round. The 61st and 66th rounds of the NSSO include a nationally representative sample covering all states and union territories in the country (except those inaccessible throughout the year due to infrastructure or conflict). The 2004-2005 panel includes 124,680 households, and the 2009-2010 panel includes 100,957 households.

The Employment and Unemployment surveys gather data on three key areas critical to this research study. First, the survey includes questions on educational participation for all members in sampled households. This includes information on "current attendance" (for those below 30 years of age) as well as "highest level of education completed". Second, the survey captures fine-grained information on educational participation detailing the kind of education (general, technical or vocational) that was accessed, the type of institution that was attended, the field of training, the duration of training, and consequent employment outcomes. Third, the survey collects detailed information on employment outcomes of all household members above 15 years of age, including occupational and wage details, and unemployment spells. Background and demographic information from the survey is linked to household characteristics,

educational participation, and employment outcomes using unique household and person identifiers.

In addition to the data described above, two additional sources of data were accessed. First, information on the supply of TVET institutions was gathered from the website of the Directorate General of Education and Training (DGET)[15] in India. These data include information on the number of institutions (public and private) in each district of the country. Second, district-level data on rainfall since the 1950s was accessed from official records. For each district for which rainfall data was available the average rainfall over the past ten years was computed and used in the present analysis. Table 2 provides a description of all the variables used in the empirical analysis.

Table 2

Description of variables from Employment and Unemployment Survey (Round 61 - 2004-05 & Round 66 - 2009-10)

Variable	Source: Employment & Unemployment Survey (NSS)	Description
OUTCOME VARIABLES		
Vocational Education	Round 61 (2004-05), Round 66 (2009-10)	Categorical variable indicating participation in vocational education - coded '1' if participated in formal vocational education, '2' if participated in informal vocational education, and '0' otherwise
Duration of TVET	Round 61 (2004-05), Round 66 (2009-10)	Continuous variable indicating duration of training program in weeks
Field of TVET	Round 61 (2004-05), Round 66 (2009-10)	Categorical variable indicating field of training
DEMOGRAPHIC VARIABLES		
Age	Round 61 (2004-05), Round 66 (2009-10)	Age in years

[15] http://dget.gov.in/

Age squared	Round 61 (2004-05), Round 66 (2009-10)	The quadratic term for age
Female	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable for gender - coded '1' for female and '0' for male
Urban	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable for location - coded '1' for urban and '0' for rural
Marital status	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable indicating marital status - coded '1' if married at the time of survey and '0' if otherwise
Other Backward Class	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable indicating social exclusion - coded '1' if OBC and '0' otherwise
Dalit	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable indicating social exclusion - coded '1' if Dalit and '0' otherwise
Adivasi	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable indicating social exclusion - coded '1' if Adivasi and '0' otherwise
Muslim	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable indicating religious affiliation - coded '1' if Muslim and '0' if non-Muslim
INDIVIDUAL CHARACTERISTICS		
Years of schooling	Round 61 (2004-05), Round 66 (2009-10)	Continuous variable indicating years of schooling (Range: 0 to 17)
HOUSEHOLD CHARACTERISTICS		
Gender of head of the household	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable indicating the household head's gender - coded '1' if female and '0' otherwise
Head of the household's years of schooling	Round 61 (2004-05), Round 66 (2009-10)	Continuous variable indicating years of schooling of the head of the household
Agricultural Household	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable indicating household type - coded '1' if agriculture is the main occupation, and '0' otherwise
Salaried Household	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable indicating household type - coded '1' if the main occupation is salaried, and '0' otherwise
Labor Household	Round 61 (2004-05), Round 66 (2009-10)	Dummy variable for household type - coded '1' if the main occupation is casual labor, and '0' otherwise

POSTSECONDARY AND MACROECONOMIC CONTEXT		
Supply of TVET institutions	Directorate General of Education and Training (India)	Continuous variable indicating district-wise institutions offering TVET programs
Unemployment rate	Round 61 (2004-05), Round 66 (2009-10)	Continuous variable indicating the rate of unemployment at the district level
Rainfall		Continuous variable indicating millimeters of average rainfall over a 10-year period
Weight	Round 61 (2004-05), Round 66 (2009-10)	Probability weights to account for sampling design
District ID	Round 61 (2004-05), Round 66 (2009-10)	Unique ID for districts in the sample

4.2.2 Analytic Sample

The analytic sample was restricted to all those between 15 and 29 years of age. The lower bound of 15 years was motivated by the fact that TVET programs in India can be accessed as early as high school (Sharma, 2010). More importantly, the surveys gathered TVET participation information from all 15-29 year olds in 2004-05 and from all those between 15-59 in the 2009-10 round. Although a wider age range was available for study in the 2009-10 panel, the analytic sample was restricted to those between 15-29 years in order to make meaningful comparisons in predictive patterns across the two panels.[16]

The NSSO surveys gather information on participation in technical education programs. These programs are available at the undergraduate and graduate levels and cover several fields of study (see NSSO, 2013; p8 for a description). Technical education programs offering a diploma or certificate in “crafts” or “other subjects” (excluding

[16] Descriptive and multivariate analysis on the entire sample of 15-59 year olds is included in Appendix A.

engineering, medicine, and agriculture) at the undergraduate levels were considered equivalent to TVET programs for the purposes of this study[17] and individuals who had participated (or were currently enrolled, at the time of survey) in these programs were classified as TVET participants.

Table 3 shows the sample sizes for the relevant age groups from each round of the survey 2004-2005 and 2009-2010. This sample was further trimmed due to missing data. Cases with missing information on key variables (those shown in Table 2) were removed from the sample. Thus, the size of the analytic sample was 133,841 individuals in the case of the 2004-05 panel, and 102,216 individuals in the 2009-10 panel.

Table 3

Analytic sample as proportion of full survey sample

Survey Panel	Full Sample	Relevant Age Range[a]	Missing individual data (%)	Missing district data (%)	Analytic Sample	Proportion of relevant age range
2004-05	602833	162779	1.91	15.87	133841	82.22
2009-10	459784	125378	0.48	18.09	102216	81.53

Note. [a] The relevant age range implies all those who were surveyed for participation in TVET. This included 15 to 29 year olds in 2004-05 and 15 to 59 year olds in 2009-10 (288662 cases). For comparability, only 15-29 year olds from the 2009-10 panel have been included here. See Appendix A for descriptive statistics and empirical estimates on the sample of 15-59 year olds.
Data on the number of TVET institutions in each district and district-level rainfall were available for 505 and 556 out of 585 districts in 2004-05 and 508 and 559 out of 612 districts in 2009-10.

[17] See section 3.1 in Chapter 3 for definitions of TVET programs.

Technical and vocational education and training programs in India show low participation rates. The tables presented below show the proportion of TVET participants and non-participants in the data. Overall, TVET participants constituted 12 percent of the relevant age-group in 2004-05, and about eight percent in 2009-10. This dip in TVET participation in 2009-10 was driven mainly by lower informal TVET participation rates in 2009-10 as compared to those in 2004-05. Of those participating in TVET in 2004-05, four percent accessed formal TVET programs while 7.72 percent were in informal TVET programs. The respective figures in 2009-10 were about 3.5 percent and 4.8 percent, respectively.

Table 4a
Weighted percent of analytic sample participating in TVET

	Any TVET	Formal TVET	Informal TVET
2004-05 (15-29 year olds)	11.84	4.12	7.72
2009-10 (15-29 year olds)	8.33	3.55	4.78
2009-10 (15-59 year olds)	7.80	2.74	5.05

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)

Table 4b
Weighted percent of TVET participants by gender and location (15-29 year olds)

	Formal TVET		Informal TVET	
	2004-05	2009-10	2004-05	2009-10
Urban Males	35.38	35.02	17.98	24.81
Urban Females	22.47	22.74	8.27	9.93
Rural Males	25.51	26.09	44.92	45.21
Rural Females	16.64	16.15	28.83	20.05

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)

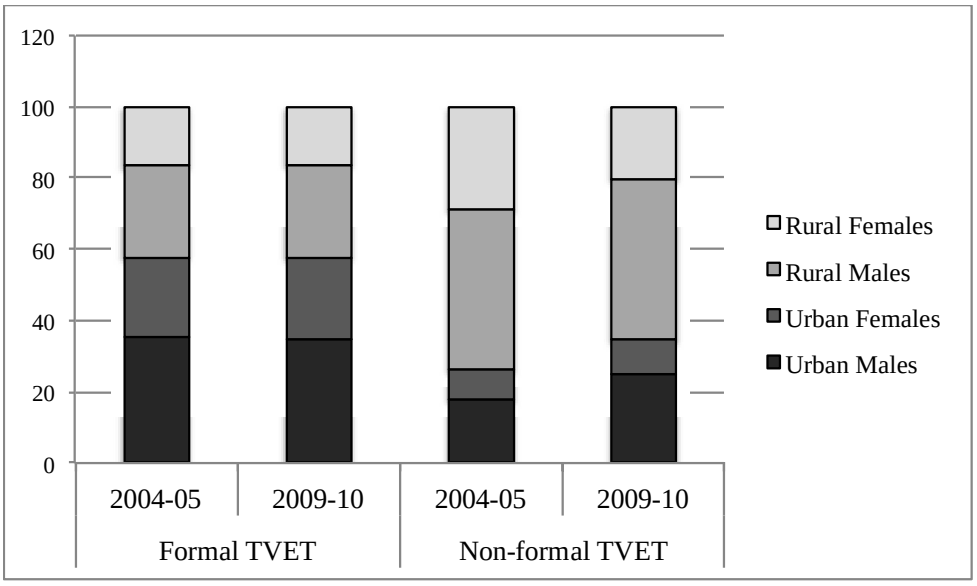

Figure 3. Percent of formal and informal TVET participations between 15-29 year olds, by gender and location (2004-05 & 2009-10).

Participation rates for formal and informal TVET by gender and location are presented in Table 4b and Figure 3. Males participated in formal TVET at higher rates than females in urban and rural areas with urban areas showing higher participation rates, in general. The proportion of male and female TVET participants by urbanicity did not undergo much change between 2004-05 and 2009-10. Informal TVET participation however showed some differences. Rural males formed the largest group of informal TVET participants. While this is true for 2004-05 and 2009-10, informal TVET participation among rural females showed some decline in 2009-10 (rural females comprised the second largest group of informal TVET participants in 2004-05. The 2009-

10 data show that urban males participated in informal TVET at a higher rate than in 2004-05.

These changes in informal TVET participation between 2004-05 and 2009-10 could be a function of changes in the way data on informal TVET participation was collected in 2009-10. In 2009-10, informal TVET was defined as that taking place within the family, through "self-learning", "on the job", or in other ways; whereas in 2004-05, informal TVET was classified as that acquired within the family or in other ways. The significant difference observed in informal TVET participation could also be explained on the basis of changes in labor force participation rates between 2004-05 and 2009-10. See Section 7.1 in Chapter 7 for a discussion.

Tables 5a to 5e provide descriptive statistics on the relevant variables for the two cross-sectional panels. The analytic samples are compared to each other in Table 5a and the subsequent tables compare subgroups on the basis of gender and urban-rural location. The average age in both panels is about 21 years with half the panel comprising females and about a third living in urban areas. A slightly higher proportion report being married (46 percent) in 2004-05 than in 2009-10 (41 percent).

The OBC group comprises the largest social group followed by Dalits and Adivasis. Muslims comprise about 14 percent of the panels.

Dummy variables for various levels of completed education provide a sense of how the panels are distributed across various education levels. (Also see Figures 4 and 5 for graphical displays of educational attainment in each panel). The largest educational attainment group across both panels was those with at least 5 years of schooling while those with a graduate (Master's) degree comprised the smallest group.

Table 5a indicates that there are significant differences in educational attainment among 15-29 year olds surveyed in 2004-05 versus those surveyed in 2009-10. For starters, the proportion of the sample with no schooling has significantly reduced over the 5-year period from 24 percent in 2004-05 to 15 percent in 2009-10. Similarly, the proportion in each of the educational attainment categories (from 5 years of schooling to those with a Bachelor's degree) has increased over this period. The proportion completing 10 years of schooling increased from 30 percent in 2004-05 to 41 percent in 2009-10. There were even slight increases in the proportion receiving Bachelor's degrees (from 6 percent to 8 percent).

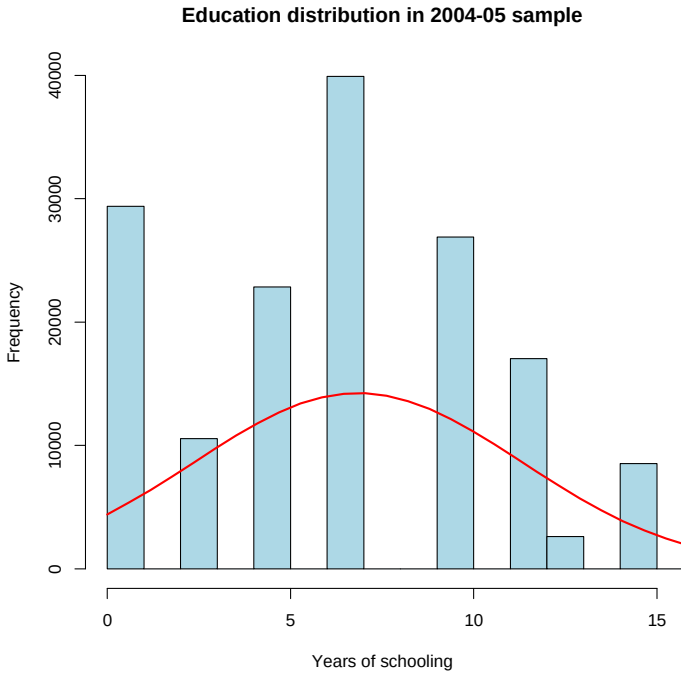

Figure 4. Years of schooling in the 2004-05 analytic sample

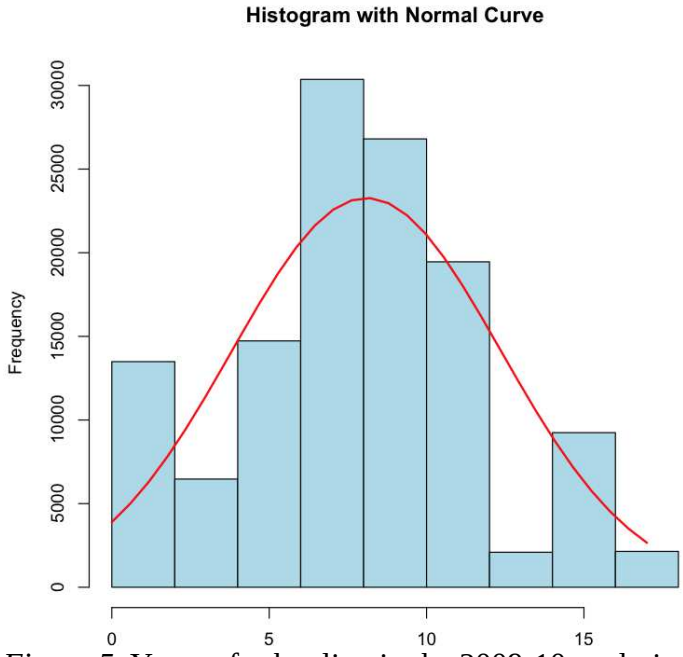

Figure 5. Years of schooling in the 2009-10 analytic sample (15-29 year olds)

The data also show that the average income of households (measured by the consumption expenditure at the household level) went up significantly between 2004-05 and 2009-10. Although there was no change in the proportion of households engaged in salaried work or self-employment, in 2009-10, the proportion of households engaged in waged work increased from 28 percent (in 2004-05) to 32 percent.[18]

At the district level, in 2009-10, the unemployment rate showed a decrease over that reported in 2004-05; from 5.83 percent to 4.6 percent. See Appendix A for plots showing distributions of other district level characteristics (number of TVET institutions and rainfall) in 2004-05 and 2009-10.

[18] As discussed in Section 7.1 in Chapter 7, the increase in the proportion of waged workers could be attributed to the implementation of a large public works employment program for rural households between 2006 and 2008.

Differences in educational attainment across the 2004-05 and 2009-10 panels are also observed when examined by subgroups – rural males and females and urban males and females. These differences followed the same trend as discussed above – a reduction in the proportion not receiving any schooling and significant increases in educational attainment up to grade 12. Amongst urban males and females, the proportion earning a Bachelor's degree also increased significantly over 2004-05.

Table 5a

Weighted descriptive statistics for select variables

	2004-05 panel		2009-10 panel	
Predictors	Mean	SE	Mean	SE
Individual Characteristics				
Age	21.40	0.02	21.43	0.03
Age Squared	475.45	0.86	477.58	1.29
Female (Dummy)	0.49	0.00	0.48	0.00
Urban (Dummy)	0.28	0.01	0.29	0.00
Marital status (Dummy)	0.46	0.00	0.41	0.00
Social Group – Dalit (Dummy)	0.20	0.00	0.20	0.00
Social Group – Adivasi (Dummy)	0.08	0.00	0.09	0.00
Social Group – OBC (Dummy)	0.41	0.00	0.41	0.01
Religious Minority (Muslim Dummy)	0.13	0.00	0.14	0.00
No education (Dummy)	0.24	0.00	0.15	0.00
5 years of education (Dummy)	0.68	0.00	0.78	0.00
10 years of education (Dummy)	0.30	0.00	0.41	0.00
12 years of education (Dummy)	0.16	0.00	0.22	0.00
Bachelor's degree (Dummy)	0.06	0.00	0.08	0.00
Master's degree (Dummy)	0.01	0.00	0.01	0.00
Monthly Household Expenditure	3770.54	22.59	5802.85	49.24
Household Head's Education	4.55	0.04	5.20	0.05
Female-headed Household (Dummy)	0.09	0.00	0.10	0.00
Household occupation: Business (Dummy)	0.23	0.00	0.23	0.00
Household occupation: Salaried (Dummy)	0.12	0.00	0.12	0.00
Household occupation: Wage Work (Dummy)	0.28	0.00	0.32	0.00
N	*159670*		*124795*	
District Characteristics				
District TVET Capacity*	19.97	21.55	19.90	21.51

Average District Rainfall**	101.10	54.44	103.26	55.30
District Unemployment Rate	5.83	4.76	4.60	3.76
N	*585*		*612*	

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)

Note. * N=505; ** N=556

Table 5b

Weighted descriptive statistics – Rural Males

	2004-05		2009-10	
	Mean	SE	Mean	SE
Age	21.16	0.03	21.04	0.05
Age Squared	465.65	1.25	460.87	2.14
Female (Dummy)	0.00	0.00	0.00	0.00
Urban (Dummy)	0.00	0.00	0.00	0.00
Marital status (Dummy)	0.35	0.00	0.30	0.01
Social Group – Dalit (Dummy)	0.21	0.00	0.23	0.01
Social Group – Adivasi (Dummy)	0.10	0.00	0.11	0.00
Social Group – OBC (Dummy)	0.42	0.01	0.41	0.01
Religious Minority (Muslim Dummy)	0.11	0.00	0.12	0.01
No education (Dummy)	0.19	0.00	0.11	0.00
5 years of education (Dummy)	0.72	0.00	0.82	0.01
10 years of education (Dummy)	0.28	0.00	0.39	0.01
12 years of education (Dummy)	0.13	0.00	0.19	0.01
Bachelor's degree (Dummy)	0.04	0.00	0.05	0.00
Master's degree (Dummy)	0.01	0.00	0.01	0.00
Monthly Household Expenditure	3300.40	20.88	4965.15	46.27
Household Head's Education	3.49	0.04	4.10	0.07
Female-headed Household (Dummy)	0.08	0.00	0.09	0.00
Household occupation: Business (Dummy)	0.16	0.00	0.16	0.00
Household occupation: Salaried (Dummy)	0.00	0.00	0.00	0.00
Household occupation: Wage Work (Dummy)	0.35	0.00	0.39	0.01
N	*52158*		*38103*	

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)

Table 5c

Weighted descriptive statistics – Rural Females

	2004-05		2009-10	
	Mean	SE	Mean	SE
Age	21.51	0.03	21.65	0.05
Age Squared	480.29	1.25	486.77	2.02
Female (Dummy)	1.00	0.00	1.00	0.00

Urban (Dummy)	0.00	0.00	0.00	0.00
Marital status (Dummy)	0.64	0.00	0.59	0.01
Social Group – Dalit (Dummy)	0.21	0.00	0.23	0.01
Social Group – Adivasi (Dummy)	0.11	0.00	0.12	0.00
Social Group – OBC (Dummy)	0.43	0.01	0.41	0.01
Religious Minority (Muslim Dummy)	0.12	0.00	0.12	0.00
No education (Dummy)	0.40	0.00	0.26	0.01
5 years of education (Dummy)	0.52	0.00	0.65	0.01
10 years of education (Dummy)	0.19	0.00	0.27	0.01
12 years of education (Dummy)	0.08	0.00	0.13	0.00
Bachelor's degree (Dummy)	0.02	0.00	0.03	0.00
Master's degree (Dummy)	0.00	0.00	0.01	0.00
Monthly Household Expenditure	3257.16	22.57	4903.58	43.53
Household Head's Education	3.64	0.04	4.19	0.05
Female-headed Household (Dummy)	0.09	0.00	0.09	0.00
Household occupation: Business (Dummy)	0.17	0.00	0.16	0.00
Household occupation: Salaried (Dummy)	0.00	0.00	0.00	0.00
Household occupation: Wage Work (Dummy)	0.35	0.00	0.40	0.01
N	*50956*		*36934*	

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)

Table 5d

Weighted descriptive statistics – Urban males

	2004-05		2009-10	
	Mean	SE	Mean	SE
Age	21.49	0.05	21.54	0.06
Age Squared	479.31	2.21	481.67	2.44
Female (Dummy)	0.00	0.00	0.00	0.00
Urban (Dummy)	1.00	0.00	1.00	0.00
Marital status (Dummy)	0.24	0.01	0.22	0.01
Social Group – Dalit (Dummy)	0.16	0.01	0.16	0.01
Social Group – Adivasi (Dummy)	0.03	0.00	0.03	0.00
Social Group – OBC (Dummy)	0.36	0.01	0.39	0.01
Religious Minority (Muslim Dummy)	0.17	0.01	0.16	0.01
No education (Dummy)	0.09	0.00	0.06	0.00
5 years of education (Dummy)	0.86	0.01	0.91	0.00
10 years of education (Dummy)	0.49	0.01	0.59	0.01
12 years of education (Dummy)	0.30	0.01	0.37	0.01
Bachelor's degree (Dummy)	0.12	0.00	0.15	0.01
Master's degree (Dummy)	0.02	0.00	0.03	0.00
Monthly Household Expenditure	4907.10	57.85	7671.44	121.89

Household Head's Education	6.94	0.10	7.65	0.10
Female-headed Household (Dummy)	0.10	0.00	0.09	0.00
Household occupation: Business (Dummy)	0.40	0.01	0.38	0.01
Household occupation: Salaried (Dummy)	0.42	0.01	0.41	0.01
Household occupation: Wage Work (Dummy)	0.12	0.00	0.13	0.00
N	*29225*		*25796*	

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)

Table 5e

Weighted descriptive statistics – Urban females

	2004-05		2009-10	
	Mean	SE	Mean	SE
Age	21.62	0.05	21.81	0.05
Age Squared	485.04	2.23	493.57	2.33
Female (Dummy)	1.00	0.00	1.00	0.00
Urban (Dummy)	1.00	0.00	1.00	0.00
Marital status (Dummy)	0.51	0.01	0.48	0.01
Social Group – Dalit (Dummy)	0.15	0.01	0.15	0.01
Social Group – Adivasi (Dummy)	0.03	0.00	0.03	0.00
Social Group – OBC (Dummy)	0.36	0.01	0.39	0.01
Religious Minority (Muslim Dummy)	0.17	0.01	0.17	0.01
No education (Dummy)	0.16	0.01	0.10	0.00
5 years of education (Dummy)	0.80	0.01	0.86	0.00
10 years of education (Dummy)	0.47	0.01	0.58	0.01
12 years of education (Dummy)	0.29	0.01	0.37	0.01
Bachelor's degree (Dummy)	0.13	0.00	0.17	0.01
Master's degree (Dummy)	0.03	0.00	0.04	0.00
Monthly Household Expenditure	5158.49	64.75	8106.44	167.54
Household Head's Education	7.23	0.09	7.80	0.09
Female-headed Household (Dummy)	0.11	0.00	0.12	0.01
Household occupation: Business (Dummy)	0.43	0.01	0.40	0.01
Household occupation: Salaried (Dummy)	0.41	0.01	0.40	0.01
Household occupation: Wage Work (Dummy)	0.12	0.00	0.14	0.00
N	*27332*		*23962*	

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)

4.2.3 Analytic Methods

The analysis of the data proceeded in two stages. The first stage included a descriptive analysis of TVET participation by various individual and context-level characteristics. Graphical displays of the data produced as part of the descriptive analysis provided a first look at the trends in TVET participation across the variables of interest. The graphs provided some insight into the trends likely to be observed in the multivariate analysis.

In the second stage, multivariate analysis was used to estimate the predictive relationships between the various socio-demographic and contextual factors and participation in TVET, controlling for other factors. The dependent outcome – participation in TVET – is defined as a categorical variable and therefore requires multivariate techniques that model the *logit* or log-odds of the outcome or event (Allison, 2001).[19] Hierarchical generalized linear modeling (HGLM) was used to model the probability of participation in TVET, taking in to account the clustered nature of the data (Raudenbush & Bryk, 2002). In the case of this study, individuals and households (level 1 units) are nested within districts (level 2 units). Not taking in to account this multilevel structure can lead to aggregation bias, miscalculation of standard errors, and heterogeneity of regression (Raudenbush & Bryk, 2002).

The dependent variable in the HGLM was defined as a categorical variable with three levels - formal TVET, informal TVET, and no TVET. The outcome was expressed in log-odds and examined using a multinomial logit link function with fixed and random intercepts.

[19] Modeling categorical outcomes using linear regression methods would violate OLS assumptions and give inconsistent and inefficient estimates (Allison, 2001).

At level 1, the model examined the relationship between types of TVET enrollment and individual and household level characteristics controlling for various demographic variables. At the second level, district characteristics were added to the model to explain additional variation in TVET participation.

The models can be expressed as follows –

Level 1

$$\eta_{ij(m)} =$$

$$\beta_{0j(m)} + \beta_{1ij(m)}(Demographic\ Controls) +$$

$$\beta_{2ij(m)}(Individual\ Characteristics) + \beta_{3ij(m)}(Household\ Characteristics)$$

Level 2

$$\beta_{0j(m)} = \gamma_{00} + \gamma_{01(m)}(District\ Characteristics) + r_{0j}$$

In the equations above, *i* denotes the individual, *j* denotes the district, and *m* denotes the type of enrollment (formal, informal or no TVET). $\eta_{ij(m)}$ is the probability of individual *i* in district *j* participating in either formal or informal TVET (compared to the reference category of not participating in either). The β terms give the coefficient estimates for each level 1 predictor in log-odd units and γ is the coefficient estimate for level 2 predictors. $\beta_{0j(m)}$ and γ_{00} are the intercept terms at level 1 and level 2, respectively. r_{0j} is the random effect at level 2.

The predictors at level 1 were group-mean centered while the remaining variables were grand-mean centered to improve interpretation. Fixed effects were used at the

individual level (assuming that all individuals in a district are influenced in the same way by district-level variables) and random effects were used at the district level to allow for differences between districts.

The models were estimated in SAS 9.3 using the GLIMMIX procedure designed for HGLM models with categorical outcomes. The procedure allows for the use of sampling weights to generate representative regression estimates and computes sampling errors of estimators based on the complex sample design.

Using the procedures described above two separate models were estimated for each response option (binary and unordered categorical) – a pooled model consisting of the entire analytic sample, and separate models for males and females.[20]

4.3 Results

At the outset, a pooling test was carried out to determine if the data from the 2004-05 panel and the 2009-10 panel should be pooled for the empirical analysis. For the pooling test a linear probability model was estimated with the all the predictors (identified in Table 2) fully interacted with a dummy variable for panel. Statistically significant estimates of the interaction terms suggested different underlying models thus making the case for analyzing separate models by panel.

[20] The empirical analysis does not include separate models by urbanicity. See Section 4.4 for a discussion on this and other limitations of the empirical analysis.

4.3.1 Descriptive Results

Plots generated as part of the descriptive analysis are shown in Figures 6 to 13. The age distribution of participants and non-participants, in general, is relatively similar in the female group. Among males, TVET participants are clustered in the 19-21 age and the 24-26 age groups. The average age of non-participants is about 20-21 years, and those in TVET are closer to 23 years, on average. Females in informal TVET programs are slightly younger than male participants and also younger than those in formal TVET programs.

A similar trend is observed for years of schooling completed by TVET participants and non-participants. Figure 8 shows that a smaller proportion of females than males complete more than six years of schooling and a larger proportion remain unschooled. While the differences in schooling levels are not that apparent among male and female TVET participants in Figure 8, the differences are more pronounced when comparing formal and informal TVET participants to non-participants (see Figure 9). On average, those in formal TVET programs (males and females) are seen to complete over 12 years of schooling. This is a significant difference from those in informal TVET programs where males and females show an average of seven and five years of schooling, respectively. The average years of schooling for non-participants is a little over six years.

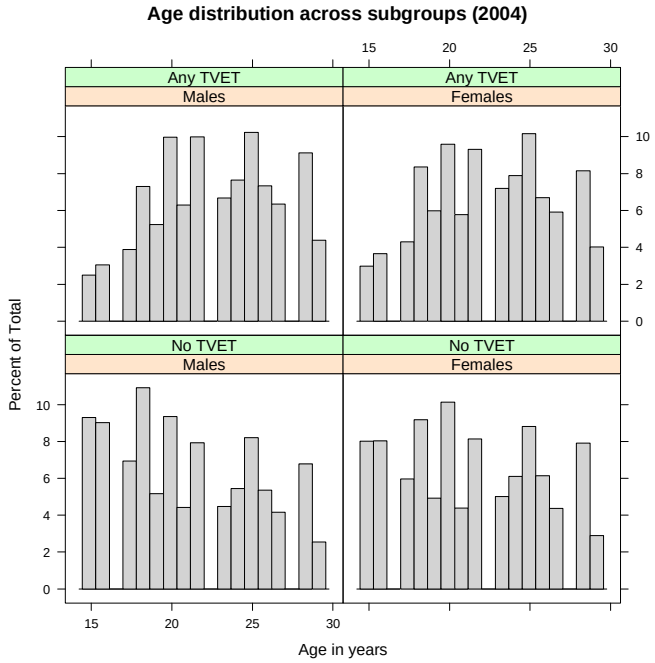

Figure 6. Age distribution by gender and TVET status in the 2004-05 analytic sample

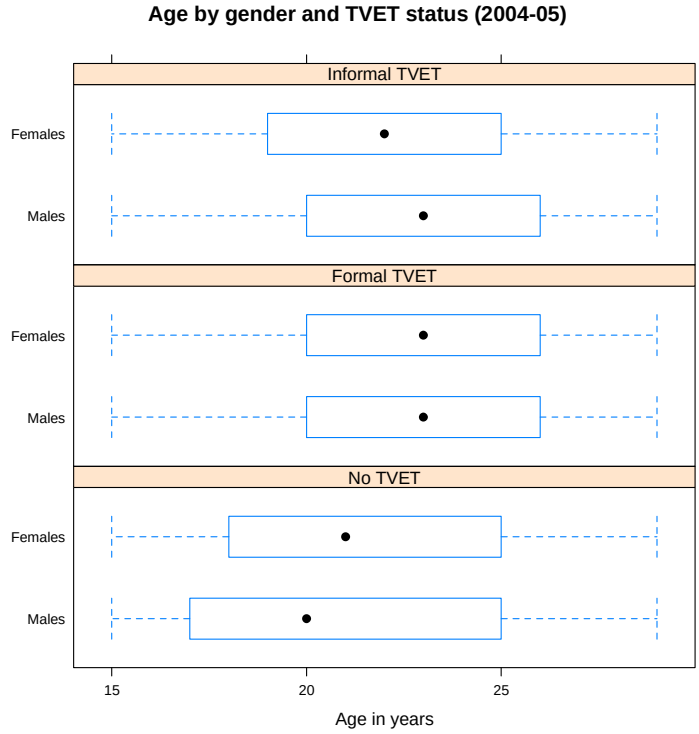

Figure 7. Average age of formal and informal TVET participants (2004-05 analytic sample)

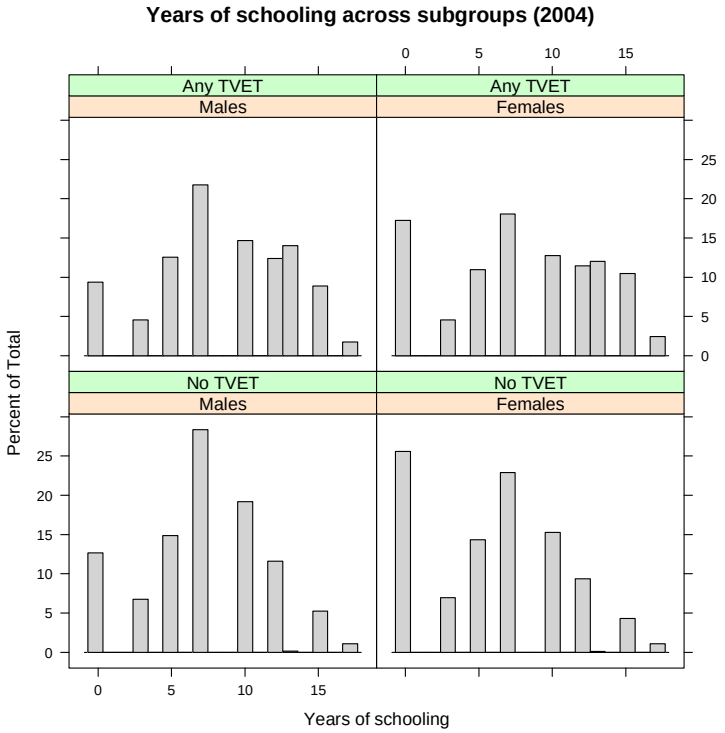

Figure 8. Years of schooling by gender and TVET status in the 2004-05 panel

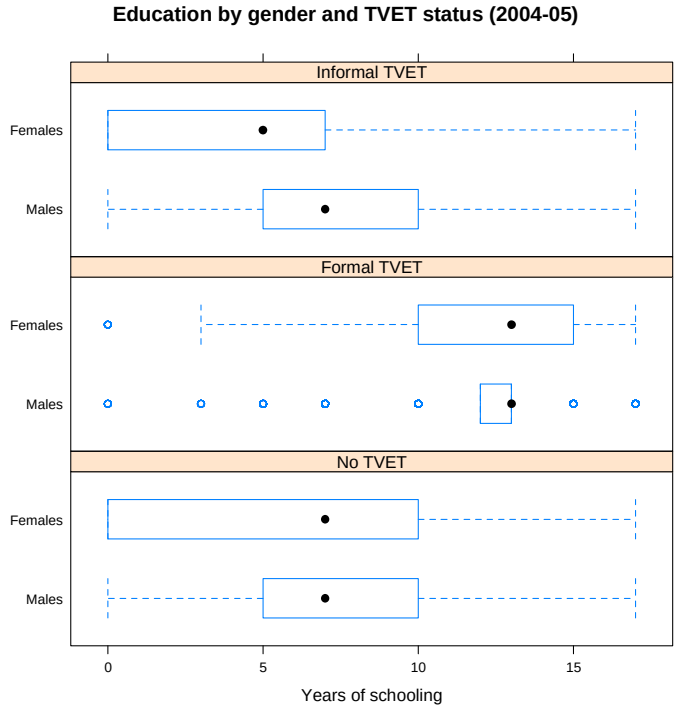

Figure 9. Average years of schooling among formal and informal TVET participants (2004-05)

Graphical displays of the 2009-10 data show similar age distributions among the participant and non-participant groups and the formal and informal TVET groups as those observed in the 2004-05 panel. The similarities in average age of participants and non-participants are more apparent in Figure 11. The plot shows formal TVET participants are about 24, while those in informal TVET are slightly older and those not participating in TVET are on average, younger.

The difference between TVET participants and non-participants in terms of the years of schooling completed is presented in Figures 12 and 13. The only differences observed are for the formal TVET group. While the average years of schooling across participants and non-participants is around seven years, formal TVET participants (males and females), complete more years of schooling than those not participating in TVET or those participating in informal TVET; the average years of schooling for this group is about 12 years.

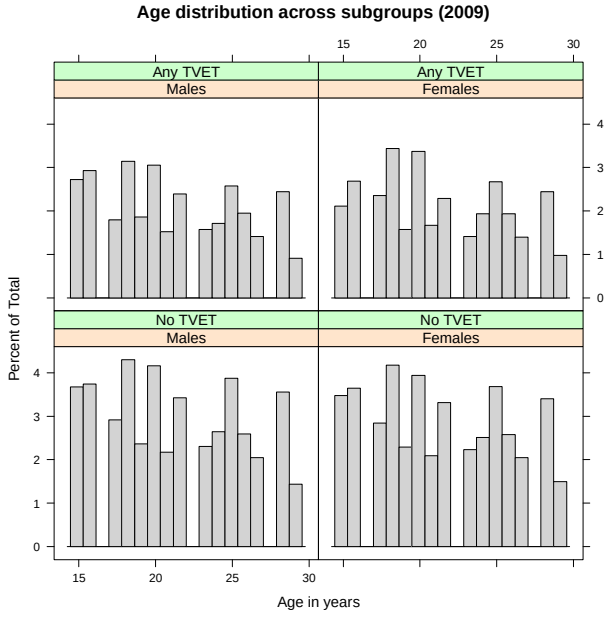

Figure 10. Age distribution by gender and TVET status in the 2009-10 panel (15-29 year olds)

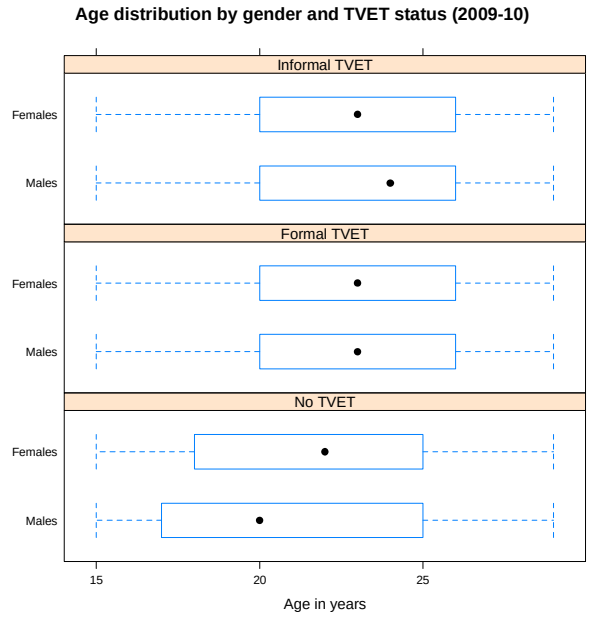

Figure 11. Average age of formal and informal TVET participants (15-29 year olds; 2009-10)

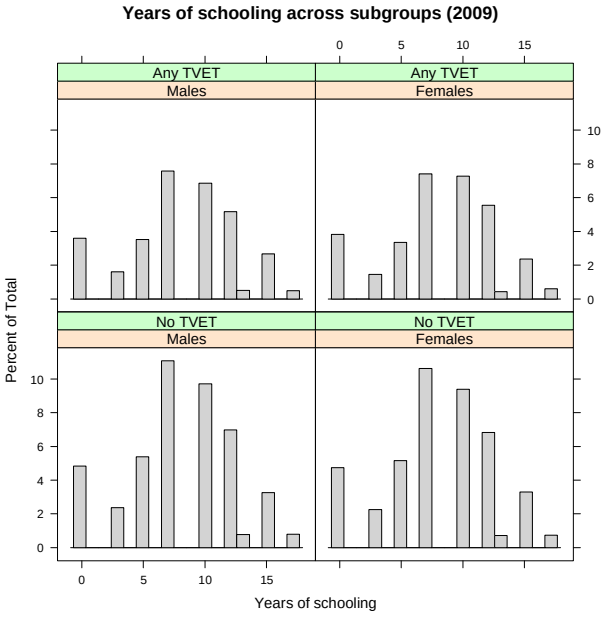

Figure 12. Years of schooling by gender and TVET status in the 2009-10 panel (15-29 year olds)

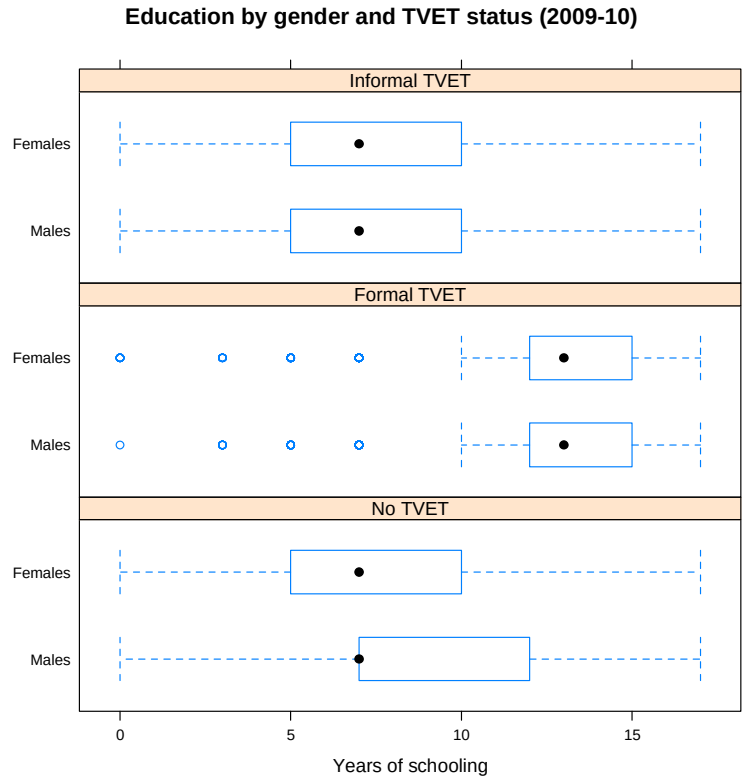

Figure 13. Average years of schooling among formal and informal TVET participants (15-29 year olds; 2009-10)

4.3.2 HGLM Results (Binary Outcome)

The binary HGLM examines the variables predicting participation in any TVET. The results for the entire analytic sample are presented in Table 6a; Table 6b includes results of the analysis by gender. The tables show marginal effects for each of the predictors in the regression. Marginal effects are population-averaged measures and denote the associated change in the response for small (discrete, in the case of categorical predictors) changes in the predictor variables. The main effects are highlighted below along with differences along gender dimensions. Model fit and classification accuracy are discussed at the end of this section.

Table 6a

Marginal effects of factors predicting participation in any TVET

	2004-05	2009-10
Demographic Controls:		
Age	0.04***	0.04***
Age Squared	-0.00***	-0.00***
Female (Dummy; Ref: Male)	-0.02***	-0.03***
Urban (Dummy; Ref: Rural)	0.01**	0.00**
Marital Status (Dummy; Ref: Unmarried)	-0.02***	-0.02***
Social group - OBC (Dummy; Ref: Other)	0.01***	0.01***
Social group - Dalit (Dummy; Ref: Other)	0.01**	0.01**
Social group - Adivasi (Dummy; Ref: Other)	0.01**	0.01**
Religious group - Muslim (Dummy; Ref: Other)	0.00*	0.01*
Individual Characteristics:		
5 years of schooling (Dummy; Ref: No schooling)	0.01***	0.01***
10 years of schooling (Dummy; Ref: Less than 10 years)	-0.00	-0.01
12 years of schooling (Dummy; Ref: Less than 12 years)	0.05***	0.03***
Bachelor's degree (Dummy; Ref: Less than a bachelor's)	-0.02***	-0.01***
Master's Degree (Dummy; Ref: Less than a master's)	0.01	0.00
Household Characteristics:		
Log of Consumption Expenditure	0.01***	0.01***
Household Head's Schooling	0.00	0.00
Female Household Head (Dummy)	0.00	0.00
Household Size	-0.00***	-0.00***
Household Occupation: Self-employment (Dummy)	0.01***	0.01***
Household Occupation: Salaried (Dummy)	0.01***	0.01***
Household Occupation: Wage Work (Dummy)	-0.01***	-0.00***
Context Characteristics:		
Number of TVET institutions	0.00	0.00
Unemployment rate	0.01***	0.00***
Average 10-year rainfall	0.00	0.00
N	*133841*	*102216*

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)
*p<.05, **p<.01, ***p<.001

This study hypothesized TVET participation as a function of several individual and household characteristics as well context-level factors like the district unemployment rate and the supply of TVET institutions in the district. The empirical models found

evidence supporting some of these hypotheses. Results suggest that educational attainment and demographic characteristics are significantly related to TVET participation decisions. While household characteristics also show significant effects these are relatively small. The relationship between unemployment rate and TVET participation is consistently significant, but the magnitude of this relationship is very small once all other factors have been accounted for.

The empirical analysis found that demographic variables are related to TVET participation in the expected direction with females showing a significantly lower likelihood of participation than males. Similarly, belonging to the OBC, Dalit and Adivasi group is associated with slightly higher likelihood of TVET participation than belonging to the majority group.

The relationship between schooling and TVET follows a somewhat U-shaped pattern. Those with five years of education are significantly more likely to participate in TVET than those with no education. The likelihood of TVET participation is highest among those with 12 years of schooling – this relationship is stronger in 2004-05 than in 2009-10 - and lowest among those with a Bachelor's degree.

In terms of household characteristics, parents' occupational background is consistently significantly related to TVET participation after controlling for educational attainment but the magnitude of these effects is quite small. While self-employed and salaried households have a positive effect on TVET participation, households engaged in casual labor showed the reverse relationship with individuals belonging to waged households having a lower likelihood of participating in TVET.

The district-level predictors used in this study do not seem to matter to TVET participation over and above the other predictors in the model. The regional unemployment rate however, shows very small but significant associations with TVET participation.

Figure 14 shows the distribution of odds ratios in the 2004-05 model with 95 percent confidence intervals around the odds ratio estimates. Odds ratio estimates around 1 indicate that that variable does not increase or decrease the odds of TVET participation.

Gender differences in TVET enrollment patterns across the 2004-05 and 2009-10 panels are presented in Table 6b. Some differences are observed with respect to demographic dimensions; age is a stronger predictor of TVET participation among males than females. Surprisingly, this is not the case in 2009-10 but could be related to lower participation rates for women in 2009-10. Being married decreases an individual's likelihood of enrolling in TVET, not surprisingly, this association is stronger in the case of females.

The biggest gender differences are observed in terms of levels of completed schooling. Small differences are also observed across the two panels. In 2004-05, the magnitude of the associations between educational attainment and TVET participation are stronger for males while the reverse is true in 2009-10, where the effects of certain levels of education on TVET enrollment are slightly larger for females. Specifically, both males and females with a high school degree were more likely to enroll in TVET than those without a high school degree. This relationship was larger in magnitude in 2004-05 and larger in the case of the males. In 2009-10 however, the association was much

smaller in magnitude, although still positive and significant, and slightly larger in the case of females.

Household income has a stronger association with male TVET participation than female TVET participation. This relationship was consistent across both years.

Among demographic characteristics, the direction of the relationships were as expected. It is worth noting that among individuals belonging to the *OBC* category, in 2004-05, the likelihood of TVET participation was slightly higher in case of males than females. Among all other social groups, males and females showed similar associations between TVET participation and group affiliation.

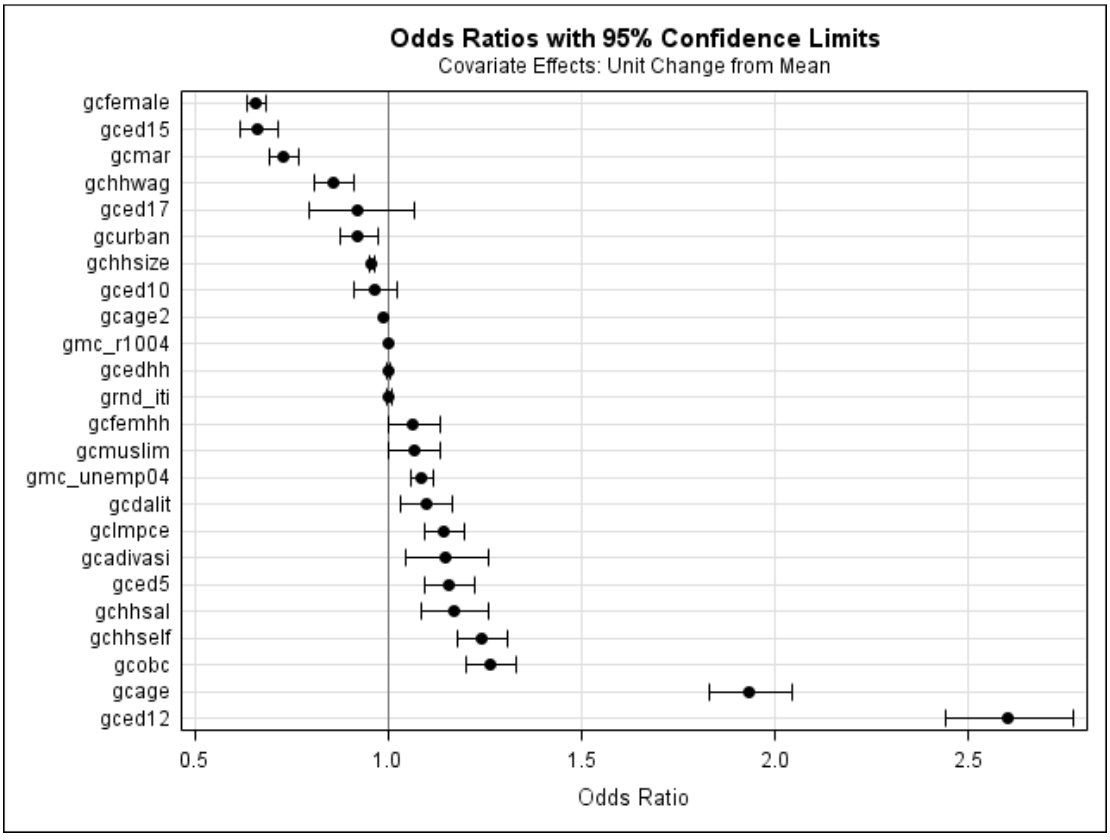

Figure 14. Distribution of odds ratio estimates of predictors of TVET enrollment (2004-05 survey panel)

Table 6b

Marginal effects of factors predicting participation in any TVET, by gender

	2004-05		2009-10	
	Males	Females	Males	Females
Demographic Controls:				
Age	0.05***	0.02***	0.01***	0.02***
Age Squared	-0.00***	0.00***	0.00***	0.00***
Urban (Dummy; Ref: Rural)	-0.00	-0.00*	0.00	0.00
Marital Status (Dummy; Ref: Unmarried)	-0.01***	-0.02***	-0.00***	-0.02***
OBC (Dummy; Ref: Other)	0.02***	0.01***	0.00***	0.00**
Dalit (Dummy; Ref: Other)	0.00	0.00**	0.00	0.01***
Adivasi (Dummy; Ref: Other)	0.01	0.01**	0.00	0.01**
Muslim (Dummy; Ref: Other)	0.00**	-0.00	0.00***	0.00
Individual Characteristics:				
5 years of schooling (Dummy; Ref: Unschooled)	0.01*	0.00***	0.00*	0.01***
10 years of schooling (Dummy; Ref: <10 years)	-0.01	0.00	-0.00***	0.00
12 years of schooling (Dummy; Ref: <12 years)	0.07***	0.04***	0.01***	0.02***
Undergraduate Degree (Dummy; Ref: <15 years)	-0.04***	-0.00**	-0.00***	-0.00
Masters Degree (Dummy; Ref: < master's)	-0.01	-0.00	0.00	0.00
Household Characteristics:				
Log of Consumption Expenditure	0.01***	0.00	0.00***	0.00**
Household Head's Schooling	0.00	0.00	0.00	0.00
Female Household Head (Dummy)	0.00	0.00	0.00	0.00*
Household Size	0.00***	-0.00***	0.00***	-0.00***
Household Occupation: Self-employment	0.02***	0.00**	0.00***	0.01**
Household Occupation: Salaried (Dummy)	0.01***	0.00	0.00***	0.01*
Household Occupation: Wage Work (Dummy)	-0.01***	-0.00*	-0.00*	-0.00
Context Characteristics:				
Number of TVET institutions	-0.00	0.00	0.00	0.00
Unemployment Rate	-0.01***	0.00***	0.00**	0.00***
Average 10-year rainfall	-0.00	0.00	0.00**	0.00
N	*68159*	*65682*	*49922*	*52294*

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)

*p<.05, **p<.01, ***p<.001

Table 6c

Post-estimation classification table of predicted probabilities from Binary HGLM

	Correct		Incorrect			Percentages			
Cutoff Value	TVET	No TVET	TVET	No TVET	Correct	Sensitivity	Specificity	False +ve	False -ve
Classification Results for 2004-05									
0.35	3766	114000	4325	11865	87.9	24.1	96.3	53.5	9.4
0.4	3162	115000	3334	12469	88.2	20.2	97.2	51.3	9.8
0.45	2674	116000	2624	12957	88.4	17.1	97.8	49.5	10.1
0.5	2200	116000	2054	13431	88.4	14.1	98.3	48.3	10.4
0.55	**1809**	**117000**	**1607**	**13822**	**88.5**	**11.6**	**98.6**	**47**	**10.6**
0.6	**1475**	**117000**	**1269**	**14156**	**88.5**	**9.4**	**98.9**	**46.2**	**10.8**
0.65	**1205**	**117000**	**966**	**14426**	**88.5**	**7.7**	**99.2**	**44.5**	**11**
0.7	925	118000	728	14706	88.5	5.9	99.4	44	11.1
0.75	709	118000	548	14922	88.4	4.5	99.5	43.6	11.3
Classification Results for 2009-10									
0.35	1201	91484	1950	7581	90.7	13.7	97.9	61.9	7.7
0.40	965	91959	1475	7817	90.9	11	98.4	60.5	7.8
0.45	771	92316	1118	8011	91.1	8.8	98.8	59.2	8
0.50	608	92589	845	8174	91.2	6.9	99.1	58.2	8.1
0.55	483	92780	654	8299	91.2	5.5	99.3	57.5	8.2
0.60	382	92960	474	8400	91.3	4.3	99.5	55.4	8.3
0.65	**274**	**93101**	**333**	**8508**	**91.4**	**3.1**	**99.6**	**54.9**	**8.4**
0.70	**195**	**93217**	**217**	**8587**	**91.4**	**2.2**	**99.8**	**52.7**	**8.4**
0.75	**142**	**93279**	**155**	**8640**	**91.4**	**1.6**	**99.8**	**52.2**	**8.5**

To examine model fit the classification table and ROC curve were examined. The classification table highlights the extent of errors made by the model in predicting TVET and non-TVET participants while the ROC curve plots the true positive rate (sensitivity) against the false positive rate (specificity) for a selected cut-point in the dataset. A cutoff of 0.5 was selected for these data. This means that cases with a predicted probability greater than 0.5 are classified as TVET participants and those with predicted probabilities under 0.5 are classified as non-participants.

As seen in Table 6c and Figures 15 and 16 below, the estimated models correctly classified about 80 percent of the cases across both panels. The proportion of false positive and false negative classifications for various cutoff values is presented in Table 6c. In the case of the 2004-05 panel, cutoff values ranging from 0.55 to 0.65 showed the greatest classification accuracy. About 11 percent of the sample of TVET participants were incorrectly classified as not being in TVET while 46 percent of the sample was incorrectly classified as receiving TVET. These results are depicted graphically in Figure 15, which shows that, for a cutoff value of 0.5, the area under the curve is about 82 percent; implying a good fit of the model with the data.

The model fit for 2009-10, as seen in Table 6c and Figure 16, shows that higher cutoff values (0.65 to 0.75) correctly classify nearly 91 percent of the individuals in the analytic sample. The differences in the classification accuracy for different cutoff values are very small. On average, the model wrongly classifies nearly half of all participants as enrolled in TVET and incorrectly classifies about eight percent of TVET participants as not receiving any TVET. The area under the ROC curve for the 2009-10 model is about 80 percent, again indicating a good fit with the data.

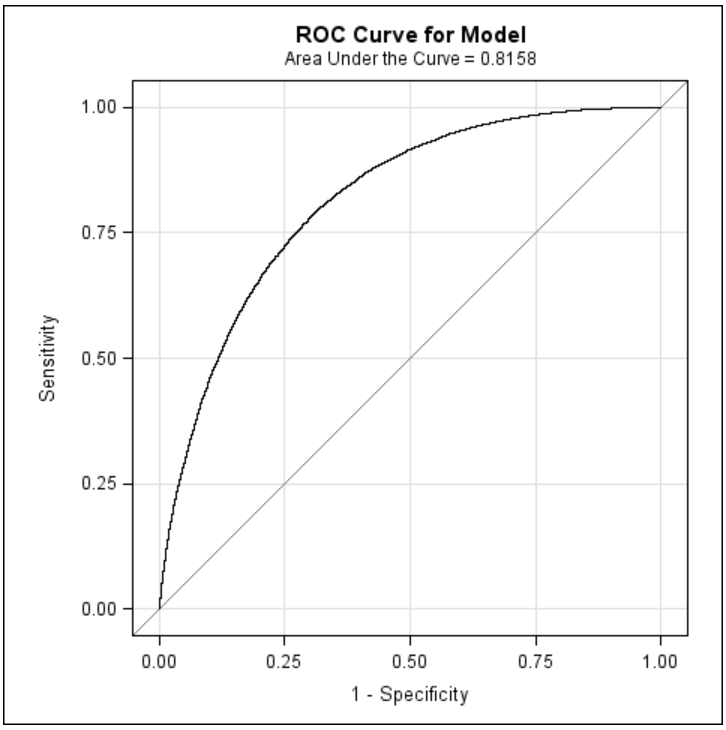

Figure 15. ROC curve for 2004-05 model

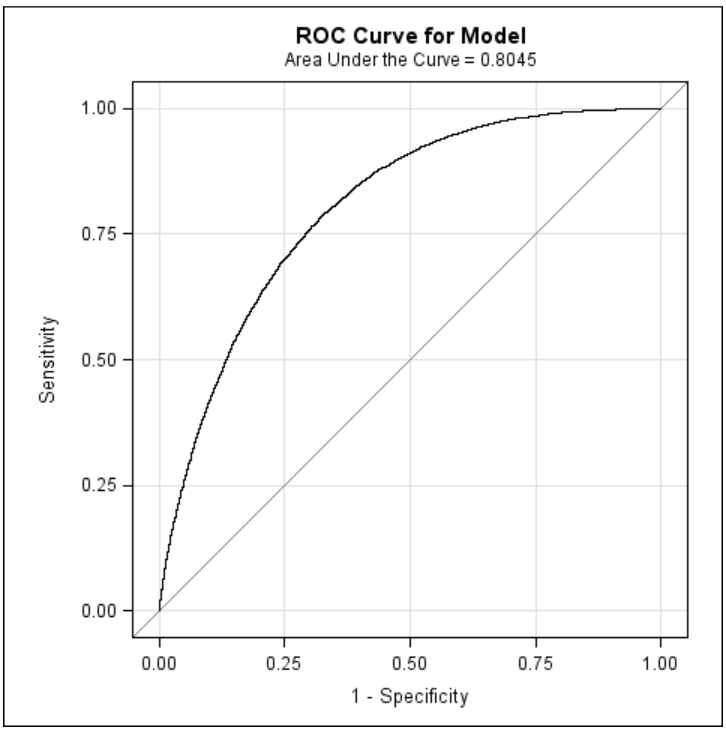

Figure 16. ROC curve for 2009-10 model

4.3.3 HGLM Results (Multinomial Outcome)

Given the clear differences between formal and informal TVET participants observed in descriptive analysis the next stage of the multivariate work focused on examining differences in the factors that predict enrollment in formal and informal TVET programs. As noted earlier, the outcome variable was defined as discrete with three unordered categories – participation in formal TVET, participation in informal TVET and no TVET – and a hierarchical generalized linear model was estimated. Results are presented in Table 6d. These are discussed and compared to highlight differences across the panels, followed by the results from the gender-wise regression analysis in Tables 6e and 6f.

Results from the multivariate analysis (presented in Table 6d) show a few significant differences in formal and informal TVET participation patterns as predicted by individual and household characteristics. Formal TVET participation is associated with unmarried males, with 10-12 years of completed schooling belonging to salaried households. Participation in informal TVET while also more likely to be male dominated is associated with lower levels of education (those with higher levels of education shower lower probabilities of participation). Individuals from lower income households and those belonging to households where the primary occupation is self-employment have higher odds of enrolling in informal TVET programs.

Comparisons between the predictive patterns across the two time points show that a student with 10 years of completed schooling was more likely to enroll in formal TVET in 2004-05 than in 2009-10, while a student with 12 years of completed schooling was

more likely to enroll in formal TVET in 2009-10 than in 2004-05, as compared to a student who did not have 12 years of schooling. Thus, formal TVET enrollment was associated with higher levels of education in 2009-10 than in 2004-05.

Across the two panels, informal TVET participation also shows some changes in predictive patterns. Most notably, compared to males, females showed lower odds of TVET participation (formal and informal) in 2009-10 than in 2004-05. Those from salaried and self-employed households, on the other hand, were much more likely to participate in informal TVET in 2009-10 than they were in 2004-05.

To explain district level variation in formal and informal TVET participation, the analysis examined the effect of three district-level characteristics – average rainfall in the district, the size of the TVET sector, and the unemployment rate. As expected, the size of the TVET sector was found to be significant in the case of formal TVET enrollments but increased the odds of participation in a negligible way. The unemployment rate improved the odds of participating in formal and informal TVET and the magnitude of this relationship was similar across both types of TVET. There were no significant gender differences with respect to district level factors.

Table 6d

Odds ratio estimates of factors predicting participation in formal and informal TVET

	2004-05		2009-10	
	Formal	Informal	Formal	Informal
Demographic Controls:				
Age	1.73***	1.99***	1.78***	2.53***
Age Squared	0.99***	0.99***	0.99***	0.98***
Female (Dummy; Ref: Male)	0.85***	0.56***	0.78***	0.39***
Urban (Dummy; Ref: Rural)	1.08	0.90**	1.13*	0.98

Marital Status (Dummy; Ref: Unmarried)	0.55***	0.86***	0.59***	0.79***
Dalit (Dummy; Ref: Other)	1.30***	0.98	1.32***	1.05
Adivasi (Dummy; Ref: Other)	1.06	1.10	1.03	1.30**
OBC (Dummy; Ref: Other)	1.17***	1.31***	1.18***	1.22***
Muslim (Dummy; Ref: Other)	0.94	1.11**	1.05	1.21***
Individual Characteristics:				
5 years of schooling (Dummy)	1.77***	1.27***	2.04***	1.42***
10 years of schooling (Dummy)	3.30***	0.71***	2.46***	0.71***
12 years of schooling (Dummy)	4.77***	0.69***	5.10***	0.59***
Undergraduate Degree (Dummy)	0.60***	0.74***	0.84***	0.71***
Masters Degree (Dummy)	0.85*	0.94	1.02	0.99
Household Characteristics:				
Log of Consumption Expenditure	1.38***	0.89***	1.26***	1.06
Household Size	0.92***	1.00	0.94***	0.98**
Household Occupation: Self-employment	0.97	1.38***	0.89*	1.96***
Household Occupation: Salaried	1.15**	0.91	1.14*	1.54***
Household Occupation: Wage Work	0.85**	0.85***	0.89	1.04
Household Head's Schooling	1.02***	0.98***	1.01*	0.98***
Female Household Head (Dummy)	1.07	1.07	1.15*	1.00
Context Characteristics:				
Number of TVET institutions	1.01**	1.00	1.01***	1.00
Unemployment Rate	1.11***	1.08***	1.08***	1.05*
Average 10-year rainfall	1.00	1.00*	1.00	1.00*
N	*6129*	*9502*	*4222*	*4560*

Source: Employment and Unemployment Survey of India (2004-05 & 2009-10)
*p<.05, **p<.01, ***p<.001

Patterns of TVET participation were also examined by gender; this enabled isolating differences in relationships between predictors and type of TVET for males and females separately. The results for the 2004-05 panel are presented in Table 6e and for the 2009-10 panel in Table 6f.

For males and females alike, being married significantly reduced the odds of formal TVET participation more than the odds of informal TVET participation. The

association, not surprisingly, is stronger in the case of females. This pattern is also evidenced in 2009-10.

The effect of education on formal TVET participation was slightly different among males and females. Males with 10 and 12 years of schooling consistently showed higher odds of enrolling in formal TVET as compared to females with the same levels of schooling. While this pattern is consistent across both panels, comparisons between the two panels show that females with 10 years of schooling had higher odds of participating in formal TVET in 2004-05 than in 2009-10. Further differences between 2004-05 and 2009-10 indicate that higher levels of education made informal TVET participation less likely in 2009-10 than in 2004-05 but more likely amongst females with five years of education. Thus, the results show that the association between education and informal TVET participation was weaker in 2009-10.

The association between household wealth and formal TVET participation among males and females also show some small but significant differences. While there is a positive association between household wealth and formal TVET for both groups, the magnitude of this relationship is larger for males. Over time – in 2009-10, this association also weakens in the case of males.

Other household characteristics show similar patterns by gender and by type of TVET. Self-employed households are related to higher odds of informal TVET participation among males and females. This relationship between informal TVET participation and belonging to a self-employed household is stronger in 2009-10 for both males and females. Individuals in households involved in casual labor are less likely to

enroll in any TVET program. This association is not significant in the case of females in either panel.

Table 6e

Odds ratio estimates of factors predicting participation in formal and informal TVET by gender (2004-05)

	Males		Females	
	Formal	Informal	Formal	Informal
Demographic Controls:				
Age	1.68***	2.20***	1.89***	1.84***
Age Squared	0.99***	0.99***	0.99***	0.99***
Urban (Dummy; Ref: Rural)	1.06	0.92	1.06	0.86**
Marital Status (Dummy; Ref: Unmarried)	0.66***	0.98	0.44***	0.80***
Dalit (Dummy; Ref: Other)	1.26***	0.95	1.39***	1.04
Adivasi (Dummy; Ref: Other)	1.08	1.00	1.11	1.23
OBC (Dummy; Ref: Other)	1.25***	1.38***	1.08	1.22***
Muslim (Dummy; Ref: Other)	1.03	1.15**	0.86	1.05
Individual Characteristics:				
5 years of schooling (Dummy)	1.18	1.29***	3.23***	1.27***
10 years of schooling (Dummy)	3.77***	0.70***	2.29***	0.74***
12 years of schooling (Dummy)	5.39***	0.64***	4.47***	0.80*
Undergraduate Degree (Dummy)	0.53***	0.64***	0.73***	0.91
Masters Degree (Dummy)	0.83	0.76	0.85	1.05
Household Characteristics:				
Log of Consumption Expenditure	1.61***	0.91*	1.16**	0.88*
Household Size	0.91***	1.00	0.93***	1.00
Household Occupation: Self-employment	0.95	1.54***	1.07	1.18**
Household Occupation: Salaried	1.19*	0.96	1.20	0.84
Household Occupation: Wage Work	0.76***	0.82***	0.94	0.87*
Household Head's Schooling	1.03***	0.98***	1.01	0.98***
Female Household Head (Dummy)	1.04	1.06	1.06	1.11
Context Characteristics:				
Number of TVET institutions	1.01***	1.00	1.00	1.00
Unemployment Rate	1.10***	1.08***	1.11***	1.07***
Average 10-year rainfall	1.00	1.00*	1.00*	1.00
N	*3633*	*5857*	*2496*	*3645*

Source: Employment and Unemployment Survey of India (2004-05)

*p<.05, **p<.01, ***p<.001

Table 6f

Odds ratio estimates of factors predicting participation in formal and informal TVET by gender (2009-10)

	Males		Females	
	Formal	Informal	Formal	Informal
Demographic Controls:				
Age	1.66***	2.88***	2.03***	2.03***
Age Squared	0.99***	0.98***	0.99***	0.99***
Urban (Dummy; Ref: Rural)	1.08*	0.96	1.17*	1.04
Marital Status (Dummy; Ref: Unmarried)	0.76***	0.88***	0.45***	0.74**
Dalit (Dummy; Ref: Other)	1.18***	0.94**	1.48***	1.30**
Adivasi (Dummy; Ref: Other)	1.01	1.06***	1.04	1.83***
OBC (Dummy; Ref: Other)	1.27	1.19***	1.06	1.31**
Muslim (Dummy; Ref: Other)	1.08	1.26	1.01	1.07
Individual Characteristics:				
5 years of schooling (Dummy)	1.08***	1.32***	2.43***	1.42***
10 years of schooling (Dummy)	3.23***	0.66*	1.81***	0.81**
12 years of schooling (Dummy)	5.82***	0.53**	4.34***	0.72**
Undergraduate Degree (Dummy)	0.80	0.67	0.90	0.78
Masters Degree (Dummy)	0.94	0.91	1.10	1.08
Household Characteristics:				
Log of Consumption Expenditure	1.35**	1.06	1.17**	1.09
Household Size	0.93***	0.98*	0.95***	0.97*
Household Occupation: Self-employment	0.88	2.23***	0.94	1.51***
Household Occupation: Salaried	1.20	1.75	1.13	1.16
Household Occupation: Wage Work	0.79	1.11	1.02	0.91
Household Head's Schooling	1.01	0.98***	1.01	0.97***
Female Household Head (Dummy)	1.14	0.95	1.14	1.13
Context Characteristics:				
Number of TVET institutions	1.01*	1.00	1.01**	1.00
Unemployment Rate	1.08***	1.04	1.01***	1.04
Average 10-year rainfall	1.00	1.00	1.00	1.05
N	*2524*	*3188*	*1698*	*1372*

Source: Employment and Unemployment Survey of India (2009-10)

*p<.05, **p<.01, ***p<.001

4.4 Limitations

The analysis and results presented above suffer from two main limitations. These are discussed here along with details on the methods used to address some of the limitations.

The district level variables included in the empirical analysis (see Section 4.2.3) do not capture much variation in TVET enrollment decisions. The distribution of these variables across 50 randomly selected districts is presented in Figure A.7 in Appendix A. Although the plot shows significant variation in the supply of TVET institutions and in the unemployment rate across districts, and lesser variation in average rainfall across districts, these variables fail to explain significant variation in TVET enrollment in the multivariate analysis. Other context-level characteristics like sector-specific job growth in the region, district density, distance to TVET institutions and/or other educational institutions, and availability of infrastructure (roads and electricity) might have better served the analysis. The data used in this study, however, do not support the inclusion of these variables in the analysis.

Further, as discussed in Section 4.1 and shown in Table 1, previous evidence shows that TVET participation is affected by factors not included in the present analysis. These include occupational prestige, social networks, and the cost and quality of TVET options. The omission of these variables from the model introduces bias in the estimates. This omitted variable bias has been addressed to some extent with the use of district-level fixed effects. Fixed effects regression provides some control again bias due to omitted variables assuming that these omitted variables and their effects are time-invariant.

Second, the binary and multinomial models estimated have not been examined by urbanicity. The GLIMMIX procedure used in SAS did not converge when estimating models separately for urban and rural locations. These models are computationally intensive and non-convergence can arise due to several reasons; complexity of the model, misspecification or overspecification, scaling of data values, and so on. Attempts to simplify and use different estimation techniques, however, were not successful. All the multivariate analysis presented in Section 4.3 examined TVET participation along urban-rural dimensions (a dummy variable indicating urban location was included in all models). In most cases, urbanicity was found not to be a significant predictor of TVET participation over and above all the other variables in the model.[21] One can therefore make the case that separate models along urban-rural dimensions were not required.

The descriptive statistics showed some differences in TVET participation rates between 2004-05 and 2009-10. These differences are more pronounced in the case of informal TVET. As noted earlier (see Chapter 7 for a fuller discussion of differences in participation rates and predictive patterns), these differences could be a function of changes in the way TVET participation data have been collected in 2009-10. Attempts have been made to understand these differences by examining education and labor force participation rates during the same period. This is discussed further in Section 7.1 in Chapter 7.

[21] See the discussion in Chapter 7 on blurring urban-rural differences in India.

Chapter 5: Empirical Estimates of the Returns to Vocational Education and Training in India

The perceived benefits of education and training are a key determinant of participation decisions (Grubb, 1988). Since information on perceived benefits is often unavailable, data on actual benefits or returns is used to assess the efficiency of education and training systems with respect to costs (Billett, 1998). While private returns to general academic education have been examined extensively[22], relatively little attention has been paid to estimating the returns to TVET. The difficulty in assessing returns to TVET is related to the heterogeneity and variation in TVET offerings, and data constraints due to several contextual issues that impact the delivery and outcome of TVET programs (Hoeckel, 2008). This is especially true in the context of developing countries.

This chapter addresses this gap and estimates the returns to TVET in India. Building on findings from previous research in developing countries, the study models wages as a function of several individual, household, and context level variables. It corrects for selection bias using Heckman's selection correction (Heckman, 1979), as well as instrumental variables (Wooldridge, 2010) and compares the estimates of returns to TVET to the estimates of returns to general academic education in India.

The study focuses on the following research questions –

1. What are the returns to TVET in India, controlling for educational attainment, ability, and individual and household characteristics?

[22] See Section 3.5 in Chapter 3 and Duraisamy, 2002 for evidence of returns in the Indian case.

2. Controlling for individual and household characteristics, what are the returns to general education in India?

5.1 Data

The data used to examine the private, economic returns to individuals participating in vocational education and training programs came from the first round of the Indian Human Development Survey (IHDS). The IHDS is a nationally representative survey of urban and rural households across all states and union territories of India (except Andaman, Nicobar and Lakshwadeep islands) and includes 41,554 households and 215,754 individuals across 384 districts, 1503 villages and 971 urban blocks (Desai et al., 2008).

The fieldwork for the first round of the IHDS was carried out in 2004-2005 in 13 different languages, and involved surveying a knowledgeable informant, typically the male head of the household. The data include information on the socio-economic condition of the household, education and employment outcomes for all household members, and the extent of the household's social networks and relationships (Desai et al., 2008). The IHDS also includes a village-level questionnaire administered to a knowledgeable member of the village and includes information on infrastructure and resources, availability of health and educational facilities, and employment opportunities (Desai et al., 2008).

The household and individual-level data in conjunction with contextual information at the village-level are used to estimate returns to vocational education and training. These separate data files were linked in two stages. In the first stage the

household and individual-level data were linked using an 11-digit unique code (IDHH) identifying the state, district, primary sampling unit (PSU), and household and individual. In the second stage, the merged household and individual file was linked to the village-level file using a derived 'village ID'. The village ID was created by concatenating the state, district and PSU identifiers.

Table 7 provides a description of the demographic, individual, household and village level covariates used to estimate returns to academic and vocational education.

Table 7

Description of variables used to estimate returns to TVET and general education

Variable	Source (IHDS 2004-2005)	Description
Log annual wages	Household/Individual data file	Natural log of annual wages/earnings
DEMOGRAPHIC INDICATORS		
Age	Household/Individual data file	Age in years
Age squared	Household/Individual data file	The quadratic term for age
Age group 1	Household/Individual data file	Includes ages 15 to 25
Age group 2	Household/Individual data file	Includes ages 26 to 40
Age group 3	Household/Individual data file	Includes ages 41 to 65
Female	Household/Individual data file	Dummy variable for gender - coded '1' for female and '0' for male
Urban	Household/Individual data file	Dummy variable for geographic location - coded '1' for urban and '0' for rural
Marital status	Household/Individual data file	Dummy variable indicating marital status - coded '1' if married at the time of survey and '0' if otherwise
Other Backward Class	Household/Individual data file	Dummy variable indicating social group - coded '1' if OBC and '0' otherwise
Dalit	Household/Individual data file	Dummy variable indicating social group - coded '1' if Dalit and '0' otherwise

Adivasi	Household/Individual data file	Dummy variable indicating social group - coded '1' if Adivasi and '0' otherwise
Muslim	Household/Individual data file	Dummy variable indicating religious affiliation - coded '1' if Muslim and '0' otherwise
INDIVIDUAL CHARACTERISTICS		
Years of schooling	Household/Individual data file	Continuous variable indicating years of schooling
BA	Household/Individual data file	Dummy variable indicating completion of undergraduate/Bachelor's degree – coded '1' if completed an undergraduate degree and '0' otherwise
MA	Household/Individual data file	Dummy variable indicating completion of Master's degree – coded '1' if completed an undergraduate degree and '0' otherwise
Professional	Household/Individual data file	Dummy variable indicating completion of a Professional degree – coded '1' if completed an undergraduate degree and '0' otherwise
TVET	Household/Individual data file	Dummy variable TVET participation - coded '1' if participated in TVET after grade 10 and '0' if not
Grade 10 Performance (1st Division)	Household/Individual data file	Dummy variable indicating performance on grade 10 - coded '1' if respondent scored in the highest division and '0' otherwise
Grade 10 Performance (2nd Division)	Household/Individual data file	Dummy variable indicating performance on grade 10 - coded '1' if respondent scored in the second division and '0' otherwise
Unemployed	Household/Individual data file	Dummy variable indicating employment status - coded '1' if respondent is unemployed and '0' otherwise

Salaried	Household/Individual data file	Dummy variable indicating employment status - coded '1' if respondent is a salaried worker and '0' otherwise
Working in household enterprise	Household/Individual data file	Dummy variable indicating employment status - coded '1' if respondent works in a household enterprise (business, farm or non-farm) and '0' otherwise
Informal sector worker	Household/Individual data file	Dummy variable indicating employment status - coded '1' if respondent is an informal sector worker and '0' otherwise
HOUSEHOLD CHARACTERISTICS		
Gender of head of the household	Household/Individual data file	Dummy variable indicating the household head's gender - coded '1' if female and '0' otherwise
Head of the household's years of schooling	Household/Individual data file	Continuous variable indicating years of schooling of the head of the household
Head of the household’s TVET participation	Household/Individual data file	Dummy variable indicating if the head of the household participated in TVET – coded ‘1’ if he/she participated in TVET and ‘0’ otherwise
Household size	Household/Individual data file	Continuous variable indicating size of the household (number of persons living in the household for a continuous period of 30 days or more)
Number of children	Household/Individual data file	Continuous variable indicating the number of children (below the age of 15 years) living in the household
Household Assets	Household/Individual data file	Continuous variable indicating the assets owned by the household. The variable ranges from 0 to 30.
VILLAGE CHARACTERISTICS		
Distance to higher secondary school	Village level data file	Continuous variable indicating distance of higher secondary school from village (in

		kilometers) if higher secondary school not available in the village
Distance to college	Village level data file	Continuous variable indicating distance of college from village (in kilometers) if no college available in the village
Distance to TVET institution	Village level data file	Continuous variable indicating distance of vocational institution from village (in kilometers) if no vocational institution available in the village
Number of higher secondary schools	Village level data file	Continuous variable indicating number of higher secondary schools (public and private) in the village
Number of TVET institutions	Village level data file	Continuous variable indicating number of TVET institutes (public and private) in the village
Weight	Household/Individual data file	Probability weights to account for sampling design
District ID	Household/Individual/Village	Unique ID for districts in the sample
Village ID	(Derived variable)	Unique ID for villages in the sample

5.1.1 Analytic Sample (For returns to general education)

The analytic sample for this study was limited to individuals between 15-65 years of age, constituting the working age population in India (United Nations, 2010). The lower bound of 15 years was motivated by the fact that vocational education and training programs in India can be accessed as early as high school and programs can be pursued as a part-time option. Of the 215,754 cases in the IHDS data, 138,776 or nearly 55 percent were between 15-65 years. Of the remaining, 889 cases (about 0.6 percent) were missing data on years of completed schooling, 329 cases (about 0.2 percent) were

missing data on head of the household's schooling, and 3,457 (2.5 percent) were missing information on English fluency. Observations with missing information on schooling were deleted from the sample.[23]

Information on wages (the dependent variable) is central to estimating returns and the IHDS data were chosen for the extensive coverage of various types of income-generating activities. The survey focuses on various sources of household income including farm and non-farm work, self-employment activities, agricultural and non-agricultural labor, and salaried work. The survey gathers information on each household member's participation in any of the listed activities in addition to other wage/salaried work the individual might be engaged in outside of household-based work. Thus, each individual in the dataset could have multiple sources of income.

In order to accurately represent the wages of individuals in the analytic sample it was important to account for income generated across various activities. As per the IHDS survey, an individual was classified as "working" in any given activity if he/she was engaged in that activity for 240 hours (about one month of full-time work) or more during the previous year. For individuals engaged in waged and salaried work, annual earnings were available for each individual. But in the case of individuals engaged in farm work, non-farm work and/or other household businesses, the total annual income was only available at the household level. In these cases, (i.e. for those engaged in any household-based activity), individual earnings were calculated by dividing the total income from that activity by the number of persons within the household engaged in that

[23] Observations missing information on English language ability were retained in the analytic sample. English fluency and its relationship with earnings is not a focus of this study. Models that include English fluency on the RHS have an effective sample size of 78572 cases.

activity. Thus, for example, for an individual engaged in agricultural labor and non-farm work, the total annual earnings included those acquired from agricultural wage labor as well as his/her share of earnings from non-farm work.

About two percent of households in the IHDS survey reported a negative total income due to losses or failed crops in the previous year. After computing annual individual income (i.e. aggregating earnings across all income-generating activities) for all cases in the analytic sample, those cases with a negative annual income were deleted from the sample (2.6 percent of 15-65 year olds).

The trimmed analytic sample thus includes 134,036 cases of which 40.25 percent are unemployed. The distribution of this sample by type of economic activity is presented in Table 8. About 70 percent of the sample is located in the rural sector where 34 percent of those employed are engaged in farm or non-farm work and/or other household businesses. In urban areas, those engaged in some kinds of household-based work constitute just four percent of the sample. Salaried workers form the largest employed group in urban areas making up about six percent of the sample. The proportion of casual labor is higher in rural areas than urban areas (10 percent versus three percent).

As seen in Table 8 the proportion of males and females in the sample is about the same. Nearly 28 percent of females in the sample are unemployed whereas 11 percent of the males report unemployment. The proportion of males is higher across all types of income-generating activities – salaried work, casual labor, and household-based enterprises. The skewed gender distribution is indicative of the reported gender discrimination in the labor market in India (Kingdon, 1997).

Table 8

Proportion of analytic sample by employment status, gender and sector

Employment Status	Rural	Urban	Male	Female	Total
Unemployed	0.23	0.16	0.11	0.28	0.39
Salaried	0.03	0.06	0.08	0.02	0.09
Casual Worker	0.11	0.03	0.10	0.04	0.14
Household Enterprise Worker	0.34	0.04	0.22	0.16	0.38
Total	0.71	0.30	0.51	0.50	1.00
N					*134,036*
N (PSUs)					*2,473*

Source: Indian Human Development Survey, 2004-05.

The analytic sample used to estimate returns to general education includes all cases with reported earnings. Figure 17 shows the distribution of *log wages* in the analytic sample. The distribution is approximately normal but skewed to the left. It should be noted that four cases in the analytic sample showed extreme values on annual income.[24] These cases were above the 75th percentile of the wage distribution. In keeping with standard econometric practice, the annual wages for these four cases were recoded to the fifth largest value at the 75th percentile.

[24] See Figure B.1 in Appendix B for a boxplot of annual wages in the untrimmed sample.

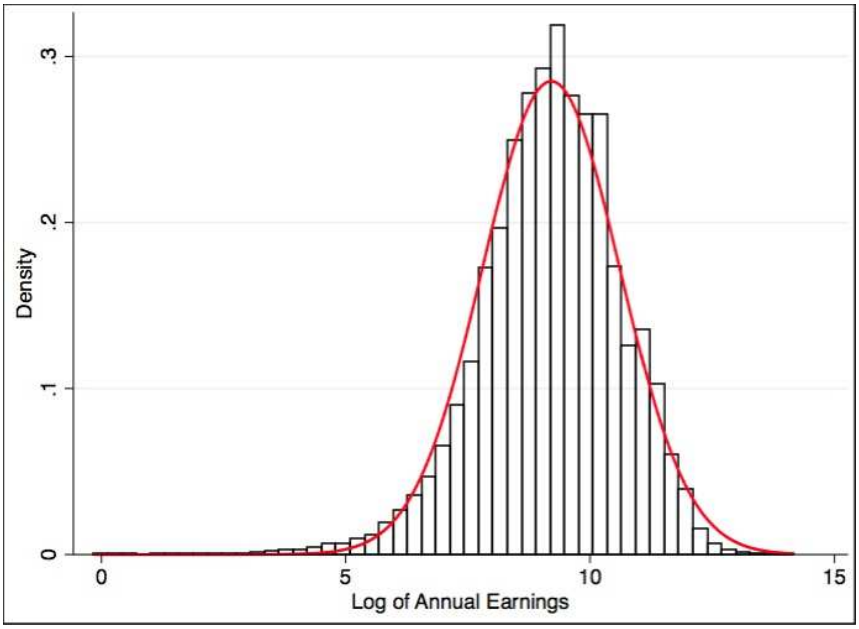

Figure 17. Distribution of log annual earnings among 15-65 year olds across all types of occupation.
Source: Indian Human Development Survey, 2004-05.

Table 9 shows the mean values for key variables used in the empirical analysis of returns to schooling by rural and urban sector. There is significant variability in annual wages across the sample. The average across the entire sample is about 29,300 (in Indian Rupees) with the average for rural areas at 13,213 and the average for urban areas at 45,563.

Table 9

Weighted means of predictors of annual wages among 15-65 year olds, by location

	Rural		Urban	
	Mean	SE (Mean)	Mean	SE (Mean)
Annual Income	13213.81	260.055	45563.87	1076.482
Log Annual Income	8.79	0.018	10.18	0.024

Years of schooling	4.40	0.060	7.89	0.103
Age	36.13	0.090	36.97	0.113
Age Squared	1474.33	7.026	1504.41	8.568
Age-Between 15-21 years)	0.15	0.003	0.09	0.003
Age-Between 22-28 years)	0.19	0.003	0.19	0.003
Age-Between 29-39 years)	0.27	0.003	0.31	0.004
Age-Between 40-65 years)	0.40	0.003	0.42	0.004
Female (Dummy; Ref: Male)	0.40	0.003	0.21	0.004
Marital Status (Dummy; Ref: Unmarried)	0.77	0.004	0.77	0.004
Social Group - OBC (Dummy)	0.38	0.011	0.32	0.010
Social Group - Dalit (Dummy)	0.24	0.008	0.18	0.009
Social Group - Adivasi (Dummy)	0.11	0.008	0.03	0.004
Religious Group - Muslim (Dummy)	0.09	0.007	0.15	0.009
Ability (Dummy; >60% in Grade X)	0.02	0.002	0.12	0.005
Ability (Dummy; < 60% in Grade X)	0.09	0.003	0.23	0.006
English Fluency	0.13	0.005	0.35	0.009
Head of the Household's Schooling	4.01	0.064	7.19	0.107
Household Size	6.23	0.055	5.50	0.047
Number of children in the household	2.03	0.033	1.51	0.027
Household Assets	9.29	0.097	16.13	0.139
N	*57,752*		*20,820*	

Source: Indian Human Development Survey, 2004-05.

In terms of age, the average age across urban and rural residents is about the same – 36 years. Surprisingly, the proportion of females in rural areas is almost double (40 percent) of that in urban locations (21 percent).

There are significant differences in the schooling outcomes across rural and urban residents. While the average years of completed schooling in rural areas is only four years, urban residents report an average of eight years of schooling. Similarly, on average, eight percent of rural residents have a high school degree and three percent have a college degree. The corresponding figures in urban areas are 27 percent and 17 percent, respectively.

The education of the head of the household also follows a similar trend; on average, rural households report that the head of the household has four years of schooling while urban households report that the head of the household has completed seven years of schooling.

The distribution of social groups across urban and rural locations is similar except for the *Adivasi* groups that tend to be located largely in rural areas (11 percent of the rural sample) than urban areas (three percent of the urban sample); and *Muslims* who form a larger proportion of the urban sample (15 percent) than the rural sample (nine percent).

Finally, the average household size is slightly larger in rural areas (6.23 persons) than in urban areas (5.49 persons), as is the number of children in the household (1.51 in urban areas and 2.03 in rural areas).

In order to better understand the heterogeneity in annual earnings, the log earnings were plotted by education level, separately for the two gender groups, and across urban and rural dimensions. The boxplots in Figures 18 and 19 show these distributions.

Research has consistently found that in the Indian case, female earnings are significantly lower than those of males, across locations, and notwithstanding education levels (Kingdon, 1997). This is evidenced in the figures below. Although the wage differences between men and women reduce at higher levels of education (over 12 years of schooling), females continue to earn significantly less than their male counterparts.

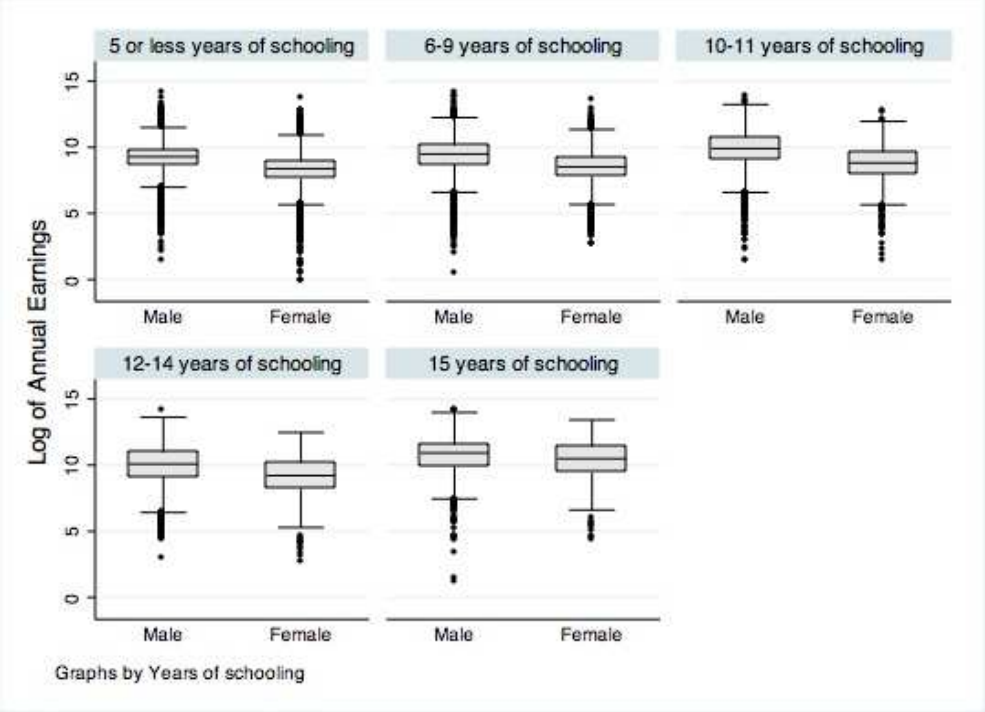

Figure 18. Distribution of log annual earnings among 15-65 year olds, by gender and education level.
Source: Indian Human Development Survey, 2004-05.

Figure 19 shows wage distributions for men and women by urban and rural locations. In rural areas, both men and women, on average, have lower wages than men and women in urban areas. Wages in rural areas also show a higher degree of variability than urban wages.

The wage distributions by education levels and urban and rural status, as seen in Figure 20 below, show that the urban-rural wage gap is widest amongst those with 10-15 years of education. As the educational attainment goes up, urban wages increase notably. In rural areas however, increasing years of schooling are not associated with the same increase in wages.

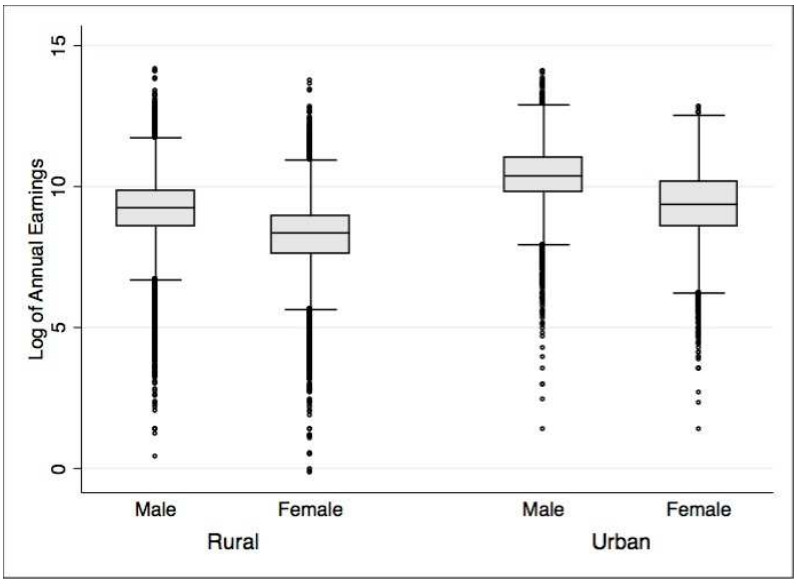

Figure 19. Distribution of log annual earnings among 15-65 year olds by gender and urban-rural status.
Source: Indian Human Development Survey, 2004-05.

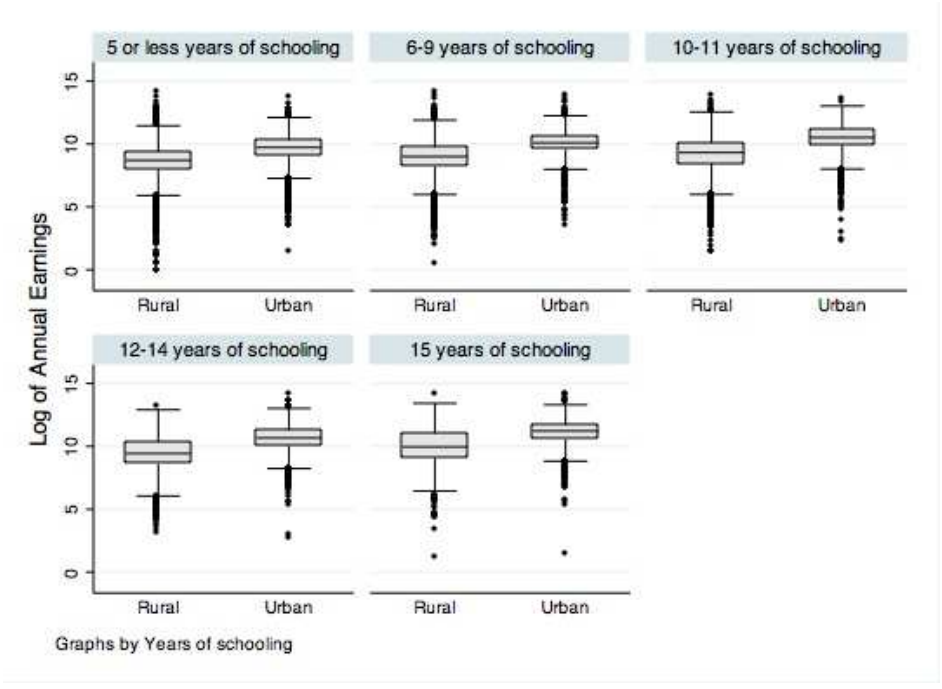

Figure 20. Distribution of log annual earnings among 15-65 year olds by education level and urban-rural status.
Source: Indian Human Development Survey, 2004-05.

5.1.2 Analytic Sample (For returns to TVET)

To estimate returns to TVET only those cases reporting participation in a vocational or technical education program were selected. A binary variable indicating TVET participation was derived using two questions; subject of study at the postsecondary level and highest level of education completed. These questions were asked to a subset of survey respondents; the first question on postsecondary subject was asked to all respondents with 10 or more years of education, and the second question was asked to respondents who attended college. Both questions included "vocational" as one of the response options. The TVET indicator thus created included cases that had participated in TVET at any point after grade 10.

Cases with less than ten years of schooling (about 20 percent of the sample had 10 or more years of schooling) were excluded from the sample. Cases missing information on postsecondary subject and highest level of education were also removed from the sample (24,100 cases). Vocational education and training participants constituted a significantly small proportion of the sample – 0.3 percent.

The same steps as noted in Section 5.1.1 were followed to further trim the sample. Cases outside the working age range (985 cases), those missing the household head's level of education, and those with negative total earnings (558 cases) were removed from the sample. This resulted in a sample size of 15,270 cases half of which were employed.

Schooling or education was defined in terms of education levels unlike in the previous case where it was defined as a continuous variable measuring years of completed schooling. In addition to the variable indicating TVET participation, three

dummy variables were created to indicate completion of a Bachelor's degree, a Master's degree, or a Professional degree.

The distribution of log annual earnings in the analytic sample is presented below in Figure 21. Appendix B includes boxplots of annual wages with extreme values in the untrimmed sample. See Figures B.3 and B.4 in Appendix B.

Figure 21 shows that the earnings for the TVET sample are left skewed indicating that a large proportion of the sample reported low annual earnings. Other than the left skew, the distribution is approximately normal.

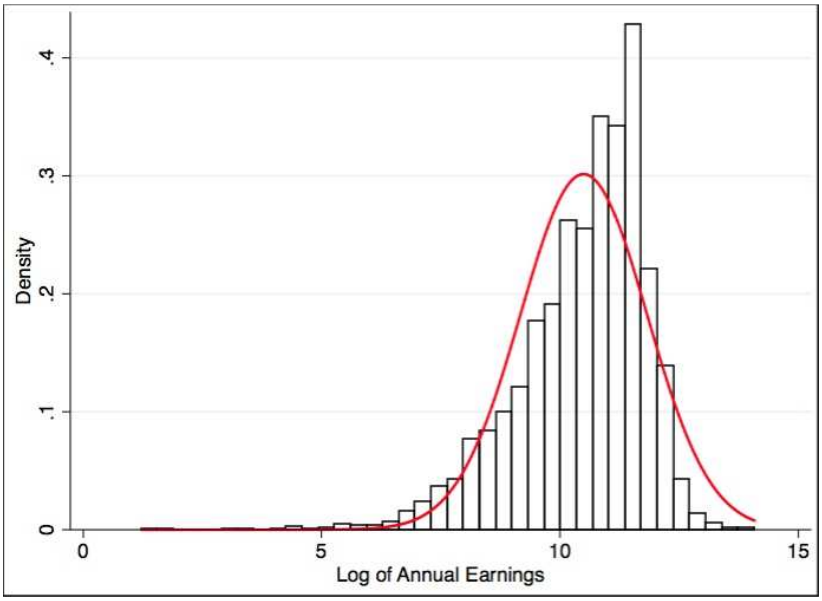

Figure 21. Distribution of log annual earnings among 15-65 year olds with 10 or more years of schooling.
Source: Indian Human Development Survey, 2004-05.

The proportion of the analytic sample participating in TVET is presented in Table 10. The weighted proportions are presented along gender and sectoral dimensions, as well as by employment status.

Table 10

Weighted proportion of TVET participants in the sample, by gender, sector and employment status

	No TVET	Some TVET	Total
Male	0.596	0.056	0.651
Female	0.336	0.013	0.349
Rural	0.396	0.032	0.427
Urban	0.536	0.037	0.573
Unemployed	0.436	0.033	0.469
Salaried Worker	0.240	0.018	0.259
Casual Worker	0.030	0.003	0.034
Household Enterprise Worker	0.224	0.015	0.239
Total	0.931	0.069	1.000
N			*15,270*
N (PSUs)			*1,999*

Source: Indian Human Development Survey, 2004-05.

Males constituted about 65 percent of the analytic sample of which five percent participated in some type of TVET. About one percent of the females in the sample participated in TVET. The proportion of TVET participants across rural and urban sectors was about the same (just over three percent). In terms of employment status, about 47 percent of the analytic sample reported being unemployed, followed by 26 percent in salaried work, 24 percent working in a household enterprise, and 3.4 percent in wage work. In terms of employment status of TVET participants, the largest proportion reported being unemployed (about three percent). Of those employed, the majority were in salaried work, followed by self-employment, and a very small proportion in casual wage work.

Table 11 shows weighted means for the variables used in the analysis by urban and rural sectors. There is significant difference in annual earnings reported across rural and urban areas. The average annual income in urban areas is more than twice the average in rural areas, with urban income showing more variation than rural incomes.

The distribution of urban and rural residents by education level follows an expected pattern with Bachelor's degree holders forming the largest group, followed by those with a Master's degree, and the smallest proportion with a professional degree. In urban areas, TVET participants are the smallest group making up nearly seven percent of the sample. In rural areas, professional degree holders form the smallest group (three percent of the sample), followed by TVET participants who again constitute seven percent of the sample.

The urban and rural samples also show significant differences in average age. Rural residents are, on average, about three years younger than their urban counterparts. When examined by various age groups, the biggest differences are observed among the 15-21 year olds and 40-65 year olds. In rural areas, 9.6 percent and 27 percent of the sample fall within 15-21 years and 40-65 years respectively. The corresponding figures for urban areas are 3.6 and 41 percent, respectively.

The proportion of females in the urban and rural sample is about the same – 15 percent. The distribution of social religious groups shows some differences across rural and urban locations – *OBCs* constitute about 38 percent of the rural sample and nearly 27 percent of the urban sample; *Dalits* make up about 14 percent of the rural and 8.7 percent of the urban sample; and *Adivasis* constitute about five percent of the rural sample and

2.7 percent of the urban sample. The proportion of *Muslims* in rural and urban areas is about the same.

With regard to the ability measures, there are significant differences in the proportion achieving more than 60 percent marks in grade 10 and the proportion fluent in English across urban and rural areas. As expected, the proportions are higher in urban locations than rural areas (38 percent versus 18 percent in the case of grade 10 scores and 84 percent versus 73 percent in the case of English language fluency). There is little difference in the proportion achieving less than 60 percent in grade 10 across sectors.

Finally, in terms of household characteristics, the average household size and the number of children in the household are similar across rural and urban sectors. Urban residents report, on average, higher household assets than rural residents and higher education levels for the head of the household.

Table 11

Weighted means of key variables used to predict returns to TVET among 15-65 year olds

	Rural		Urban	
	Mean	SE (Mean)	Mean	SE (Mean)
Annual Income	35321.10	1290.26	86403.61	2317.83
Log Annual Income	9.719	0.040	10.977	0.027
BA Degree (Dummy)	0.406	0.016	0.484	0.011
MA Degree (Dummy)	0.109	0.010	0.164	0.008
Professional Degree (Dummy)	0.036	0.006	0.082	0.006
TVET (Dummy)	0.072	0.008	0.068	0.005
Age	33.804	0.301	37.346	0.207
Age Squared	1262.186	23.372	1505.600	16.114
Age-Between 15-21 years	0.096	0.007	0.036	0.003
Age-Between 22-28 years	0.287	0.012	0.201	0.007
Age-Between 29-39 years	0.348	0.012	0.354	0.009
Age-Between 40-65 years	0.269	0.011	0.410	0.010

Female (Dummy)	0.156	0.010	0.155	0.006
Marital Status (Dummy)	0.698	0.014	0.776	0.008
Social Group - OBC (Dummy)	0.379	0.020	0.268	0.012
Social Group - Dalit (Dummy)	0.145	0.014	0.087	0.008
Social Group - Adivasi (Dummy)	0.054	0.008	0.027	0.005
Religious Group - Muslim (Dummy)	0.080	0.010	0.072	0.008
Ability (> 60% in Grade 10)	0.187	0.013	0.387	0.013
Ability (< 60% in Grade 10)	0.556	0.016	0.505	0.011
English Fluency	0.732	0.015	0.846	0.010
Head of the Household's Schooling	9.279	0.160	12.348	0.091
Head of the Household's TVET Participation	0.033	0.005	0.044	0.004
Household Size	6.523	0.127	5.197	0.067
Number of children in the household	1.780	0.065	1.172	0.031
Household Assets	14.576	0.180	20.816	0.134
N				*7,877*
N (PSUs)				*1,818*

Source: Indian Human Development Survey, 2004-05.

In order to further examine the variation in annual earnings across the sample the earnings were graphed by education level and along gender and sector dimensions. The distribution of earnings for TVET participants by gender is presented in Figure 22. Plots for Bachelor's, Master's and Professional degree holders can be found in Appendix B; see Figures B.3, B.4, and B.5.

Figure 22 and plots for other degree holders show approximately normal distributions of log annual earnings by gender. The distributions are left skewed and in some cases leptokurtic (kurtosis=4. 83). The distribution of log annual earnings for female TVET participants, however, does not fit the normal distribution.

Figure 23 shows boxplots of log annual earnings for TVET participants by urban-rural location. The plot indicates that urban residents, on average, irrespective of TVET participation, show higher earnings than their rural counterparts. Within urban and rural

areas, TVET participants show a slight disadvantage in earnings when compared to non-TVET participants. Earnings of those not in TVET show significantly greater variation than the earnings of TVET participants.

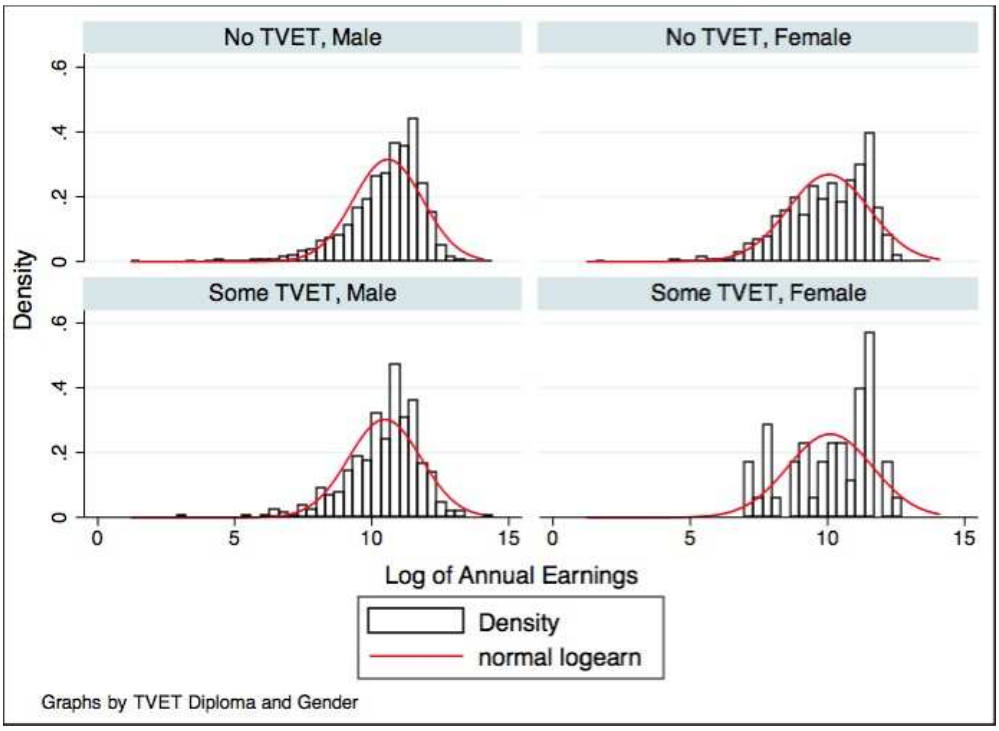

Figure 22. Distribution of log annual earnings by gender and TVET status among 15-65 year olds with 10 or more years of education.
Source: Indian Human Development Survey, 2004-05.

In Figure 24, log annual earnings are plotted by education/training. The figure shows that average earnings increase slightly with each additional credential. Earnings for those without a TVET or higher credential show the lowest mean earnings and those with a professional degree have the highest mean earnings. The variation in earnings is significant amongst those without a credential and those with a Bachelor's degree.

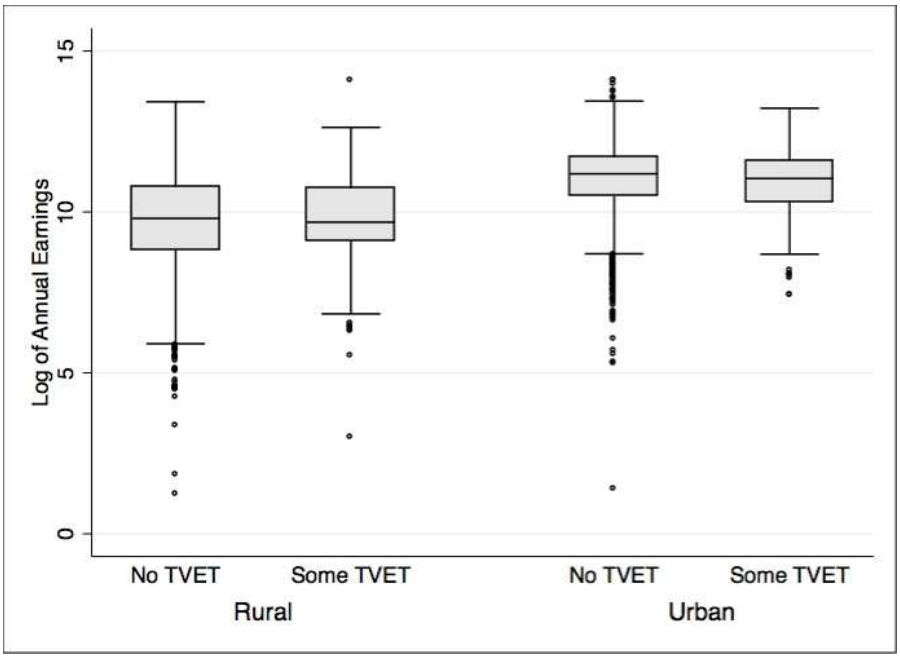

Figure 23. Boxplot of log annual income of TVET and non-TVET participants, by urban-rural location. The analytic sample includes 15-65 year olds with 10 or more years of education.
Source: Indian Human Development Survey, 2004-05.

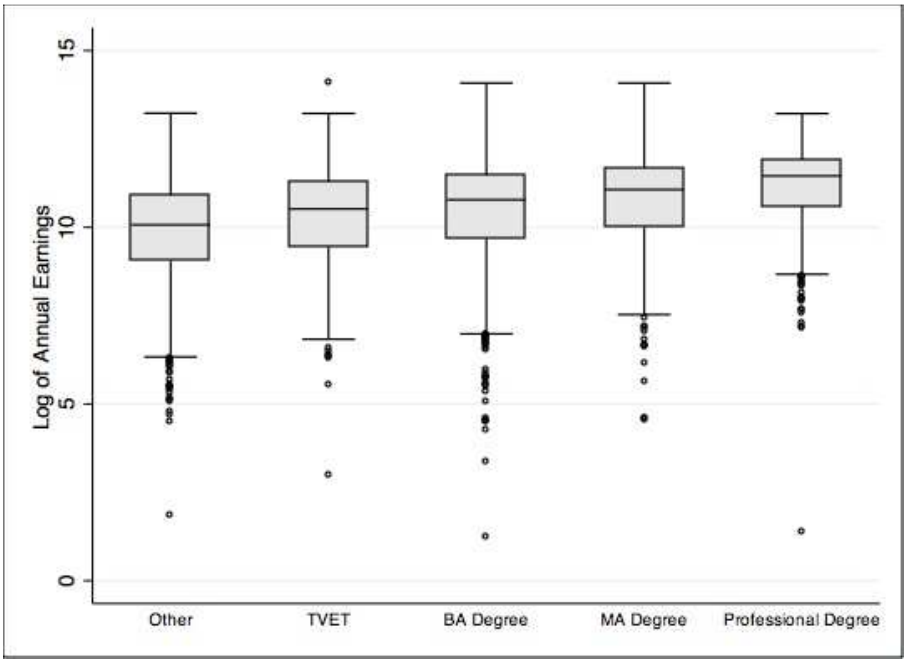

Figure 24. Boxplot of log annual income by education/training among 15-65 year olds with 10 or more years of education.
Source: Indian Human Development Survey, 2004-05.

5.2 Analytic Methods

The estimation of returns in this paper was based on the standard Mincerian approach of estimating wage functions to compute rates of return to education (Mincer, 1975). The relationship between wages and years of schooling, and wages and vocational education is expressed as:

$$lnW_i = \alpha + \beta_1 S_i + \gamma_1 A_i + \gamma_2 A_i^2 + \delta_{1k} \sum_{k=1}^{N} X_{1ik} + \delta_{2k} \sum_{k=1}^{N} X_{2k} + \delta_{3k} \sum_{k=1}^{N} X_{3k} + u_i$$

$$lnW_i =$$

$$\alpha + \beta_1 S_i + \beta_2 V_i + \gamma_1 A_i + \gamma_2 A_i^2 + \delta_{1k} \sum_{k=1}^{N} X_{1ik} + \delta_2 \sum_{k=1}^{N} X_{2k} + \delta_{3k} \sum_{k=1}^{N} X_{3k} + u_i$$

In the first equation, lnW_i is the log of hourly wages for individual i, S_i is years of schooling, A_i and A_i^2 represent age (in years) and it's quadratic, X_{1ik} is a vector of observed individual characteristics, X_{2k} is a vector of observed household characteristics, X_{3k} is a vector of observed district-level characteristics, and u_i represents the individual-specific error. In the second equation, V_i represents participation in vocational education and takes a value of 1 if an individual participates in TVET and 0 otherwise.

Ordinary Least Squares (OLS) provides unbiased estimates of the coefficient on schooling and vocational education if the error term is uncorrelated with each of the regressors. However, in the case of wage functions, OLS estimates can significantly over or underestimate the effect of schooling on wages (Card, 2001). The overestimation is a

result of endogeneity of the schooling variable, while underestimation is attributed to measurement error in years of schooling.

According to Card (2001), there are three sources of endogeneity – omitted variables, simultaneity, and measurement error. It can be argued that ability, unobserved in the equations above, is a determinant of years of schooling (and wages), and its absence in the equation results in inconsistent estimates of returns. The coefficient of education represents the causal effect of education on wages only when observed differences in wages can be attributed to varying years of schooling and not any underlying, unobserved differences in ability.

Self-reported measures of education often include errors due to various reasons; social desirability, inaccurate memory, and so on. The difference between the true value and the reported or measured value is called measurement error. Within the OLS framework, measurement error in years of schooling (i.e. the difference between the true level of education and the reported level of education) has been shown to be correlated with observed years of schooling causing significant attenuation of the OLS estimate on schooling (Wooldridge, 2010).

Finally, as wages are only observed for those employed in the labor force, estimates of returns to education are based on a non-random sample of the population. This results in sample selection bias and inconsistent OLS estimates (Wooldridge, 2010). It can be shown that sample selection bias is similar to the bias from omitted variables, and can be addressed by least squares methods (Heckman, 1979).

The extant literature on returns to education has employed various techniques to address inconsistencies in OLS estimates caused due to endogeneity. Card (2001) reviews

these and finds that 15 percent of the reviewed studies used Heckman's two-step correction, while 80 percent used instrumental variables. None of the reviewed studies used repeated observations or household fixed effects methods.

This study employs Heckman's correction and IV methods in estimating the returns function. Household fixed effects and repeated observations, although a possible solution to endogeneity, cannot be used due to data limitations.[25]

5.2.1 Heckman Selection Correction

One of the key assumptions underlying regression equations is sample randomness. When this assumption is violated due to nonrandom missing observations on the dependent variable, the coefficient estimates are biased. The intuition behind Heckman's correction for sample selectivity or selection bias is to construct a model that jointly represents the regression equation to be estimated, as well as the process that determines if the dependent variable is observed (Olsen, 1980).

In the case of this study, in equation (1) and (2), wages are only observed for those currently employed in the labor force i.e. where $W_i > 0$. Employing the Heckman correction entails estimating the probability of 'labor force participation' for the sample, followed by estimating the returns while controlling for selection, which is equivalent to addressing selection on observables. More specifically, in the first step, probit regression is used to estimate the propensity of being "waged" based on a vector of explanatory variables. This equation is the selection equation and can be formally represented as:

$$Prob(LFP = 1|Z) = Prob(\varepsilon_i > -Z_i\gamma_1) = \Phi(Z_i\gamma_1)$$

[25] 85 percent of the analytic sample represents households with one observation

Here, labor force participation (*LFP*) is a latent binary indicator of being employed in paid work and depends on a vector of explanatory variables *Z*. The explanatory variables in *Z* are different from those included in the vector *X* described in equations (1) and (2) and include household size, number of children in the household and household assets. In the equation above, Φ represents the standard cumulative distribution function (C.D.F) and γ_1 represents the associated parameter vector.

The predicted probabilities resulting from the selection equation are used to compute the 'Inverse Mills Ratio' or *lambda*, which is added to the returns equation as an additional explanatory variable.

The wage equation is then represented as:

$$lnW_i|LFP_i > 0 = \alpha + \beta_i S_i + \gamma_1 A_i + \gamma_2 A_i^2 + \delta_1 \mathrm{X}_{1i} + \delta_2 \mathrm{X}_2 + E(u_i|\varepsilon_i > -Z_i\gamma_1)$$

$$= \alpha + \beta_i S_i + \gamma_1 A_i + \gamma_2 A_i^2 + \delta_1 \mathrm{X}_{1i} + \delta_2 \mathrm{X}_2 + \rho\sigma_u \lambda_i + error$$

The null hypothesis that the coefficient on the selectivity term, lambda (λ_i), is zero provides a test for sample selectivity (Heckman, 1979; Wooldridge, 2010). If the null is rejected, it suggests there is sample selection bias.

5.2.2 Instrumental Variables

Endogeneity causes one or more explanatory variables to be correlated with error terms in a regression equation. The instrumental variable (IV) approach to addressing endogeneity is based on introducing an *instrument* or instrumental variable in the regression equation that is correlated with the endogenous regressor conditional on the

other covariates in the model. Weak correlation between the IV and endogenous regressor results in a larger bias and inconsistency in the IV estimates than that obtained using OLS (Murray, 2006). The Kleibergen-Paap Wald statistic is used as a test for validity of IVs and is robust to the presence of heteroskedasticity, autocorrelation, and clustering (Kleibergen & Paap, 2006).

Further, the IV must be uncorrelated with the error term in the second stage regression. A test of over-identifying restrictions – the Hansen J statistic – is used as a test. It should be noted that when multiple IVs share a common rationale, the over-identifying restrictions test might not be meaningful.[26]

Previous studies that have used the IV approach in estimating returns to education have included natural experiments as well as nonexperimental IVs such as family background variables (Card, 2001). This study uses a combination of family background variables (for example, years of schooling of the head of the household and gender of the head of the household) as well as contextual indicators that capture variation at the local level (for example, proximity to various levels of schooling, and the supply of educational institutions).

Equation (1) includes one endogenous regressor, years of schooling, whereas equation (2) includes two endogenous regressors – years of schooling and vocational education. To ensure identification, the number of IVs exceeded the number of endogenous variables in equation (2). The two-stage least squares approach to IV estimation was adopted.

[26] The instruments proposed in the case of this analysis do not share a common rationale. As discussed below, supply-side indicators, household indicators, and policy shifts will be considered as possible instruments.

In the case of equation (1), the first stage involved regressing years of schooling on the instruments and the other exogenous predictors from equation (1). This was formalized as:

$$S_i = \beta_0 + \beta_1 A_i + \beta_2 A_i^2 + \beta_{3k} \sum_{k=1}^{N} X_{1ik} + \beta_{4k} \sum_{k=1}^{N} X_{2k} + \beta_{5k} \sum_{k=1}^{N} Z_{ik} + \varepsilon$$

where, Z_{ik} represents the vector of instruments. The predicted values from the first stage were then used in equation (1) to estimate returns. Similarly, in the case of equation (2), the linear projection of schooling and vocational education was used to estimate returns to vocational education.

5.2.3 Other Methods

In addition to the Heckman procedure and instrumental variable estimation, repeated measures and household fixed effects have been used to address endogeneity (Card, 1999) and selection bias (Behrman & Deolalikar, 1995). Repeated observations on the same individual over time or observations from multiple individuals within the same household/family are used within a fixed effects approach. The assumption underlying these approaches is that differences in unobserved ability are smaller within households than between households. The fixed effects method controls for sources of variation at the household level and the unobserved heterogeneity common to individuals within a household.

The data used to estimate the returns function in this paper does not support either of these approaches. These data are cross-sectional and therefore do not include repeated

observations on individuals. Further, these data cannot be used within a household fixed effects approach due to sample size limitations.

Table 12 lists the variables used in the various estimation models, by method.

Table 12

Variables used in the analysis of returns by different analytic methods

Variables	Analytic Methods		
	OLS	Heckman	2SLS
Dependent/2nd Stage	Log annual wages	Log annual wages	Log annual wages
Dependent/1st Stage		Wage work	Schooling/TVET
Predictors:			
Vocational Participation	TVET dummy	TVET dummy	TVET dummy
Education	Completed years of schooling	Completed years of schooling	Completed years of schooling
Ability	Grade 10 performance dummy variables	Grade 10 performance dummy variables	Grade 10 performance dummy variables
Controls	Dummy for age group 1	Dummy for age group 1	Dummy for age group 1
	Dummy for age group 2	Dummy for age group 2	Dummy for age group 2
	Female dummy	Female dummy	Female dummy
	Age*Female	Age*Female	Age*Female
	Marital status	Marital status	Marital status
	Marital status*Female	Marital status*Female	Marital status*Female
Household Characteristics	Dummy variables social group	Dummy variables social group	Dummy variables social group
	Dummy variables for religious group	Dummy variables for religious group	Dummy variables for religious group
Instruments		Lambda/IMR Household size Number of children in the household Household Assets	Number of educational institutions (schools, TVET options, colleges) in village Head of the household's schooling Head of the household's TVET participation

5.3 Results

This section reports the results of the empirical analysis described in the previous section. Returns to general education were estimated using three methods – OLS, Heckman's selection correction and IV estimation. These are reported and compared in Section 5.3.1. The returns to TVET were estimated using OLS and Heckman's sample selectivity correction methods and are reported in Section 5.3.2.[27]

5.3.1 Returns to schooling

The returns to schooling were estimated for the working age group in India (15 to 65 years). Table 13 presents the marginal effects of schooling, controlling for demographic dimensions (age, gender, urban-rural status, marital status, and social class), ability and English language fluency. Three methods were used to estimate log annual wages for the sample: OLS, Heckman's selection correction method and IV estimation. The OLS results are discussed first, followed by the estimates using the Heckman and IV methods.

Table 13 shows that controlling for the individual and household characteristics noted above, OLS estimates a 3.5 percent increase in log annual earnings for each additional year of schooling. These estimates are consistent with those found in other studies estimating returns to schooling in India (Agrawal, 2011; Azam, et al., 2010).

As discussed in *Section 5.2,* OLS estimates of earnings are biased and therefore unbiased estimates of earnings are estimated using alternate methods. The Heckman

[27] Instrumental Variable estimation was not used for returns to TVET. See Section 5.3.2 and Section 5.4 for an explanation.

estimates presented in the table below address non-randomness of the sample and correct for selection.[28] The estimate and standard error of the selection index (*rho*) is shown in the table below and indicates the presence of selection bias in the model (thus making the case for sample selection methods). The results show that the returns to an additional year of schooling reduce to 2.9 percent when estimated using this method.

The instrumental variables approach to addressing omitted variable bias and endogeneity is considered more robust than the approach suggested by Heckman, especially if there is possible collinearity in the model (Puhan, 1997). The last two columns of Table 13 present the results from the IV estimation of log annual earnings for the current sample. The endogenous schooling variable was instrumented using the household head's level of education. The equation was exactly identified since one instrument was used for one endogenous variable.[29]

In order to test the strength and validity of the instrument (in the first stage) the Kleibergen-Paap LM statistic was used. This is an appropriate test for weighted survey data and tests the null hypothesis that the matrix of reduced form coefficients is underidentified (or has rank=K1-1). The chi-square value for this test was 1058.23 and the null hypothesis of underidentification was rejected. As discussed in Section 5.2, this test is robust to heteroskedasticity, autocorrelation, and clustering.

In the second stage, in case of a just-identified model, instrument exogeneity cannot be statistically tested. The choice of household head's education as an instrument

[28] See Table B.1 in Appendix B for results of the first stage equation.
[29] Table B.2 in Appendix B provides the results for the first stage regression predicting completed years of education.

is consistent with previous research on the returns to schooling using parents' education as an instrument (Card, 2001).

The IV results showed that each additional year of schooling is associated with a six percent increase in earnings. The difference in the OLS and IV estimates of the coefficient on schooling suggests that OLS significantly underestimates the returns to schooling in this sample. This observed downward bias using OLS is in keeping with Card's (2001) findings that attribute the downward bias to endogeneity of the schooling variable.

Table 13

Marginal effects of schooling on log wages using OLS, Heckman and Instrumental Variables methods, and controlling for other variables

	OLS Estimates		Heckman Estimates		IV Estimates	
	Coef.	SE	Coef.	SE	Coef.	SE
Years of schooling	0.04***	0.002	0.03***	0.002	0.06***	0.005
Age-Between 15-21 years	-0.44***	0.034	-0.45***	0.034	-0.47***	0.035
Age-Between 22-28 years	-0.12***	0.021	-0.13***	0.021	-0.15***	0.023
Age-Between 29-39 years	0.05**	0.018	0.05***	0.018	0.03	0.019
Female	-0.54***	0.042	-0.65***	0.039	-0.48***	0.043
Age Group 1*Female	-0.10	0.055	-0.08	0.055	-0.14*	0.056
Age Group 2*Female	-0.02	0.031	-0.02	0.031	-0.03	0.031
Age Group 3*Female	-0.02	0.026	-0.01	0.026	-0.02	0.027
Urban	0.93***	0.025	0.90***	0.025	0.91***	0.026
Marital Status	0.29***	0.026	0.45***	0.028	0.30***	0.027
Marital Status*Female	-0.30***	0.041	-0.41***	0.041	-0.30***	0.041
Social Group - OBC	-0.25***	0.029	-0.23***	0.029	-0.23***	0.029
Social Group - Dalit	-0.18***	0.030	-0.15***	0.030	-0.15***	0.031
Social Group - Adivasi	-0.32***	0.051	-0.28***	0.051	-0.28***	0.053
Religious Group - Muslim	-0.12**	0.039	-0.10**	0.039	-0.08*	0.040
Ability - Grade X Performance	0.51***	0.037	0.49***	0.037	0.39***	0.043
Ability - Grade X Performance	0.16***	0.027	0.15***	0.027	0.05***	0.032
English Fluency	0.19***	0.030	0.19***	0.029	0.10**	0.034
Intercept	8.94***	0.042	8.75***	0.043	8.82***	0.045

IV (Household head's schooling)			*0.43****	*0.005*
Selection Index	*0.30****	*0.016*		
N				*78,737*

Source: Indian Human Development Survey, 2004-05. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

5.3.2 Returns to TVET

To estimate log annual earnings (Equation 2 in Section 5.2) the sample was restricted to all working age adults with 10 or more years of education. Table 14 presents these results using OLS and Heckman's selection correction method.[30] Results using OLS estimation are presented for four models; with each model adding more controls. Cluster robust standard errors (at the PSU-level) are reported for the OLS models. Results from Heckman's selection correction method are presented for the full model with all controls.

The first OLS model in Table 14 models log annual wages as a function of human capital variables; namely, education and training acquired by the individual. As described before, three dummy variables representing completion of a Bachelor's, Master's and Professional degree, respectively, control for educational attainment. Results show that TVET participants earn 38 percent more in annual wages than those who do not participate in TVET.[31] As discussed in Section 5.3.1, this is likely an underestimation of the returns to TVET in the current sample.

Controlling for the demographic characteristics of the individual (age, gender, urbanicity, marital status, and social class) results in a sizeable reduction of the estimated

[30] See *Section 5.4* for a discussion on why IV estimation was not used to estimate returns to TVET.
[31] In Indian currency units, annual wages for TVET participants are Rupees 19,894 more.

coefficient on TVET; the coefficient estimate goes down by more than 15 percentage points.

When measures of ability are added to the model, the returns to TVET further decrease with TVET participants seeing an earnings advantage of 14 percent over those not in TVET. This relationship is significant at the 0.1 level.

In the final stage of estimation, controls for English language fluency are added to the OLS model. The results show that controlling for English fluency has a modest reduction on the coefficient on TVET (0.114 log points); but this relationship is not statistically significant.

The results of the OLS regressions are based on a non-random sample of the population – those individuals for whom earnings are observed. In order to address the bias due to sample selection Heckman's method is used and reported in the last two columns of Table 14. Section 5.2.1 describes the method in detail. Results of the first step (or the labor force participation equation) are presented in Table B.1 in Appendix B.

The results in Table B.1 show that there is indeed selection in the model (as evidenced by the Wald Chi-Square test of independent equations). The estimates reported in Table 14 indicate that, controlling for ability (and English fluency), educational attainment, and demographic characteristics, TVET participants earn 18.6 percent more in annual wages than those not participating in TVET. This estimate is larger than that estimated by OLS and is statistically significant. In keeping with the results from *Section 5.3.1* and previous research on the direction of the bias in estimating earnings using OLS

(Card, 2001), the estimates produced by the Heckman method likely underestimate the returns to TVET in the current sample.

In order to test if the marginal effects of TVET participation on earnings are different from the marginal effects of having a BA, MA or professional degree, *F*-statistics and Chi-square statistics were computed for the OLS and Heckman models, respectively. These are reported in Table 14. Results for the OLS and Heckman models showed slight differences. While the OLS results indicate that the marginal effects of TVET on earnings are equal to the marginal effects of a BA degree and a MA degree, the Heckman results indicate TVET effects are different from those of a BA degree. The marginal effects of a professional degree are not equal to the other marginal effects in either the OLS or Heckman models.

Table 14

Marginal effects of TVET participation on log wages among 15-65 year olds with 10 or more years of education

	OLS (With PSU-level Cluster Robust SEs)								Heckman	
	Human Capital		Demog. Controls		Ability		English Fluency		Full Model	
Predictors of Log Annual Earnings:										
BA Degree (Ref: Other)	0.577***	0.057	0.220***	0.043	0.121**	0.044	0.051	0.045	0.062	0.045
MA Degree (Ref: Other)	0.851***	0.075	0.400***	0.060	0.232***	0.060	0.151*	0.061	0.188**	0.061
Professional Degree (Ref: Other)	1.247***	0.094	0.676***	0.069	0.468***	0.071	0.394***	0.073	0.432***	0.074
TVET (Ref: Other)	0.384***	0.107	0.223**	0.084	0.142	0.084	0.114	0.084	0.186**	0.066
Age-Between 15-21 years			-1.068***	0.109	-1.093***	0.111	-1.083***	0.107	-1.372***	0.115
Age-Between 22-28 years			-0.609***	0.063	-0.630***	0.064	-0.620***	0.063	-0.654***	0.064
Age-Between 29-39 years			-0.256***	0.043	-0.266***	0.042	-0.266***	0.041	-0.233***	0.042
Female (Ref: Male)			0.122	0.127	0.122	0.126	0.112	0.125	-0.014	0.129
Age Group 1*Female			-0.434*	0.187	-0.487**	0.188	-0.487**	0.189	-0.393*	0.190
Age Group 2*Female			-0.289*	0.119	-0.333**	0.117	-0.340**	0.116	-0.416***	0.119
Age Group 3*Female			-0.181	0.098	-0.209*	0.096	-0.192*	0.095	-0.239*	0.096
Urban (Ref: Rural)			0.994***	0.046	0.915***	0.045	0.902***	0.045	0.841***	0.045
Marital Status (Ref: Unmarried)			0.257***	0.067	0.256***	0.068	0.257***	0.066	0.381***	0.066
Marital Status*Female			-0.412***	0.117	-0.404***	0.115	-0.394***	0.115	-0.574***	0.117
Social Group - OBC (Ref: Non-OBC)			-0.292***	0.046	-0.279***	0.045	-0.274***	0.045	-0.275***	0.045
Social Group - Dalit (Ref: Non-Dalit)			-0.269***	0.070	-0.241***	0.069	-0.231***	0.068	-0.198**	0.064

Social Group - Adivasi (Ref: Non-Adivasi)			-0.257**	0.100	-0.177	0.101	-0.159	0.098	-0.133	0.095
Religious Group – Muslim (Ref: Non-Muslim)			-0.094	0.080	-0.041	0.080	-0.045	0.079	-0.053	0.078
Ability (>60% in grade 10)					0.573***	0.058	0.537***	0.058	0.510***	0.055
Ability (<60% in grade 10)					0.164**	0.054	0.144**	0.053	0.127*	0.051
English Fluency							0.277***	0.048	0.263***	0.047
Intercept	9.898***	0.049	9.958***	0.092	9.832***	0.101	9.680***	0.103	9.508***	0.102
TVET = BA							*0.07* [a]	*(0.79)*	*4.24* [b]	*(0.04)*
TVET = MA							*1.06* [a]	*(0.30)*	*0.00* [b]	*(0.96)*
TVET = Professional Degree							*12.9* [a]	*(0.00)*	*9.24* [b]	*(0.00)*
N		*7915*		*7915*		*7915*		*7877*		*7877*

Source: Indian Human Development Survey, 2004-05. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

[a] *F*-statistic for test of equality of coefficients (Prob > *F* in parentheses).

[b] Chi-2 value for test of difference between coefficients (Prob > Chi-2 in parentheses)

The coefficient estimates on TVET and the education level variables were used to compute the returns associated with each additional year at a particular level of education. This is done by dividing the difference in two estimates by the difference in the years of education required to attain the higher level of education. So, for example, completion of a Bachelor's degree in India implies a total of 15 years of education and completion of a Master's degree implies 17 years of education. In order to compute the returns to each additional year of education in a Master's degree (over a Bachelor's degree), the difference between the coefficient estimates for a Master's degree and a Bachelor's degree is divided by two (the difference in the number of years of education required for completion of a Bachelor's and Master's). These results are presented in Table 15 and Figure 25.[32]

The table and figure below show that the returns to an additional year in a Bachelor's degree are significantly lower than postsecondary TVET. Although secondary TVET also shows higher returns than a Bachelor's degree, this difference is very small. The returns associated with an additional year in a Master's degree program although positive when compared to postsecondary TVET, are negligible. The biggest returns are observed in the case of Professional degrees.

[32] The time for completion of TVET programs varies by type of program and the previous level of education completed. Programs can require a total of 12 to 14 years of education/training.

Table 15

Returns to each additional year of education

Education level	Difference in years of education	Returns
BA over TVET*	3 years	-0.002
BA over Postsecondary TVET	1 year	-0.126
MA over Postsecondary TVET	3 years	0.001
Professional over Postsecondary TVET	3 years	0.082

Note. * Includes TVET programs that can be accessed after 10th grade and result in a Certificate.
** Includes TVET programs that can be accessed after12th grade and result in a Diploma.

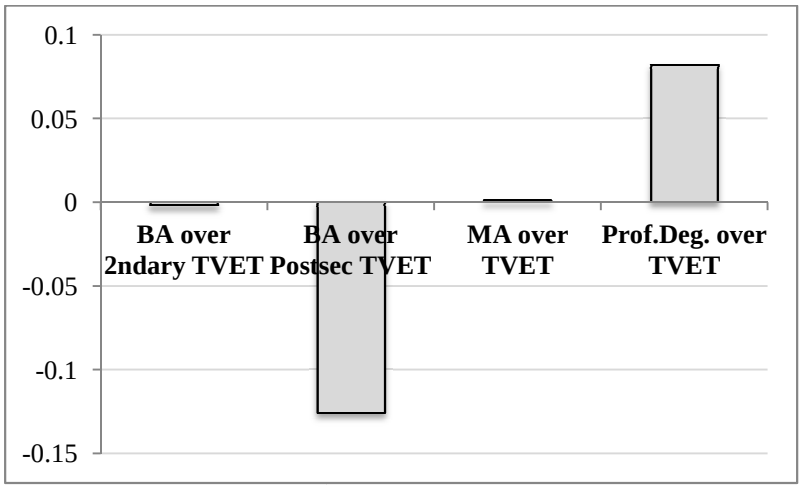

Figure 25. Returns to each additional year of education
Source: Indian Human Development Survey, 2004-05.

5.4 Limitations

This section discusses some of the limitations of the analysis and results presented in this chapter. These limitations are related to data constraints and generalizability of the findings, and to the estimation methods used.

First, TVET participants in this analytic sample represent individuals with at least 10 completed grades of schooling. Formal TVET institutions in India offer programs for individuals with a minimum of eight completed grades of schooling. The TVET participants in this sample are therefore not representative of all formal TVET participants in India. Further, the TVET sample in these data constitutes a very small proportion of the sample and represents about half the districts in the country. The estimates reported here therefore cannot be generalized to all TVET programs, or to the country as a whole.

Second, previous research on wages in the Indian context has noted the heterogeneity in earnings along gender, location and social class dimensions. Descriptive graphs of the average wages of individuals in this analytic sample also show significant differences by gender and urban-rural status. The heterogeneity in the returns to TVET by gender and urban-rural status has not been examined owing to the small proportion of TVET participants in these data.

Next, the IV approach has not been used to estimate returns to TVET. In Equation 2 in Section 5.2 schooling and TVET are both endogenous. Therefore, at least two instruments (one for each endogenous regressor) are required. Various potential instruments were tested. These included instruments representing household

characteristics (head of the household's education, head of the household's participation in TVET programs), supply-side characteristics (distance to secondary and postsecondary institutions, supply of postsecondary institutions) and infrastructural indicators (availability of roads, availability of transportation). None of these instruments passed the statistical tests for strength, validity, and overidentifying restrictions as per the Kleibergen-Paap Wald tests and the Hansen J statistic.[33] Murray (2006) notes that even in case of slight violations of IV assumptions the IV estimator tends to be imprecise, especially in large samples. Due to the lack of strong instruments for estimating returns to TVET the analysis has been limited to using OLS and Heckman methods.

A final limitation of the estimation models is the lack of context-level controls. In order to address this limitation the OLS models for returns to TVET were estimated using PSU-level fixed effects. The sampling weights could not be used in the estimation of the fixed effects models; the results are presented in Table B.6 in Appendix B. The fixed effects estimates show small differences compared to the results discussed in the previous section. With regard to the returns to TVET, the OLS model without fixed effects estimated an 11% increase in earnings while the fixed effects estimate indicates a 23% increase in earnings for TVET participants. While both estimates likely underestimate the returns to TVET in the Indian case, the weighted OLS estimates with standard errors adjusted for heteroskedasticity and clustering at the PSU-level are preferred.

[33] Tables B.4 and B.5 in Appendix B present IV results for the first stage regressions (predicting completed years of schooling and TVET participation) using the household head's schooling and TVET participation as instruments.

Chapter 6: Returns to a Secondary School TVET Program - Impact Estimates Using Propensity Score Matching

About 45 percent of the world's youth (about 700 million young people) are in the Asia and Pacific region (United Nations Economic and Social Council for Asia and the Pacific [UN ESCAP], 2013) and according to the 2013 World Development Report on "*Jobs*", globally, more than 621 million youth were neither working nor studying in 2013 (World Bank, 2013). Reports indicate that transitions from lower to upper secondary schooling, and transitions from school to work are the main obstacles facing youth globally and in developing countries in particular.

These youth challenges are somewhat exacerbated in the Indian context. About a third of India's population comprises those between 10-24 years of age. Amongst all 15-24 year olds, between 60-65 percent were enrolled in secondary school, 15 to 20 percent were enrolled at the tertiary level and about 10 percent were unemployed (Population Reference Bureau, 2013).

Policy discussions on how best to address these challenges often focus on developing links between education and careers through various types of TVET programs. While the majority of TVET programs for youth focus on school-to-work transitions (Arum & Shavit, 1995; Castellano et al., 2011; Hawley, 2008), fewer have tried to address the high dropout rates between the lower and higher secondary stages of schooling (Agodini & Deke, 2004; Kemple et al., 2000; 2004; 2008). Further, of the school-based TVET programs that do exist, few have been evaluated. To date, only one long-term evaluation of secondary level TVET has been conducted that studies the

impact on earnings over the lifecycle (Hanushek et al., 2011). Findings from the evaluation literature suggest that the effects of secondary-level TVET are ambiguous; the impacts are largely context dependent and vary based on schooling structure and type of training.

Experimental evaluations of Career Academies in the United States (Kemple et al., 2008) show no impact on high school outcomes but significant positive impacts on labor market outcomes for young men in the sample. A study of high school TVET programs in US schools found similar results with regard to preventing dropouts among high school students (Agodini & Deke, 2004). Hanushek's (2011) multi-country study also finds positive labor market outcomes for TVET participants in the short term but diminishing returns to education (as compared to those with general academic training) over the life cycle.

There are several methodological concerns when evaluating TVET programs. As Ryan (2001) notes, selection on unobservables, lack of prior labor market experience, and difficulty in conducting experimental evaluations make statistical evaluations in this area difficult.

In spite of limited and heterogeneous evidence on the effectiveness of TVET programs for youth, governments, civil society organizations and multilateral agencies are establishing youth policies focusing on training and employment through an expansion of TVET programs at the secondary and postsecondary levels (UN ESCAP, 2013). In India, "vocationalisation" at the secondary level has gained traction over the past few years and TVET programs for secondary school students are being implemented in several states.

The lack of reliable evidence on TVET programs in general, and those for young people in particular, poses a major constraint in developing effective policies that can best serve the needs of youth. This study begins to address the need for reliable evidence and examines the effectiveness of a secondary school-based TVET program in one state in rural India. Training students in foundational skills around four broad topics the program aims to improve educational and labor market outcomes of rural Indian youth. While several other innovative secondary school TVET programs are being fielded in different parts of the country, this study attempts to address the gaps in what is known about the impacts of these programs. In so doing, it focuses on one program and uses rigorous quasi-experimental methods to ask the following questions:

1. Does participation in secondary school-based TVET result in higher rates of school completion?
2. Does participation in secondary school-based TVET result in higher rates of enrollment beyond grade 10?
3. Do the effects of TVET participation on school completion and further enrollment vary by gender?

6.1 The Program - Introduction to Basic Technology (IBT)

The TVET program under study is developed for rural secondary schools and targeted at students in grades eight through ten. Designed as a supplemental whole-school program, it was first introduced in 2006 in five schools across three districts in one state of India. In the second year, the program expanded to six additional schools in two more districts in the state. Currently, the Introduction to Basic Technology (IBT) is offered in

over 60 rural schools reaching 8,000 students. Lend-A-Hand-India (LAHI), a Maharashtra state-based non-governmental organization (NGO) is responsible for the development, implementation and monitoring of the IBT program.

The objective of the IBT program is to improve participation in secondary schools in the short term and address rural unemployment in the long term. The program envisions achieving this objective through providing skills training to secondary school students with support from the school and community. As shown in Figure 26, the focus on secondary school students is a response to the high rates of dropout at the secondary stage in India.[34] "IBT schools" or schools where the IBT program is offered, provide students in the relevant grades foundational skills in four subject areas – Agriculture, Engineering, Energy & Environmental Science, and Home and Health Sciences – that are considered relevant to further academic education, TVET, and/or employment and entrepreneurial opportunities.

The IBT program is part of the State's secondary school curriculum and can be taken as an optional subject in grades nine and ten.[35] Students who complete the two-year program and successfully complete the terminal statewide examination in the required subjects, including IBT, are awarded a certificate of completion. Successful students have the option of enrolling in any postsecondary vocational program offered at public institutions run by the state. LAHI's partnership with the state government ensures that

[34] According to NSSO data, the gross enrollment ratio at the lower secondary stage in India was about 51 percent in 2007-08 while the gross attendance ratio for the same stage in the same year was about 70 percent (Biswal, 2011).
[35] Lend-a-Hand-India identifies IBT as a three-year program beginning in the eighth grade. Schools offering IBT with LAHI's support thus enroll students in IBT beginning in the eighth grade.

interested students who have successfully completed the IBT program are granted one of the reserved seats at public TVET institutions.

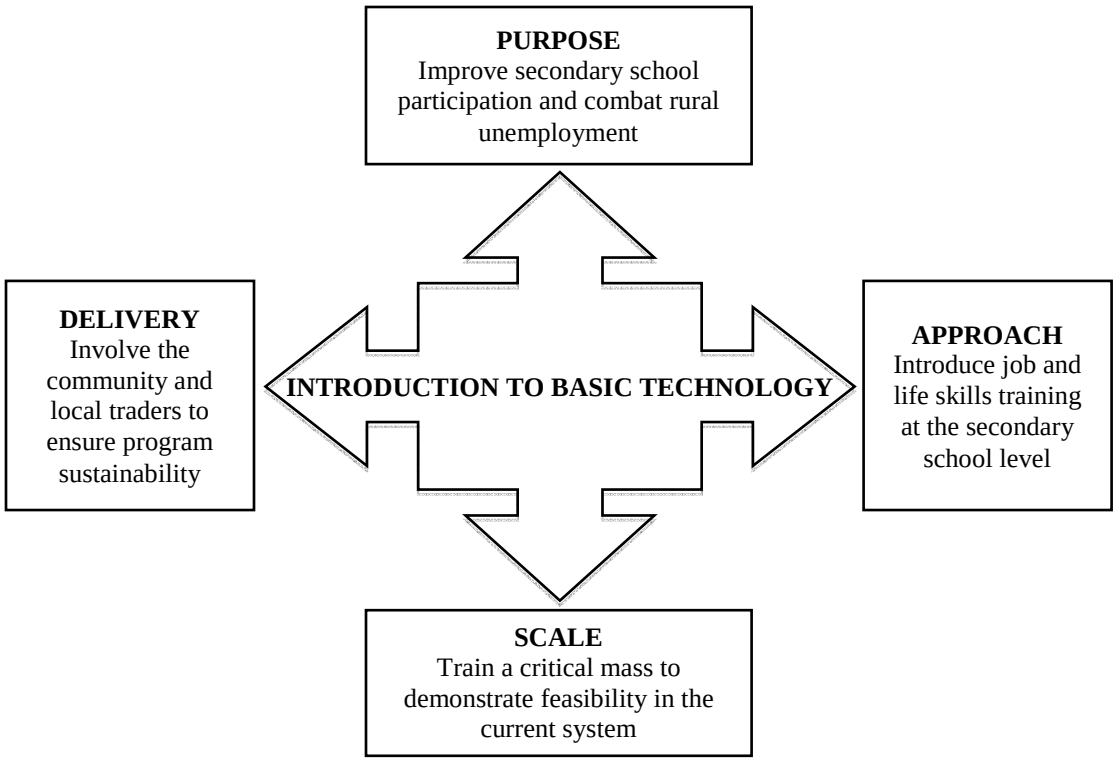

Figure 26. Graphical representation of the approach to the IBT program (adapted from Lend-A-Hand-India)

6.1.1 Implementing IBT

The IBT syllabus includes theoretical and practical modules that are integrated into the school's timetable. The math or science teacher at the school is responsible for transacting theoretical components of the IBT syllabus while locally recruited professionals with demonstrated experience in one or more of the IBT subjects lead the practical modules. School principals are responsible for recruiting local professionals –

preferably from within the community that is served by the school – who act as volunteer teachers. LAHI provides the training necessary for newly recruited volunteer teachers and others associated with the IBT program to implement the program in the school. These training programs are conducted twice a year – before the academic year begins and midway through the academic year. In addition to providing content-specific instruction, the training program focuses on providing newly recruited instructors the pedagogical skills required to teach students practical skills. Principals commit to supporting and monitoring the IBT program, modifying the school's timetable to include IBT classes during the school day,[36] and managing program funds.[37] LAHI representatives visit IBT schools three to four times during the school year to monitor progress and address implementation or content-related concerns.

The cost of implementing the IBT program in a secondary school with a cohort-size of 50 students works out to $12,000 over a period of three years. The investment in the first year includes capital improvements; namely, establishing a workshop and equipping it with the requisite tools and machinery to conduct practical modules that are part of the IBT syllabus. Schools fulfill 20 percent of this one-time cost with LAHI fulfilling the remainder. After the first three years, schools commit to taking on the financial responsibility of the program with minimal financial assistance from LAHI. Since the IBT program does not receive any public funding, schools that cannot raise additional funds from the community or private sources are not eligible to offer the

[36] LAHI offers schools a suggested timetable that can be adapted to meet their specific needs. See Table C.1 in Appendix C.
[37] Principals are required to set up a separate account for all IBT-related expenses that is used for program costs including instructor salaries.

program to students. Government-aided schools[38] instead, funded through a combination of public and private grants are the only type of schools eligible to offer IBT. Unrecognized and unaided or private schools interested in implementing IBT require necessary approvals from the state government to offer the IBT syllabus.

LAHI organizes a training program for all teachers and principals before the beginning of the academic year.

6.1.1.a School and student participation in IBT

As mentioned earlier, IBT is a whole-school program. All students in the relevant grades in participating schools receive IBT training. Schools interested in becoming "IBT schools" go through a multi-stage process of selection drawn up by LAHI in partnership with the state's department of education. Selection is based on fulfilling several eligibility criteria laid down by LAHI and the state's department of education. The process of being selected as an IBT school is illustrated in Figure 27.

First, schools receive information about the option of offering IBT training to students in their schools in one of three ways - 1) The state's Technical Board[39] publishes an advertisement in local newspapers inviting schools to offer one of three vocational programs (the IBT syllabus is classified as 'Vocational 1 (V1)'[40]) at the secondary level,

[38] Government-aided schools are very similar to government (i.e. public) schools. They are privately managed but receive public funds. The teachers at government-aided schools are public employees and paid directly by the state government at the same teacher salary rate. These schools also charge the same tuition (now mandated to be nil) as government schools (Kingdon, 2007).
[39] The Technical Board is a division of the state government responsible for vocational and technical education in the state.
[40] The state's secondary curriculum includes three optional vocational subjects – V1 (Introduction to Basic Technology), V2 (Elements of Mechanical Engineering), V3 (Elements of Electronics Engineering). The

2) LAHI conducts awareness drives and holds informational sessions from time to time, and 3) schools hear about IBT through word of mouth. Interested schools reach out to the Technical Board or LAHI and in turn receive a list of eligibility criteria that must be fulfilled in order to implement the IBT program. The eligibility criteria are as follows –

1. Availability of land for agriculture (either owned by the school or donated by the community)
2. Availability of electricity to run practical sessions of the IBT syllabus
3. Availability of two spare rooms to serve as workshops for practical modules
4. Availability of a weekly market (*bazaar*) in the village
5. Minimum cohort size of 40 students

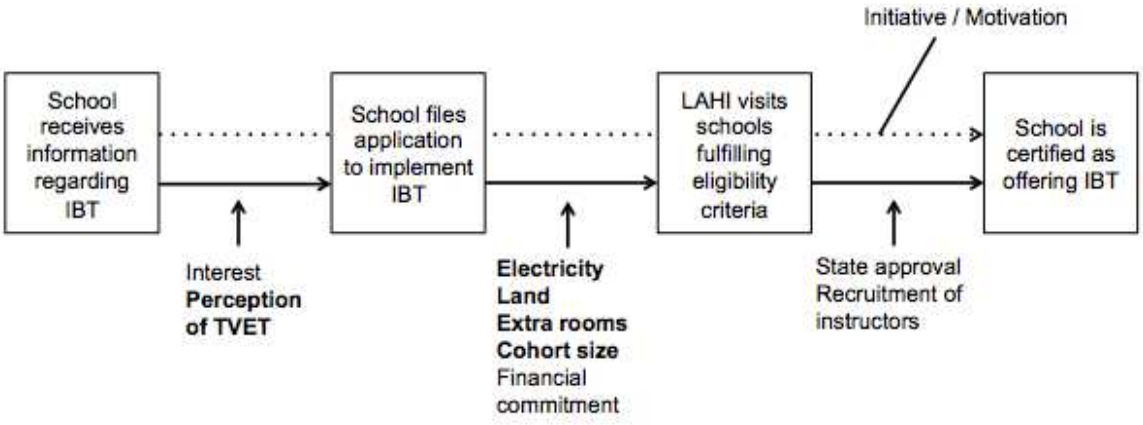

Figure 27. Graphical representation of IBT selection process

Schools meeting these requirements submit a formal application to LAHI along with an official declaration from the school's management committing to support the

courses graduate from foundational to advanced in the level of skill training offered, with V1 being foundational and V3 being advanced.

program and ensure that it is implemented with fidelity.[41] This commitment includes a promise towards making the IBT program sustainable after the first three years of the program. LAHI follows up with a visit to the school to confirm that all requirements have been met. An additional important component of LAHI's visit is orienting the school and the community to the IBT program. To this end, school and community members are invited to watch a short film on the objectives of the IBT program; it's components, advantages, and the role of the school and the community in sustaining the program.[42]

Finally, schools that meet all eligibility requirements commit to a one-time investment of $2000 towards setting up the workshop space with tools and other materials. Officials from the department of education visit selected schools (those completing all requirements including the initial investment) to gauge their level of preparedness to implement IBT. Schools receiving final approval from the state are deemed "IBT schools" and begin to receive funds from LAHI for program implementation.

Students interested in participating in IBT must be enrolled at an IBT school at the lower secondary stage (grades eight through ten). In IBT schools with one division per grade, all enrolled students participate in IBT. In larger schools (with two or more grades per division), only a subset of students can participate in the IBT program. Student selection in to IBT at these schools is based on student interest and one of two criteria – ability to pay additional tuition or meeting some predetermined academic performance standard. Interested students meeting school-specific eligibility criteria participate in IBT.

[41] See template of 'Management Resolution and Commitment from Partners' in Appendix C
[42] www.lend-a-hand-india.org

6.2 Research Design

Given the post-hoc nature of the evaluation this study adopts a non-experimental approach - propensity score matching (Rosenbaum & Rubin, 1990; Rosenbaum, 2002; Stuart, 2010). Since students are not randomly assigned to schools offering IBT, a post-hoc control group was created with a comparison group of students from comparison schools. The comparison group includes students from the same cohorts as those in the treatment group. Pre-intervention data on treated and non-treated students was used to create a matched sample of treatment and comparison students sharing similar characteristics. The "best matches" of treatment and comparison students were used to estimate impacts controlling for demographic, background, and pre-intervention factors.

6.2.1 Data

The data used in the evaluation of the secondary school-based vocational education program, IBT, was collected in 2013, over a period of ten months. The data collection targeted three types of information from three different sources – administrative records from schools, school characteristics from interviews with principals, and outcomes information from students. The same school and student survey instruments (see Sections 6.2.1.b and 6.2.1.c) were used to collect data from treatment and comparison samples. Heckman et al. (1998) note that this can significantly reduce the bias in propensity score matched program estimates since observed characteristics are measured in the same way and thus represent the same concepts.

A local data collection agency was recruited to carry out the data collection. The team included four lead persons (all of whom were certified in Human Subjects Research

as per University of Pennsylvania's Institutional Review Board requirements) and several assistant field workers not directly involved in administration of surveys or collection of records. All survey instruments were translated in the local language (Marathi) and administered in person (see Appendix C for school and student survey protocols).

6.2.1.a Administrative Records

Administrative data on students was collected from school rosters. These data included attendance and performance records (achievement scores on school tests) for students in the treatment and comparison samples for a period of five years, beginning in the fifth grade through when they left school or graduated at the end of tenth grade. The attendance data included information on total number of days attended and total number of working days for each student in each academic year. The performance data included scores on four subject tests at each grade level – Hindi, English, Mathematics and Science.

As mentioned above, these data were gathered from rosters available at each school. Rosters were scanned using portable scanners to create electronic documents containing the relevant data. The data from these electronic documents was then entered electronically in Microsoft Excel.

The administrative data thus collected were used for two purposes. First, the attendance and performance data for grades five to seven comprised pre-intervention data and were used in the estimation of propensity scores. Second, the attendance data were used to generate a list of all eighth grade students at each sampled school (comparison and treatment). Eighty percent of students in the comparison schools were randomly

selected to be part of the comparison student sample and shortlisted for administration of student surveys (a detailed description of sample selection methods is provided in the next section). All eighth graders in treatment schools were included in the treatment sample and shortlisted for survey administration.

6.2.1.b Principal Survey

Principal or school surveys (See Appendix C) were designed to collect contextual information on all treatment and comparison schools. The surveys were administered in the local language (Marathi) to principals or vice principals at sampled schools and focused on collecting information on school background characteristics (enrollment, and infrastructure), school staff (number of teachers, teacher qualifications and teacher experience), average student performance (completion rates, and average scores on tenth grade state tests), and dropout rates at grades eight, nine and ten. These data were used to compare school characteristics across all schools in the treatment and comparison groups.

6.2.1.c Student Survey

Students sampled from treatment and comparison schools (eighth graders in 2006 and 2007) were targeted for student survey data collection in the villages/towns where students reside. Apart from collecting information central to the focus of this study (educational and employment outcomes), the surveys (See Appendix C) collected information on students' background information (gender, and ethnicity), and their household characteristics (type of dwelling, household assets, education and employment

details of household members). Survey questions focusing on outcomes covered educational outcomes (10^{th} grade completion, 12^{th} grade completion, and postsecondary enrollment), and employment outcomes (employment status, type of employment, and wages/earnings). Demographic and historical information on prior achievement was also collected as part of these surveys. This information was used to estimate the propensity scores as well as control variables in the analysis of outcomes.

6.2.2 Treatment Sample

To estimate impacts of the IBT program on educational and employment outcomes the treatment group was selected such that sufficient time had lapsed for relevant outcomes to become available. The first two cohorts of IBT schools were selected for inclusion in the treatment sample. This included five schools that began the IBT program in 2006 and six schools that began the program in 2007. Thus, at the time of data collection in 2012, it had been six to seven years since the first cohort of students entered the program and three to four years since the first cohort graduated from the IBT program.

IBT is a school-based program and open to all students entering the eighth grade in IBT schools. In smaller schools that have one division per grade, all students participate in the program. In schools with more than one division per grade, students are assigned to IBT based on their interest and/or ability to pay the tuition associated with the program. Students who entered the eighth grade in 2006 (in the five treatment schools that introduced the IBT program in 2006) and those who entered the eighth grade in 2007 (in the six treatment schools that started the program in 2007) were included in the

treatment sample. In three schools with more than one division in the eighth grade, only those students in the IBT division were included in the treatment sample. The treatment sample therefore included 11 IBT schools and 555 IBT students.

Data from the student surveys revealed that of the students in the treatment group (identified as such by virtue of attending an IBT school or being in the IBT division of an IBT school), about half responded to participating in the IBT. This discrepancy between those identified as treatment group students by program definition and those self-identifying as treatment group students could be explained in a few possible ways. Because sufficient time had lapsed between program completion and survey administration it is possible that many students did not remember participating in IBT. The student survey specifically asked if the student had participated in any skill training (or other type of) program during grades 8 to 10. Since IBT is part of the school curriculum at IBT schools (or IBT divisions), it is possible that IBT students do not perceive it as a "supplementary" skill-training program. Other reasons for discrepancy in responses could have to do errors in collecting data – miscommunication between the field investigator and student or errors during data entry.

6.2.3 Comparison Sample

The selection of comparison schools and students was motivated by two main concerns – identifying schools and students that were as similar as possible to those in the treatment group, and selecting a large enough pool of comparison group members to find suitable matches using propensity scores (Heckman et al., 1998). Thus, for every treatment school, a minimum of three comparison schools were identified that were

located in the same *tehsil* (sub-district) as the treatment school[43] and that shared characteristics similar to those of the counterpart treatment school. Apart from geographical proximity to treatment schools, the following school characteristics were compared and formed the basis for selecting comparison schools for data collection –

1. School size
2. Number of grades
3. Proportion of minority students
4. Number of classrooms
5. Availability of electricity
6. Availability of land

The Secondary Education Management Information System (SEMIS)[44] is an online portal maintained by the education department of the country that provides basic information from all recognized secondary and higher secondary schools of the country. This online tool was used to access "report cards" for all treatment schools, and for up to ten additional schools belonging to the same *tehsil*. The "report cards" included all the information outlined above and were used to compare treatment and potential comparison schools on the above-outlined indicators. Schools meeting all or most of the identified criteria were selected as part of the comparison sample.[45] Since, in several cases, potential comparison secondary schools did not match their treatment counterparts on all

[43] Heckman et al. (1998) recommend that selecting treatment and control group members that face the same economic incentives (for example, belonging to same geographic area) reduces bias in PSM estimates.
[44] http://14.139.60.147:8051/Default.aspx
[45] See Table C.10 in Appendix C for a comparison of treatment and comparison schools on select characteristics.

identified indicators, a larger number of comparison schools were shortlisted for data collection. The comparison school sample thus included 39 schools.

Selection of comparison students from within comparison followed a slightly different approach than that adopted for the treatment sample. To balance time and cost constraints 80 percent of students who entered the eighth grade in the relevant years (2006, if the school was serving as a comparison to a treatment school that began the IBT program in 2006, and 2007 otherwise) were randomly assigned to be part of the comparison student sample. Further, based on previous research findings indicating lower rates of female participation at the secondary school stage, and early age of marriage among females in rural areas, female students were oversampled in the ratio 1.5 to 1. Thus, the comparison student sample included 2,654 students of which 60 percent were female.

6.2.4 Analytic Sample

This section describes the results of the data collection effort in terms of the number of respondents located and surveyed, and the size of the analytic sample i.e. those for whom complete data are available.

6.2.4.a. Surveyed sample

The data collection for this study targeted 3,209 students across 53 schools and five districts of Maharashtra state. Since students were being tracked three to four years after their expected graduation year some degree of attrition was expected. Previous data

collection efforts based on similar procedures had achieved a response rate of about 60 percent. A slightly higher response rate (70 percent) was expected for the treatment sample because of prior information on schools and students in the treatment group.

The data collection effort resulted in complete school data from all 11 treatment schools. Of the 42 comparison schools four did not provide consent to participate in the study; and school data was collected from 38 comparison schools. Table 16 shows the effective response rates at the school-level for both groups. The difference in response rates between treatment and comparison schools was 11.63.

Table 16
Targeted and surveyed sample sizes for treatment and comparison schools

	Targeted	Surveyed	Effective Response Rate	Difference
Treatment	11	11	100	11.63
Comparison	42	38	88.37	

Table 17
Targeted and surveyed sample sizes for treatment and comparison students

	Targeted	Located	Located with complete data*	Effective response rate**	Difference
Treatment	555	305	160	52.29	14.91
Comparison	2654	1895	1161	61.27	
N (Treatment schools)		*11*	*11*		
N (Comparison schools)		*35*	*33*		

Note. * Complete data includes information on educational outcomes (school completion and postsecondary enrollment, in this case) and all the control variables used in the estimation of propensity scores and effects.
** The effective response rate is calculated based on the number of students located. The response rate would be significantly lower (28.82 and 43.74 percent) if it was based on the number targeted in each group.

For the student-level data collection 555 treatment students and 2,654 comparison students were targeted. These students belonged to 11 treatment and 35 comparison schools, respectively. Of the 38 comparison schools that participated in the school-level data collection three could not provide student administrative records. So it was not possible to select comparison students from these three schools. A response rate of 70 percent and 60 percent was expected for treatment and comparison students respectively. Table 17 shows the number of students located and the number for whom complete data were available at the end of data collection.

Field investigators were able to locate 55 percent of treatment group students and 71 percent of comparison group students in their villages/communities. There was significant variation in the proportion of treatment and comparison group students located by district. One district (Thane district) in particular showed very low response rates among treatment students. The treatment schools in this district were close to the state border and discussions with school principals revealed that the school is attended by a significant proportion of out-of-state students (from Gujarat state) whose parents temporarily migrate to the district for employment. District-specific numbers are presented in Table C.2 in Appendix C.

The gender breakdown amongst those located and not located for survey administration is presented in Figure 28. The proportion of females in the comparison group was significantly higher across surveyed and non-surveyed groups. While comparison females made up about 58 percent of the sample, treatment females

constituted about 43 percent of the sample in both groups. This difference reflects the oversampling of females in the comparison group discussed in Section 6.2.3.

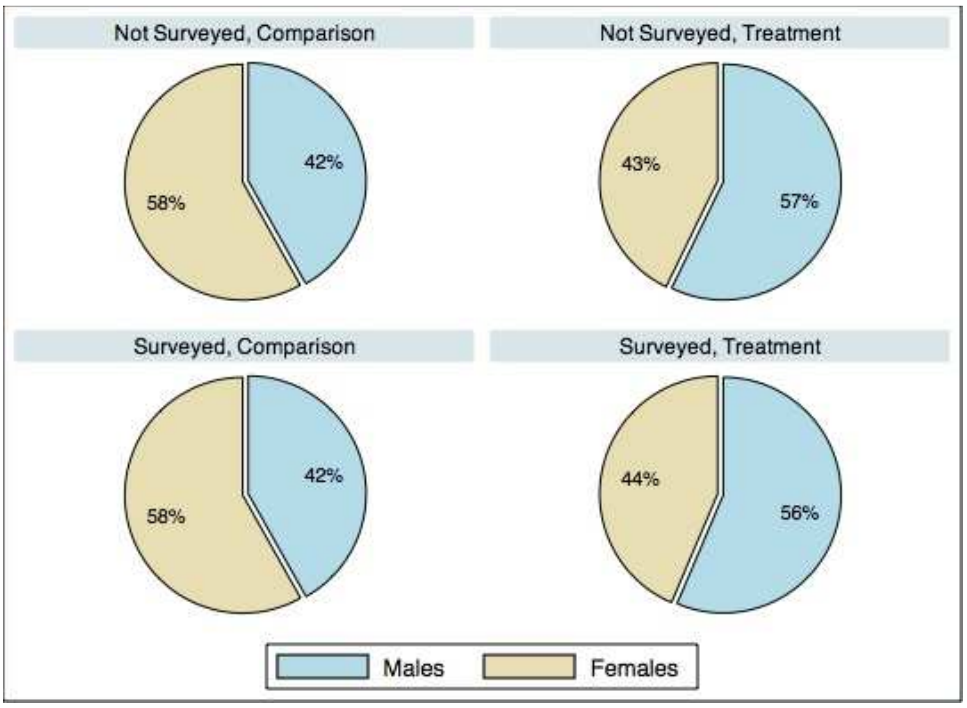

Figure 28. Gender breakdown of surveyed and non-surveyed groups by treatment status

Those located and not located for survey were compared on key indicators to check for any systematic differences between the two groups. Data from student administrative records included performance information in grades five to ten. Although these data were not available for all targeted students[46], comparisons were made on the sample for which this information was available. Table 18 presents the results from *t*-tests comparing test score means by treatment status for those located and not located. Raw test scores were converted to Z-scores within each school using the school mean and

[46] There was significant variation between schools in the availability and quality of administrative information.

standard deviation. Significant differences were observed for gender and later academic achievement. The gender difference between the treatment and comparison groups is due to the oversampling of female students discussed earlier.

Table 18

Comparisons between means on select indicators for located and not located students

	Located			Not located		
	Treatment	Comparison	*t*	Treatment	Comparison	*t*
Female (Dummy)	0.43	0.58	4.76***	0.43	0.58	4.23***
Grade 5 Hindi (Z scores)	0.43	-1.59	0.95	0.43	-1.59	-0.92
Grade 5 English (Z scores)	2.05	-1.81	1.34	2.05	-1.81	-1.96
Grade 5 Math (Z scores)	2.60	-0.92	1.72	2.60	-0.92	-1.77
Grade 5 Science (Z scores)	2.49	-0.98	1.65	2.49	-0.98	-1.62
Grade 6 Hindi (Z scores)	-2.02	-1.43	0.61	-2.02	-1.43	0.30
Grade 6 English (Z scores)	-0.10	-1.04	1.34	-0.10	-1.04	-0.49
Grade 6 Math (Z scores)	0.77	-1.00	1.41	0.77	-1.00	-0.88
Grade 6 Science (Z scores)	0.50	-0.74	0.96	0.50	-0.74	-0.66
Grade 7 Hindi (Z scores)	-2.29	-1.71	0.16	-2.29	-1.71	0.29
Grade 7 English (Z scores)	-0.01	-0.94	0.77	-0.01	-0.94	-0.52
Grade 7 Math (Z scores)	-1.24	-1.14	0.65	-1.24	-1.14	0.06
Grade 7 Science (Z scores)	-0.87	-1.35	0.49	-0.87	-1.35	-0.24
Grade 8 Hindi (Z scores)	-1.97	-1.50	0.41	-1.97	-1.50	0.33
Grade 8 English (Z scores)	-0.35	-0.95	0.62	-0.35	-0.95	-0.49
Grade 8 Math (Z scores)	-1.48	-0.93	0.34	-1.48	-0.93	0.40
Grade 8 Science (Z scores)	-0.67	-1.22	0.45	-0.67	-1.22	-0.40
Grade 9 Hindi (Z scores)	-3.88	-0.14	-1.27	-3.88	-0.14	2.78**
Grade 9 English (Z scores)	-2.65	-0.17	-0.32	-2.65	-0.17	2.23*
Grade 9 Math (Z scores)	-9.03	-1.07	-1.34	-9.03	-1.07	3.59***
Grade 9 Science (Z scores)	-3.69	0.18	-2.10*	-3.69	0.18	2.88**
Grade 10 (Standardized test)	56.32	57.43	-0.62	56.32	57.43	0.68
N	*305*	*1819*		*250*	*758*	

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The above comparisons utilize all available data. Due to missing data the comparisons are based on sample sizes ranging from 305 to 93 for the analytic treatment group, 1819 to 704 for the analytic comparison group, 250 to 61 for the non-analytic treatment group and 758 to 222 for the non-analytic comparison group.

6.2.4.b. Missing data

Of the treatment and comparison students located for survey administration, complete data on relevant indicators was available for about 50 percent. This included data on pre-intervention outcomes (specifically seventh grade achievement), individual and household characteristics and school completion. Table 17 shows the number of students in both groups for whom complete data are available; these students were included the analytic sample. The treatment group included 160 students across 11 schools and the comparison group included 1161 students across 33 schools.[47] See Tables C.3 and C.4 in Appendix C for school-wise proportions of treatment and comparison students.

Outcome data were available for an additional 113 treatment cases and 486 comparison students. These cases could not be included in the analytic sample because of missing data on critical pre-intervention characteristics. The proportion of missing data on all relevant variables is presented in Table C.5 in Appendix C. Since the proportion of missing data on pre-intervention variables was well over five percent, methods to impute missing data were not utilized (Rubin, 1987). Reasons for missing data included unwillingness to participate in the research study and inconsistent responses to key survey questions.[48]

Figure 29 presents the gender breakdown amongst those in the analytic sample and those excluded from the analytic sample due to incomplete or missing data. The

[47] One comparison school that had begun offering the IBT program was removed from the sample (including all surveyed students belonging to this school) since the date of program inception could not be confirmed.

[48] A broader discussion on the challenges encountered during survey data collection and reasons for missing data is provided in Section 6.2.4.c. Suggestions for field-based research in similar contexts are also discussed.

gender distribution is similar to that observed in Figure 28; females constitute a little over 40 percent of the treatment group and about 58 percent of the comparison group.

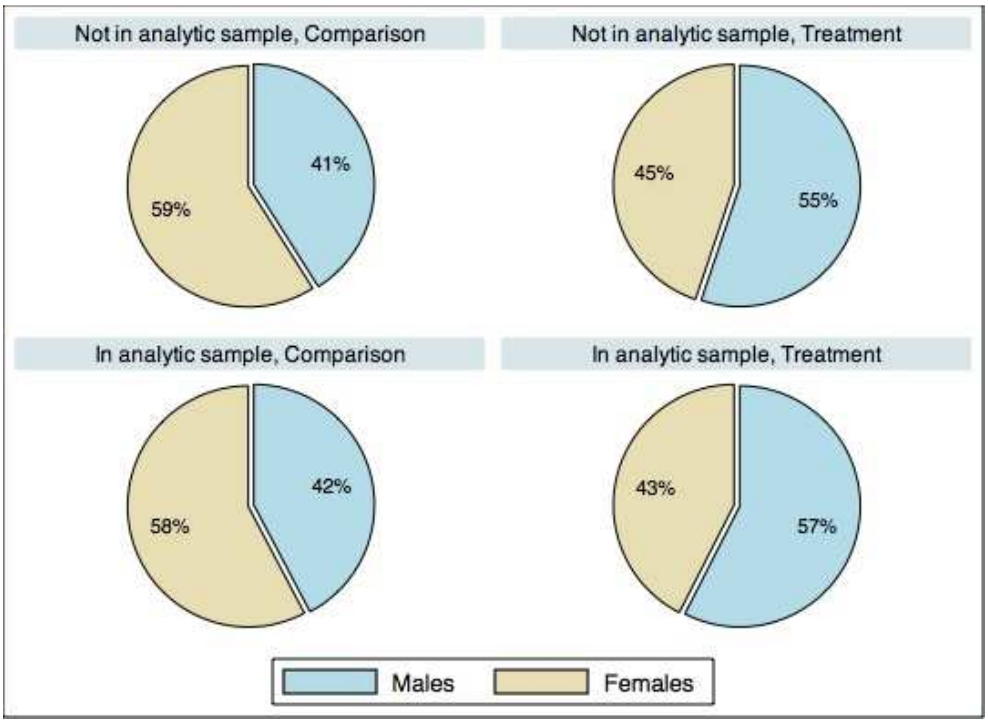

Figure 29. Gender breakdown of the analytic and non-analytic samples by treatment status

Comparisons were also made between the treatment and control groups in the analytic sample and in the sample excluded from the analytic group due to missing data. The means on select indicators for all these groups are presented in Table 19. For both samples, *t*-tests show significant differences in postsecondary enrollment (the outcome measured in this study), gender, religious affiliation, social group affiliation, household income, and context characteristics.

Table 19

Mean comparisons on select indicators for the analytic sample and those not included in the analytic sample, by treatment status

	Analytic sample			Excluded sample		
	Treatment	Comparison	*t*	Treatment	Comparison	*t*
Postsecondary enrollment (Dummy)	0.68	0.50	-4.24***	0.40	0.34	-1.26
Age	20.04	19.93	-0.98	19.18	19.76	4.17***
Female (Dummy)	0.43	0.58	3.65***	0.45	0.59	3.127**
Hindu (Dummy)	0.73	0.95	10.88***	0.71	0.97	10.50***
Social Group – Dalit (Dummy)	0.08	0.09	0.35	0.10	0.12	0.60
Social Group – Adivasi (Dummy)	0.36	0.24	-3.19**	0.29	0.24	-1.07
Social Group – OBC (Dummy)	0.18	0.30	3.24**	0.45	0.26	-4.41***
Household Size	4.81	4.41	-4.16***	4.59	4.37	-1.79
High income household (Dummy)	0.19	0.21	0.43	0.38	0.21	-3.81***
Medium income household (Dummy)	0.53	0.40	-3.02**	0.52	0.47	-1.05
Low income household (Dummy)	0.28	0.39	2.71**	0.10	0.32	4.76***
Household head in agriculture (Dummy)	0.62	0.53	-2.20*	0.38	0.45	1.30
Household head in self-employment (Dummy)	0.21	0.24	0.88	0.34	0.22	-3.22**
Household head in informal work (Dummy)	0.02	0.05	1.73	0.07	0.13	1.64
Household head in salaried work (Dummy)	0.16	0.19	0.94	0.05	0.11	1.76
Grade 7 Performance (Raw score)	63.83	58.04	-6.26***	62.42	56.10	-3.91***
Grade 7 performance (Z scores)	0.37	0.52	0.17	-0.61	-0.58	0.02
Distance to nearest town/city	17.06	10.76	-8.89***	11.63	11.36	-0.45
Access to public transport (Dummy)	0.98	0.93	-2.48*	0.91	0.99	6.18***
Village population	9186.25	6468.64	-5.73***	12492.41	4425.70	-13.45***
N	*160*	*1161*		*56*	*275*	

Note. The above comparisons utilize all available data. Due to missing data the comparisons are based on sample sizes ranging from 56 to 145 for the excluded treatment group and from 275 to 658 for the excluded comparison group.

6.2.4.c. Data collection challenges and suggestions for field-based research

As mentioned in the previous sections, the data for this study were collected from several different sources owing to the post-hoc design of the evaluation. While the data from student surveys provided outcome information necessary for estimating impacts, the data from administrative records (mainly, students' past educational achievement) were essential for creating a statistically matched group of treatment and control students who could then be compared on relevant outcomes. Data from these sources were merged such that each student record included key pieces of data from administrative records and data from field surveys. Students for whom either administrative records were missing and/or survey information was missing or incomplete could not be included in the analytic sample.

There were three main challenges encountered in gathering complete data on students' past academic achievement – missing records at the school, incomplete data in the case of students who completed middle and secondary levels at different schools, and incomplete information due to student mobility or dropping out.

All government-aided schools in India maintain records on students' attendance and achievement dating back to about 10 years. These records are typed or hand-written by teachers and school staff and stored at the school. In a handful of comparison schools, records for the years relevant to the study were either lost or misplaced and thus unavailable during data collection.

The student records were collected for the period between grade five and grade 10. In several cases students attended grades five to seven in one of the middle schools in the village and grades eight to ten in a different secondary school. Thus, for these

students academic achievement information had to be collected from two different schools. Since several middle schools can feed in to a village secondary school, collecting academic achievement information prior to grade eight implied collecting administrative data from all the feeder schools. However, this was not possible due to time and cost constraints.

As mentioned above, one reason for incomplete information on student achievement was student mobility or dropping out. Often, records were available for a student in the eighth grade but not in subsequent grades. It could not be ascertained whether this was because the student had changed schools after eighth grade, moved to a new location or dropped out of school altogether. It is possible that school records indicate reasons for absence or discontinuation and can be resolved on a case-by-case basis.

The challenges faced during survey administration included locating students in their villages three to four years after they had completed grade 10, cultural barriers preventing female participation, and perceived complexity of the survey instrument. These constraints limited the number of students for whom outcome information was available.

It was assumed based on past data collection experiences in similar contexts that although there was a high likelihood of students leaving their villages for further education or employment their families would still be living in the same villages. A majority of the students in both the treatment and comparison groups were located for survey administration except in those sub-districts that shared a border with the neighboring state. It was learned during data collection that these villages include a large

proportion of migrant workers and many schools offer residential accommodation to students. It is therefore possible that many of the students who were being sought for data collection moved away after completing school and were therefore not located.

In the case of female participants, field investigators were able to locate many students and/or their families. In several cases it was learned that female students were married. Oftentimes female participants were unwilling to participate in the research due to cultural propriety. All of the field investigators administering the survey were male and it was deemed culturally inappropriate for married women to talk to unknown men.

Feedback from field supervisors revealed that survey administrators and participants perceived the instrument to be long and complex which resulted in erroneous or incomplete responses on the survey.

The challenges discussed above largely limited the size of the analytic sample in this study, which in turn has implications for the reliability of the findings on the impact of secondary level vocational education on further enrollment. The experience from this data collection effort however provides useful recommendations for field-based research in India and other developing countries.

First, developing concise surveys with simple response options (or simple coding schemes) can ensure that surveys are not only completed in a time effective way but that responses are noted with minimal errors. Second, training and shadowing field staff prior to data collection and during the first few days/weeks is imperative to the success of the data collection effort. While training helps orient field investigators to the purpose of the research, develop familiarity with the survey instrument and coding of responses, shadowing can help field investigators learn how to respond to unexpected situations

during data collection. Third, an awareness of social and cultural norms in the region under study can significantly improve the data collection effort. Recruiting local field investigators familiar with the realities of the region and incorporating their input in the planning stages, talking to community members and key stakeholders prior to data collection, pilot testing survey instruments, and recruiting female staff members for easier access to female participants could go a long way in successfully collecting field data. Next, studies that require student achievement data must first assess availability and accessibility issues. In several developing countries including India, student achievement data are not readily electronically available. Studies should thus be designed to optimize available data, adopt innovative and efficient methods for data collection (these often require significant time and resources), and extend the size of the sample – if time and cost permits – to account for unanticipated situations that limit data availability or usability. Finally, despite all the efforts made to collect accurate and complete data from a large sample, nonresponse and missing data are frequently encountered in survey research. Collecting information on reasons for nonresponse can be useful in making inferences about the direction of nonresponse bias and can provide evidence for improving future data collection efforts.

6.2.5 Data Analysis

Matching methods, of which propensity score matching is one, provide a statistical technique to equate or balance the distribution of covariates in the treated and non-treated groups. Thus, they aim to mirror randomized experiments which guarantee that assignment to treatment remains independent of observed and unobserved

characteristics (Rosenbaum, 2002; Stuart, 2010). Propensity score matching, a type of matching method, uses propensity scores to achieve this objective. Propensity scores are estimates of the probability of being in the treatment group conditional on the observed covariates (Rosenbaum & Rubin, 1983). Formally, the propensity score for individual *i* can be expressed as:

$$e_i(X_i) = P(T_i = 1|X_i)$$

It should be noted that a key property of propensity scores is "balance". This means that at each value of the propensity score, the distribution of covariates is the same in the treatment and control groups.

Non-experimental studies using matching methods rely on two key assumptions: (1) individual *i*'s assignment to treatment *T* is independent of the potential outcome given the covariates in the model (also called the "ignorability" or "unconfoundedness" or "no hidden bias" assumption), and (2) all values of the covariates are associated with a positive probability of receiving treatment (Rosenbaum & Rubin, 1983; Rosenbaum, 2002). The implication of these assumptions along with the balancing property of propensity scores is that conditional on pre-treatment variables X_i, there are no systematic pre-treatment differences between the treatment and control groups. Further, comparisons of individuals with the same propensity score in the treatment and control group produces an unbiased estimate of the treatment effect at that propensity score value (Rosenbaum, 2002; Stuart, 2010). The treatment effect is formalized as:

$$\tau_{ATT} = E_{P(X)|T=1}\{E[Y(1)|T = 1, P(X)] - E[Y(0)|T = 0, P(X)]\}$$

Following Stuart (2010), application of the propensity score method in this study proceeded in four stages. First, pre-intervention academic performance, time-invariant household characteristics and demographic characteristics were used to estimate the propensity scores or probability of being treated. Second, several matching methods were used to find "best matches" between treatment and comparison units. These were subject to diagnostic analyses in the third stage to test the quality of the resulting matches. In the fourth and last stage, the treatment effect was estimated.

6.2.5.a. Estimating propensity scores

The choice of variables in the model estimating propensity scores is driven by the ignorability assumption. Variables related to the outcome (postsecondary enrolment) as well as those related to treatment assignment were included in the equation. Students participating in IBT could be classified as those participating in IBT because the only school in their village offered the program, or those who chose to enrol in a school offering IBT because of an underlying interest in the program and what it offers. The propensity score equation therefore attempted to model this student motivation by including student-level and context-level characteristics associated with participation in education and training. Students' performance in the seventh grade (the year just prior to IBT), their demographic (gender, social and religious affiliation) and household characteristics (including, household assets, occupation of the head of the household, household size, number of children in the household, number of employed persons in the household), and context-level characteristics (size of their village, access to public transportation and a major town/city) were used to predict treatment status. Prior

research on education and training participation decisions has shown that individual and household-level characteristics are among the strongest predictors of participation in education and training.

In the case of IBT, treatment assignment is related to attending an IBT school. Of the 8,829 rural secondary schools (recognized-aided and recognized-unaided) in the state of Maharashtra[49], less than one percent offered IBT in 2011. As mentioned in the previous section, comparison schools were selected such that they were as similar as possible to treatment group schools in terms of their size, infrastructure and resources. Nearly all comparison schools in the sample offered some type of academic or skill training program to students.

6.2.5.b. Matching

Of the various matching methods available to estimate the average treatment effect on the treated (ATT), nearest neighbor matching, optimal matching, full matching and subclassification were implemented in this study.[50] Matching was carried out for the entire analytic sample and both, matching with and without replacement, were considered.

In the case of nearest neighbor matching, each treatment unit is matched with the closest (in terms of distance measure) control unit (Rosenbaum, 2002). Although a large number of observations for which no matches are found get discarded during this

[49] Figures based on raw data from http://semisonline.net/
[50] Subclassification methods allow for measuring the average treatment effect (ATE) in addition to ATT. All matching was done using the MatchIt program in R (3.0.2)

process, with large sample sizes nearest neighbor matching results in groups that are quite similar to those obtained through optimal matching (Gu & Rosenbaum, 1993).

The advantage of optimal matching over nearest neighbor is that the method takes into account the overall set of matches when choosing individual matches (Rosenbaum, 2002). Thus, although the number of matches produced by optimal matching might not be greater than the nearest neighbour method, optimal matching ensures that the distance between matched treatment and control units is reduced.

Finally, full matching and subclassification methods create a series of matched sets with at least one unit from the treatment group and at least one unit from the control group. The sets are designed such that the distance between the treated and control units is minimal (as in the case of optimal matching). The advantage lies in the fact that these methods do not discard control units for whom data is available leading to better efficiency and precision (Forston et al., 2012).

6.2.5.c. Diagnostics

Ho et al. (2007) note that the main diagnostic of success in matching is balance and the number of observations remaining after matching. Balance, according to them, involves ensuring common support (i.e. pruning cases falling outside the empirical density of treatment and comparison units), a step often skipped in the applied literature, and adjusting densities that do overlap to have the same height. To check for common support this study applied the more conservative “convex hull” approach suggested King and Zeng (2006, 2007), as well as the less conservative comparison of propensity

score distributions.[51] To test if comparison (or treatment) observations were outside the "convex hull" of the treatment (or comparison) group, the MatchIt program in R (3.0.2) was used. The less conservative method discards units that fall outside the range of the propensity scores of the other group (Heckman et al., 1997; Dehejia & Wahba, 1999) and was carried out in Stata.

Each matching method was assessed by examining the standardized difference in means of the propensity score and the full sets of covariates in the treatment and comparison groups (Rubin, 2001). Plots of the standardized mean bias before and after matching were also examined for each matching method.

6.2.5.d. Estimating the treatment effect

In the final stage, regression analysis was used to estimate the mean impact of the program on students' postsecondary enrollment. Imbens and Wooldridge (2009) suggest including regression adjustments in the outcome analysis to safeguard against misspecification in either the propensity score model or the outcome model.

6.3 Results

The analytic sample to measure the effect of IBT on school completion included 160 treatment students and 1161 comparison students. The treatment students in the analytic sample cannot be considered representative of all treatment students who participated in IBT in 2006 and 2007 due to the proportion of cases excluded because of

[51] See Ho et al. (2007) and King and Zeng (2006, 2007) for a discussion of each method.

missing data (see Section 6.2.4.a). Nevertheless, propensity score matching was carried out with the available sample of treatment and comparison students. The results discussed in this section are thus suggestive of likely trends.

During the pre-processing stage, school and student survey responses showed that three treatment schools offered IBT to one division each in grades 8 to 10 and had two divisions that were not offered the IBT treatment. While students from the non-IBT divisions in these schools could be the best comparisons for treatment group students at the same schools, student survey responses revealed some contamination in the sample. Specifically, students not assigned to the IBT division in grade eight responded to participating in IBT. A second analytic sample was thus created excluding all surveyed students from these three schools. All further analysis was carried out for two analytic samples – sample I included students from these three schools, and sample II excluded these students.

6.3.1 Propensity Score Equation Results

Multivariate logistic regression was used to estimate the propensity score equation. The dependent variable (treatment status) was regressed on all relevant covariates (see Table C.6 in Appendix C). The outcome indicator[52] was not used in the estimation of propensity scores to avoid inducing bias (Ho et al., 2007). The logit models correctly classified 90 percent of the cases in both analytic samples I and II. The

[52] The evaluation study planned to measure the effect of IBT on grade 10 performance, postsecondary enrollment and wages. A significantly large proportion of treatment and control cases were missing information on wages. This analysis is therefore limited to measuring the impact on postsecondary enrollment along with a descriptive analysis of grade 10 performance and postsecondary enrollment patterns for a subset of school completers

predicted probabilities produced as a result of the logistic regression analysis formed the propensity scores. Results of the logistic models are presented in Table C.6 in Appendix C.

6.3.1.a. Assessing Common Support

The propensity scores for each analytic sample were examined separately for the treatment and comparison groups. The distributions were graphed and are presented in Figures 30 and 31, respectively. The figures show that the distributions do not overlap to a large extent (to a slightly lesser extent in Analytic Sample II than Analytic Sample I). Following the suggestion of Dehejia & Wahba (1999), cases that were outside the common support of the estimated propensity scores were identified for deletion. Discarding these cases reduces model dependence, variance and mean squared error (Ho et al., 2007). A total of 56 cases (52 from the comparison group and four from the treatment group) in sample I and 10 cases (9 from the comparison group and one from the treatment group) in sample II were identified for possible deletion.[53]

6.3.2 Matching Results

Given the distribution of propensity scores observed in Figures 30 and 31, a variety of matching methods were tested. One-to-one nearest neighbor matching without replacement was tested with and without cases outside of common support. One-to-many

[53] The more conservative "convex hull" approach (King & Zeng, 2007) was also carried out to assess common support and showed that only four cases were in the "convex hull" of treated and comparison units.

nearest neighbor matching was not tested because the convex hull analysis showed that several cases were outside of common support. Optimal matching with a 2:1 comparison to treatment ratio was also selected. Optimal matching is similar to nearest neighbor matching but optimizes the average absolute distance among matched pairs and is especially useful in cases where fewer appropriate comparison units are available for the treatment units (Gu & Rosenbaum, 2003). Full matching, full matching with restrictions and subclassification were also tested. These methods create matched sets such that each set has at least one treatment or comparison unit. They were tested to see if they offered better balance at the cost of increased variance. Standardized bias under 0.25 standard deviation units (Ho et al., 2007) was used as the decision criteria to assess the quality of the matches.

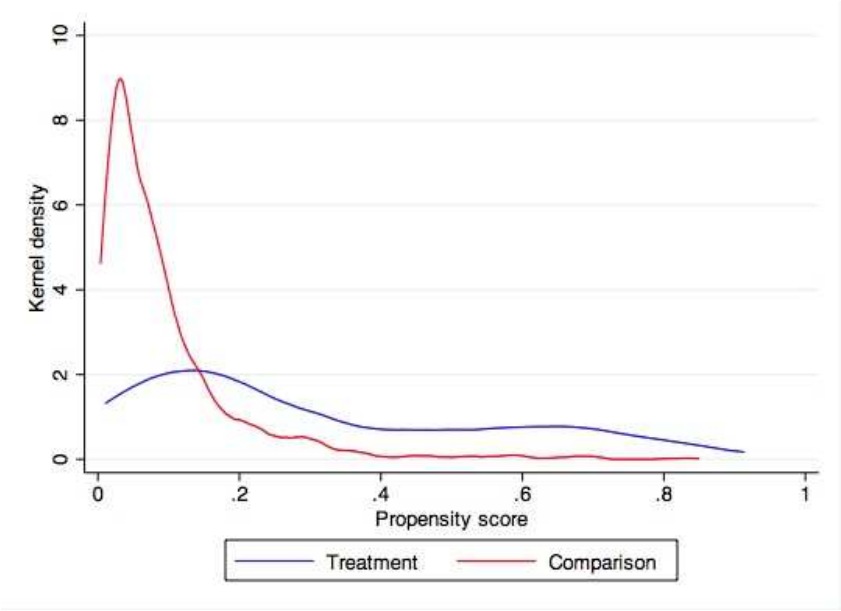

Figure 30. Distribution of propensity scores in Analytic Sample I

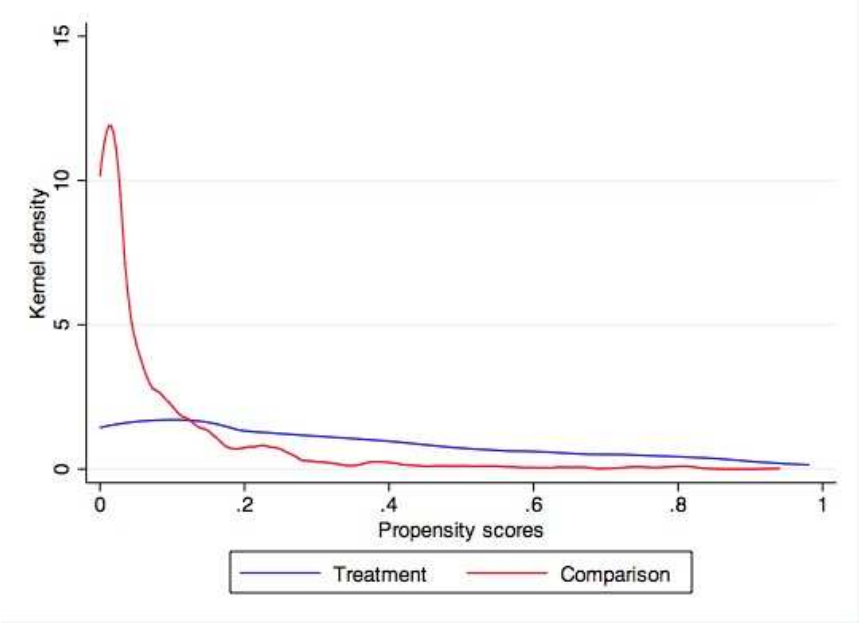

Figure 31. Distribution of propensity scores in Analytic Sample II

6.3.2.a. One-to-one Nearest Neighbor Matching (full sample)

This method discarded three treatment cases and matched the remaining 157 with 157 comparison cases for analytic sample I. In the case of analytic sample II 114 treatment cases (one treatment case was discarded) were matched. The standardized bias for all covariates was well within the 0.25 criterion for analytic sample I. For sample II however, the standardized bias for the propensity score was above the acceptable quarter of a standard deviation unit.

6.3.3.b. One-to-one Nearest Neighbor Matching (with cases in common support)

Of the 156 cases in the treatment sample (after excluding four treatment and 52 comparison cases outside common support), 155 were matched with counterpart comparison cases. In the second analytic sample, all 114 treatment cases (one was

removed from the sample before matching) were matched to 114 comparison cases. The standardized bias for analytic sample I was just under 0.25 for the propensity scores and for the dummy variable indicating whether a student was Hindu or from another religious group. In analytic sample II, the standardized bias for all covariates was within the acceptable limit.

6.3.3.c. 2:1 Optimal Matching (full sample)

This method matched all treatment units (160 in the case of sample I and 115 in the case of sample II) to 320 and 230 comparison cases respectively, such that the global distance across all matched pairs was minimized. The standardized bias for the propensity scores and the indicator for religious group was higher than the acceptable limit in both samples. In the second analytic sample the standardized bias for village size was also high - 0.27.

6.3.3.d. Full Matching without constraints (full sample)

Full matching creates matched sets of treated and comparison units with varying numbers of treated and comparison units in each set. The number of treatment and comparison units in each set depend on the relative number of treated and comparison units with similar propensity scores (Stuart & Green, 2008). One potential issue that can arise with full matching is that the varying ratios of treatment to comparison cases can increase the variance in impact estimates. As seen in Figures 32 and 33, comparison units towards the tail end of the propensity score distribution carried significantly higher

weights than comparison cases in the lower end of the distribution. While all the covariates in sample I met the balance requirements, three variables in sample II were close to or above the acceptable standardized bias criterion.

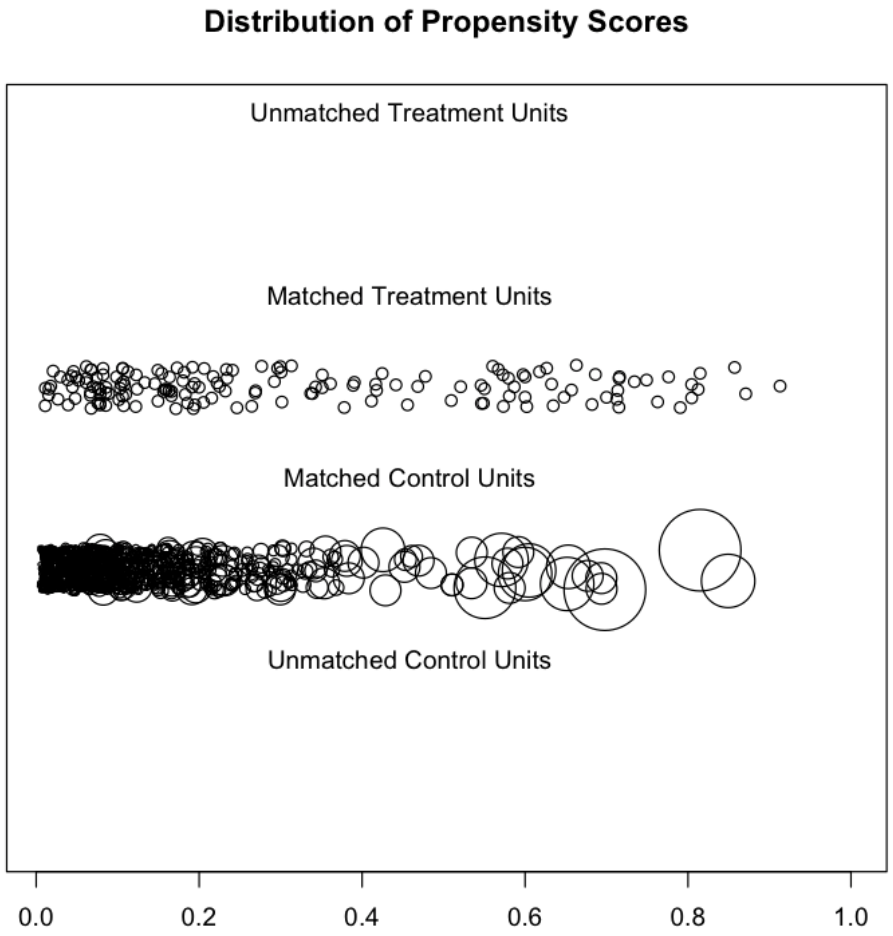

Figure 32. Distribution of propensity scores using Full Matching (sample I)

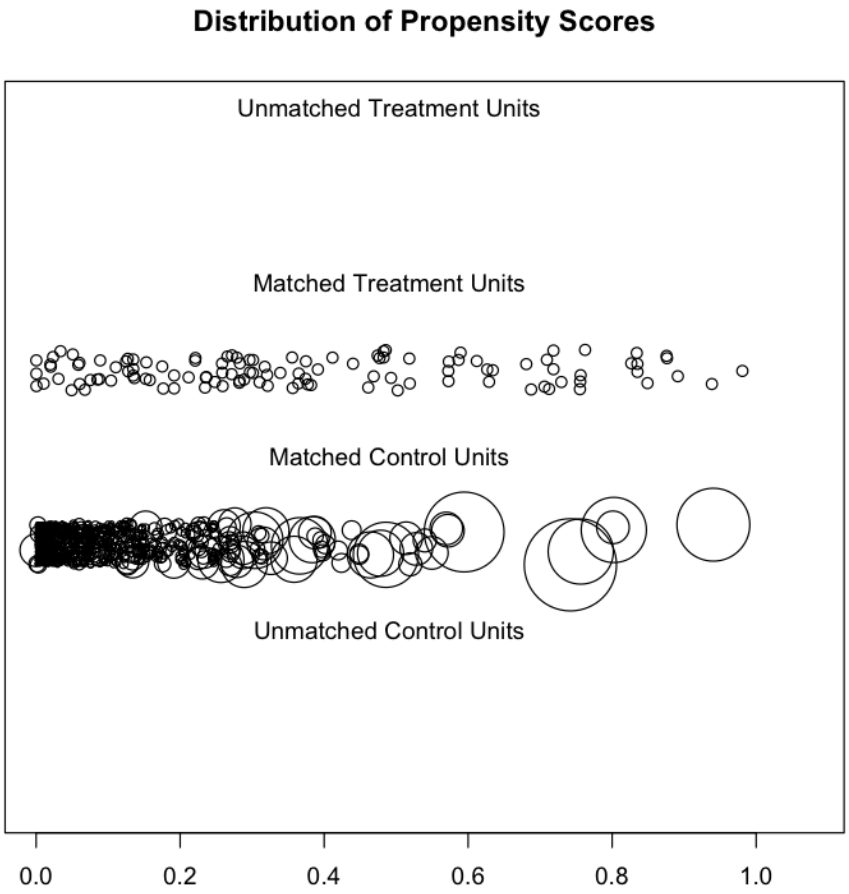

Figure 33. Distribution of propensity scores using Full Matching (sample II)

6.3.3.e. Full Matching - with constraints (full sample)

In order to address disproportionately high weights assigned to some control units (as seen in the figures above) full matching was repeated with constraints such that the ratio of treatment to comparison units was not lower than half of that in the original sample and not higher than double of that in the original sample. The original analytic sample had seven comparison units for every treatment unit. The constraints imposed a 2:7 and 1:14 ratio for treatment to control units. Similarly, for sample II, full matching

was constrained such that the ratio of treated to comparison units was between 1:4 and 1:16. This method did not provide good balance across several covariates.

6.3.3.f. Subclassification (full sample)

Six subclasses were formed using this method with each subclass having about the same number of treated units and varying number of comparison units. For both analytic samples, subclass 5 and subclass 6 had considerably fewer comparison units and showed high standardized bias across several covariates. The method was modified to collapse six subclasses in to four subclasses. However, this did not improve balance.

Tables 20a and 20b present the standardized bias for all covariates and the propensity score across each method tested above. A graphical comparison of the distribution of the standardized mean difference across all models (for both samples) is presented in Figures 34 and 35.

Table 20a

Comparisons of standardized bias across all covariates after matching for analytic sample I

	Model 1.1	Model 1.2	Model 1.3	Model 1.4	Model 1.5	Model 1.6	Model 1.7
Propensity score	0.20	0.23	**0.43**	0.00	**0.64**	**0.86**	**0.86**
Female (Dummy)	0.10	0.04	0.07	0.09	0.23	**0.31**	**0.31**
Hindu	0.14	**0.25**	**0.32**	0.05	0.40	**0.51**	**0.51**
Social Group - Dalit	0.02	0.05	0.05	0.15	0.01	0.03	0.03
Social Group - Adivasi	0.13	0.14	0.16	0.23	0.22	**0.24**	**0.24**
Social Group - OBC	0.12	0.13	0.07	0.00	0.25	**0.32**	**0.32**
Household Size	0.01	0.12	0.01	0.12	0.19	**0.35**	**0.35**
High income household	0.19	0.08	0.06	0.14	0.03	0.04	0.04
Medium income household	0.04	0.13	0.02	0.09	0.11	**0.25**	**0.25**
Household head in agriculture	0.13	0.13	0.01	0.09	0.13	0.19	0.19
Household head in self-employment	0.14	0.17	0.01	0.08	0.02	0.08	0.08
Household head in salaried work	0.03	0.02	0.00	0.01	0.08	0.08	0.08
Grade 7 performance (Z scores)	0.00	0.10	0.04	0.01	0.01	0.02	0.02
Distance to nearest town/city	0.08	0.03	0.20	0.02	**0.38**	**0.59**	**0.59**
Access to public transport	0.09	0.06	0.00	0.02	**0.28**	**0.37**	**0.37**
Distance to bus station	0.01	0.06	0.03	0.00	0.03	0.03	0.03
Village size	0.02	0.06	0.10	0.09	0.22	**0.35**	**0.35**
Female*Hindu	0.09	0.04	0.13	0.01	**0.35**	**0.48**	**0.48**
Female*Dalit	0.00	0.11	0.06	0.02	0.17	**0.29**	**0.29**
Female*Adivasi	0.12	0.05	0.05	0.18	0.08	0.09	0.09
Female*OBC	0.08	0.06	0.11	0.03	**0.39**	**0.51**	**0.51**
Female*Grade 7 performance	0.05	0.11	0.03	0.07	0.06	0.07	0.07
Distance to nearest town/city*Grade 7 performance	0.02	0.06	0.01	0.06	0.03	0.06	0.06

N	*157*	*155*	*160*	*160*	*160*	*160*	*160*

Note. The standardized bias is the weighted difference in means divided by the standard deviation in the full comparison group. Models are as follows – 1.1 = 1:1 nearest neighbor matching, 1.2 = 1:1 nearest neighbor matching without cases not on common support, 1.3 = optimal matching, 1.4 = full matching, 1.5 = constrained full matching, 1.6 = subclassification, 1.7 = subclassification with 4 subclasses.

Table 20b

Comparisons of standardized bias across all covariates after matching for analytic sample II

	Model 2.1	Model 2.2	Model 2.3	Model 2.4	Model 2.5	Model 2.6	Model 2.7
Propensity score	**0.35**	0.22	**0.62**	0.00	**1.06**	**1.15**	**1.15**
Female (Dummy)	0.14	0.08	0.18	0.19	**0.38**	**0.45**	**0.45**
Hindu	0.19	0.19	0.31	0.02	**0.49**	**0.53**	**0.53**
Social Group - Dalit	0.10	0.04	0.03	0.04	0.12	0.10	0.10
Social Group - Adivasi	0.24	0.17	0.11	**0.24**	0.11	0.11	0.11
Social Group - OBC	0.07	0.05	0.02	0.21	**0.24**	**0.30**	**0.30**
Household Size	0.03	0.11	0.03	0.01	**0.29**	**0.35**	**0.35**
High income household	0.16	0.08	0.07	0.07	0.06	0.10	0.10
Medium income household	0.09	0.10	0.06	0.03	**0.26**	**0.29**	**0.29**
Household head in agriculture	0.17	0.03	0.01	0.08	**0.29**	**0.32**	**0.32**
Household head in self-employment	0.12	0.05	0.02	0.21	0.14	0.14	0.14
Household head in salaried work	0.06	0.10	0.03	0.11	0.16	0.17	0.17
Grade 7 performance (Z scores)	0.03	0.03	0.14	0.10	0.09	0.07	0.07
Distance to nearest town/city	0.01	0.05	0.22	**0.25**	**0.37**	**0.40**	**0.40**
Access to public transport	0.07	0.04	0.03	0.03	**0.32**	**0.33**	**0.33**
Distance to bus station	0.11	0.00	0.20	0.04	**1.11**	**1.27**	**1.27**

Village size	0.22	0.15	**0.27**	0.18	**0.32**	**0.31**	**0.31**
Female*Hindu	0.15	0.02	0.17	0.08	**0.48**	**0.56**	**0.56**
Female*Dalit	0.09	0.00	0.05	0.00	**0.45**	**0.42**	**0.42**
Female*Adivasi	0.07	0.07	0.11	0.09	0.18	0.18	0.18
Female*OBC	0.04	0.04	0.04	0.06	**0.42**	**0.53**	**0.53**
Female*Grade 7 performance	0.12	0.00	0.12	**0.30**	0.02	0.02	0.02
Distance to nearest town/city*Grade 7 performance	0.09	0.09	0.14	0.18	0.05	0.05	0.05
N	*114*	*105*	*115*	*115*	*115*	*115*	*115*

Note. The standardized bias is the weighted difference in means divided by the standard deviation in the full comparison group. Models are as follows – 2.1 = 1:1 nearest neighbor matching, 2.2 = 1:1 nearest neighbor matching without cases not on common support, 2.3 = optimal matching, 2.4 = full matching, 2.5 = constrained full matching, 2.6 = subclassification, 2.7 = subclassification with 4 subclasses.

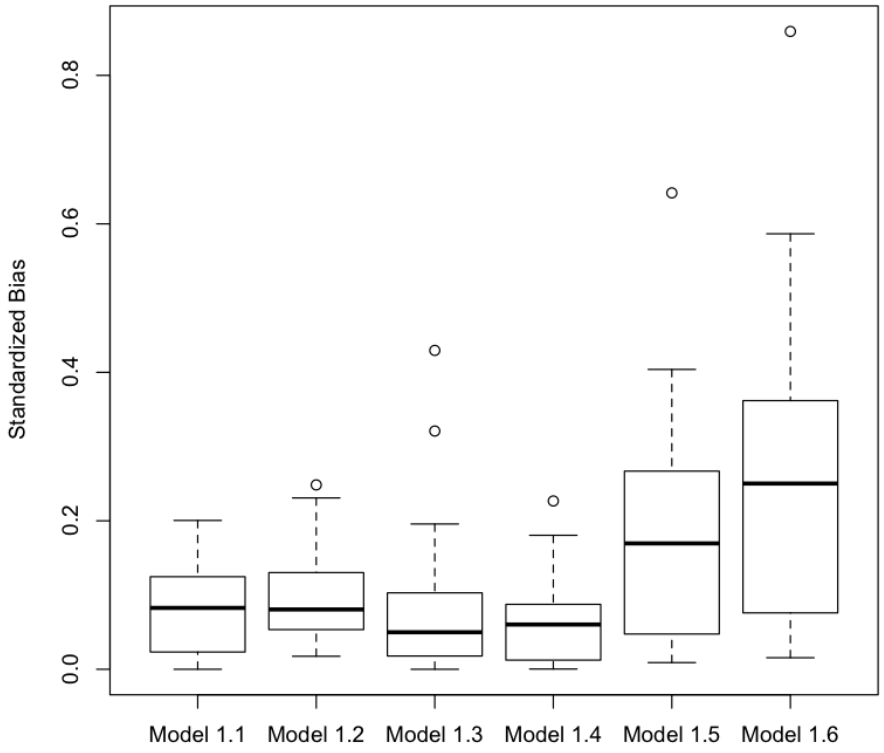

Figure 34. Boxplots of absolute standardized bias for covariates in Table 20a. The standardized bias is the weighted difference in means divided by the standard deviation in the full comparison group. Models are as follows – 1.1 = 1:1 nearest neighbor matching, 1.2 = 1:1 nearest neighbor matching without cases not on common support, 1.3 = optimal matching, 1.4 = full matching, 1.5 = constrained full matching, 1.6 = subclassification with 4 subclasses.

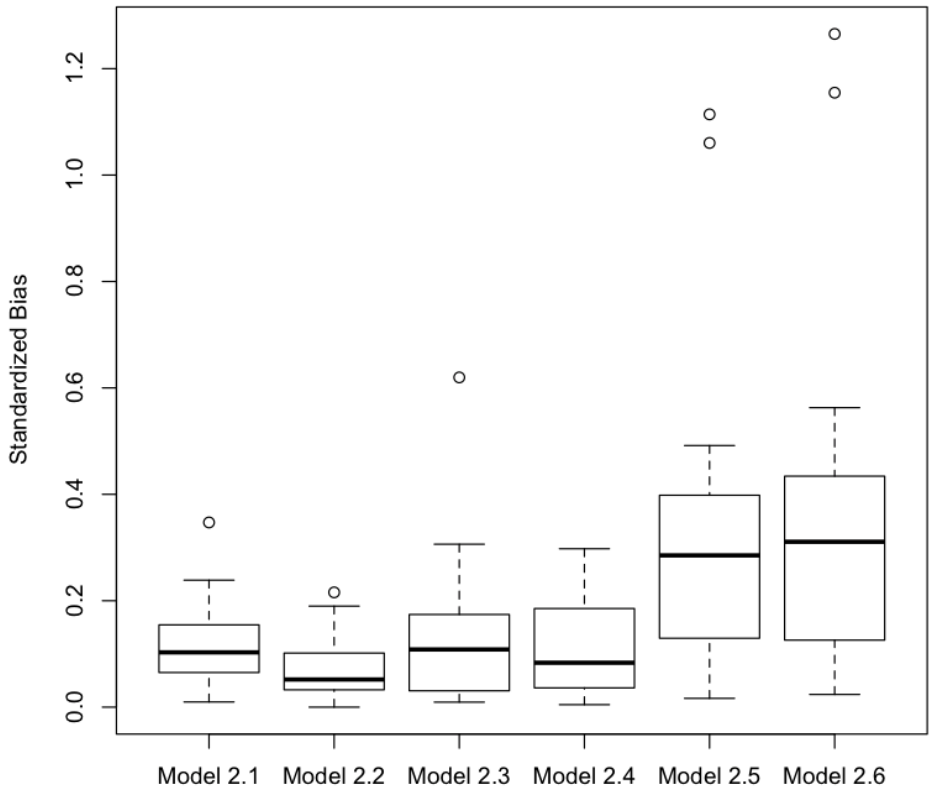

Figure 35. Boxplots of absolute standardized bias for covariates in Table 20b. The standardized bias is the weighted difference in means divided by the standard deviation in the full comparison group. Models are as follows – 2.1 = 1:1 nearest neighbor matching, 2.2 = 1:1 nearest neighbor matching without cases not on common support, 2.3 = optimal matching, 2.4 = full matching, 2.5 = constrained full matching, 2.6 = subclassification with 4 subclasses.

Based on the results discussed above, the matched sample from model 1.1 (one-to-one nearest neighbor matching using the entire original analytic sample) and the matched sample from model 2.2 (one-to-one nearest neighbor matching after discarding cases not on common support) were used for the outcome analysis. Figures 36 and 37 show the

change in standardized bias across all covariates in the original data and the matched data for both samples.[54]

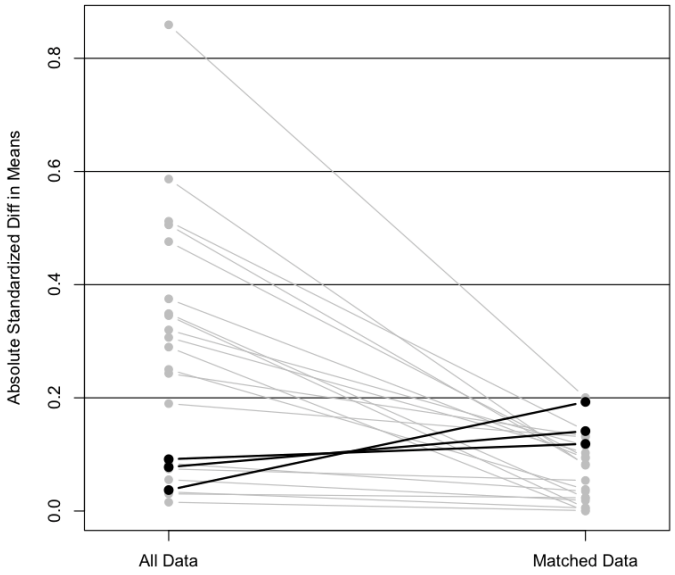

Figure 36. Change in absolute standardized bias after 1:1 nearest neighbor matching (sample I)

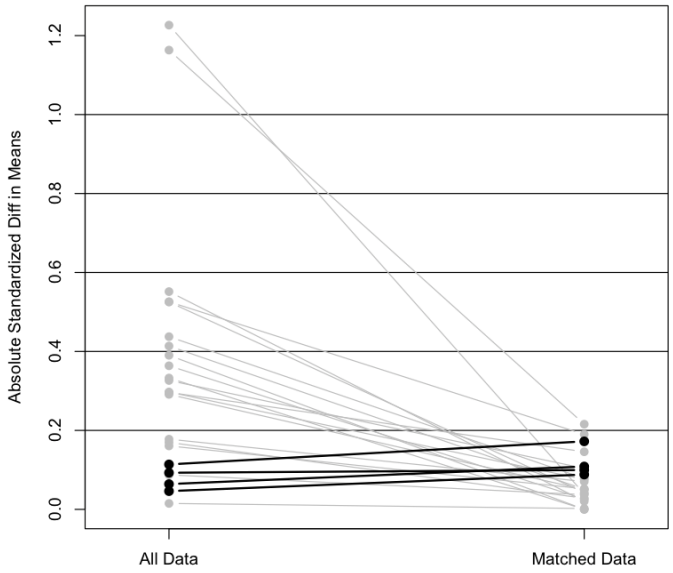

Figure 37. Change in absolute standardized bias after 1:1 nearest neighbor matching and discarding cases not on common support (sample II)

[54] See Appendix C for jitter plots and histograms for the matched data.

6.3.3 Descriptive Results for Matched Data

Tables 21a and 21b show the means and standard deviations for select indicators in the original analytic sample and the matched analytic sample for the treatment and control groups, respectively. The matched sample included three fewer cases in the treatment group than the original analytic sample and therefore there were no significant differences between the means in the two samples. The comparison sample however reduced from 1161 cases to 157 cases and Table 21b shows the corresponding change in means of selected variables.

A little over half of all students in the analytic sample reported enrolling in some further education (beyond grade 10). The proportion of treatment students enrolling in further education was significantly higher than the proportion reported in the comparison group.

The proportion of female students in the treatment and comparison groups was similar before matching. In the matched sample, the proportion of comparison females was significantly lower (27 percent) than that in the treatment group (43 percent).

The distribution of social groups across the treatment and comparison samples showed some differences but these were unchanged in the matched sample. While most of the comparison group students self-identified as Hindu, the treatment group was a little more mixed with about 70 percent identifying as Hindu. Similarly, the proportion of students in the *OBC* group was higher in the comparison group than the treatment group. The treatment group had a larger proportion of *Adivasi* students than the comparison group.

Students were classified into one of three household income categories. The number of assets in the household was used as a proxy for household income. Both groups, treatment and comparison, consisted of a larger proportion of students from medium income households (about 50 percent) followed by about 30 percent in the lowest income category.

In terms of the household head's occupation, the majority of students reported this as agricultural work. Over 50 percent of the students in the treatment and comparison group belonged in this category. Informal work was the smallest occupational category with 1 to 2 percent making up this group.

The average previous academic performance for the comparison group was similar to that of the treatment group. The comparison group however showed more variation in grade 7 performance.

With regard to context variables, the distance and access to public transport measures indicate that treatment school students were located slightly farther away from major towns and bus stations than their comparison group counterparts. Comparison group students tended to belong to, on average, slightly larger villages – although there was significantly higher variation in village size in the matched treatment group.

Table 21a

Means of relevant indicators for the treatment group in the original and matched data

	Unmatched sample I		Matched sample I	
	Mean	SD	Mean	SD
Postsecondary enrollment (Dummy)	0.68	0.47	0.68	0.47
Age	20.04	1.38	20.05	1.39
Female (Dummy)	0.43	0.50	0.43	0.50
Hindu (Dummy)	0.73	0.45	0.74	0.44

Social Group – Dalit (Dummy)	0.08	0.27	0.08	0.28
Social Group – Adivasi (Dummy)	0.36	0.48	0.34	0.48
Social Group – OBC (Dummy)	0.18	0.39	0.18	0.39
Household Size	4.81	1.16	4.76	1.10
High income household (Dummy)	0.19	0.40	0.20	0.40
Medium income household (Dummy)	0.53	0.50	0.52	0.50
Low income household (Dummy)	0.28	0.45	0.29	0.45
Household head in agriculture (Dummy)	0.62	0.49	0.61	0.49
Household head in self-employment (Dummy)	0.21	0.41	0.21	0.41
Household head in informal work (Dummy)	0.02	0.14	0.02	0.14
Household head in salaried work (Dummy)	0.16	0.36	0.16	0.37
Grade 7 Performance	63.83	10.57	63.78	10.60
Grade 7 performance (Z scores)	0.37	9.35	0.42	9.31
Distance to nearest town/city	17.06	10.74	16.85	10.73
Access to public transport (Dummy)	0.98	0.14	4.83	6.64
Distance to bus station	4.80	6.59	0.98	0.14
Village size	9186.25	7790.34	9151.59	7858.34
N	*160*		*157*	

Note. Sampling weights are used in the calculation of means and account for the oversampling of female students in the comparison group.

Table 21b

Means of relevant indicators for the comparison group in the original and matched data

	Unmatched sample I		Matched sample I	
	Mean	SD	Mean	SD
Postsecondary enrollment (Dummy)	0.50	0.50	0.52	0.50
Age	19.93	1.32	20.34	1.29
Female (Dummy)	0.45	0.49	0.27	0.49
Hindu (Dummy)	0.95	0.21	0.80	0.40
Social Group – Dalit (Dummy)	0.09	0.29	0.09	0.29
Social Group – Adivasi (Dummy)	0.24	0.43	0.28	0.45
Social Group – OBC (Dummy)	0.30	0.46	0.23	0.42
Household Size	4.41	1.14	4.77	1.37
High income household (Dummy)	0.21	0.41	0.12	0.33
Medium income household (Dummy)	0.40	0.49	0.54	0.50
Low income household (Dummy)	0.39	0.49	0.34	0.48

Household head in agriculture (Dummy)	0.53	0.50	0.55	0.50
Household head in self-employment (Dummy)	0.24	0.43	0.27	0.44
Household head in informal work (Dummy)	0.05	0.22	0.01	0.11
Household head in salaried work (Dummy)	0.19	0.39	0.17	0.38
Grade 7 Performance	58.04	11.11	59.74	11.58
Grade 7 performance (Z scores)	0.52	10.30	0.43	10.50
Distance to nearest town/city	10.76	8.03	15.98	11.05
Access to public transport (Dummy)	0.93	0.25	4.79	4.89
Distance to bus station	5.02	5.01	0.97	0.18
Village size	6468.64	5261.78	9333.34	6626.28
N	*1161*		*157*	

Note. Sampling weights are used in the calculation of means and account for the oversampling of female students in the comparison group.

The means of all relevant variables in sample II are presented in Tables C.7 (treatment group) and C.8 (comparison group) in Appendix C. The demographic and household characteristics are similar to those discussed in the case of sample I. Appendix C also includes a comparison of the schools offering IBT and those not offering IBT. (See Table C.9).

6.3.4 Results of the Outcome Analysis

To estimate the effect of IBT participation on postsecondary enrollment a logit model was estimated using the matched data. The outcome indicator (a binary variable indicating postsecondary enrollment) was regressed on treatment status and the entire set of covariates used in the propensity score equation. The regression adjustment was used to adjust for any misspecification in the model (Imbens & Wooldridge, 2009). Results

show (see Table 22) that controlling for all other variables in the model, students participating in the IBT program have higher odds of being enrolled in some type of educational program after grade 10. The odds are two times higher for those in the treatment group (compared to comparison students) in sample I and five times higher for treatment group students in sample II.

Data on performance in the grade 10 standardized test and enrollment patterns after successful completion of grade 10 were examined for a subset of students in matched sample I (for whom these data were available). Figure 38 shows the distribution of grade 10 scores for the treatment and comparison groups in matched sample I. Grade 10 test scores were available for 90 percent of the matched sample. On average, treatment group students reported scoring one percentage point higher than the comparison group on the school exit exam in grade 10 (63 percent versus 62 percent).

Table 22

Odds ratio estimates of factors predicting postsecondary enrollment in the matched samples

	Analytic Sample I		Analytic Sample II	
	Odds Ratio	SE	Odds Ratio	SE
Treatment status	2.23**	0.59	5.25***	1.85
Female (Dummy)	0.25*	0.18	0.27	0.27
Hindu (Dummy)	1.68	0.73	0.88	0.51
Social Group – Dalit (Dummy)	0.62	0.36	1.72	1.57
Social Group – Adivasi (Dummy)	0.40*	0.18	0.29	0.20
Social Group – OBC (Dummy)	0.83	0.43	0.96	0.62
Household Size	0.94	0.11	1.12	0.16
Household head in agriculture (Dummy)	0.74	0.60	1.71	2.01
Household head in self-employment (Dummy)	0.97	0.83	3.22	4.02
Household head in salaried work (Dummy)	0.92	0.79	1.97	2.45

High income household (Dummy)	1.95	0.83	1.99	1.12
Medium income household (Dummy)	1.88*	0.58	1.14	0.51
Grade 7 performance (Z scores)	1.05	0.03	1.08	0.04
Distance to nearest town/city	1.04**	0.02	1.03	0.03
Distance to bus station	0.95*	0.03	1.18	0.11
Access to public transport (Dummy)	1.53	1.44	5.80	10.30
Village size	1.00	0.00	1.00	0.00
Female*Hindu	4.29*	2.97	3.53	3.44
Female*Dalit	2.52	3.55	-	-
Female*Adivasi	1.52	0.99	1.18	1.11
Female*OBC	1.03	0.85	0.58	0.61
Female*Grade 7 performance	0.99	0.03	0.96	0.04
Distance to town/city*Grade 7 performance	1.00	0.00	1.00	0.00
N	*314*		*208*	

Note. Sampling weights are used in the estimation to account for oversampling of female students in the comparison group.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

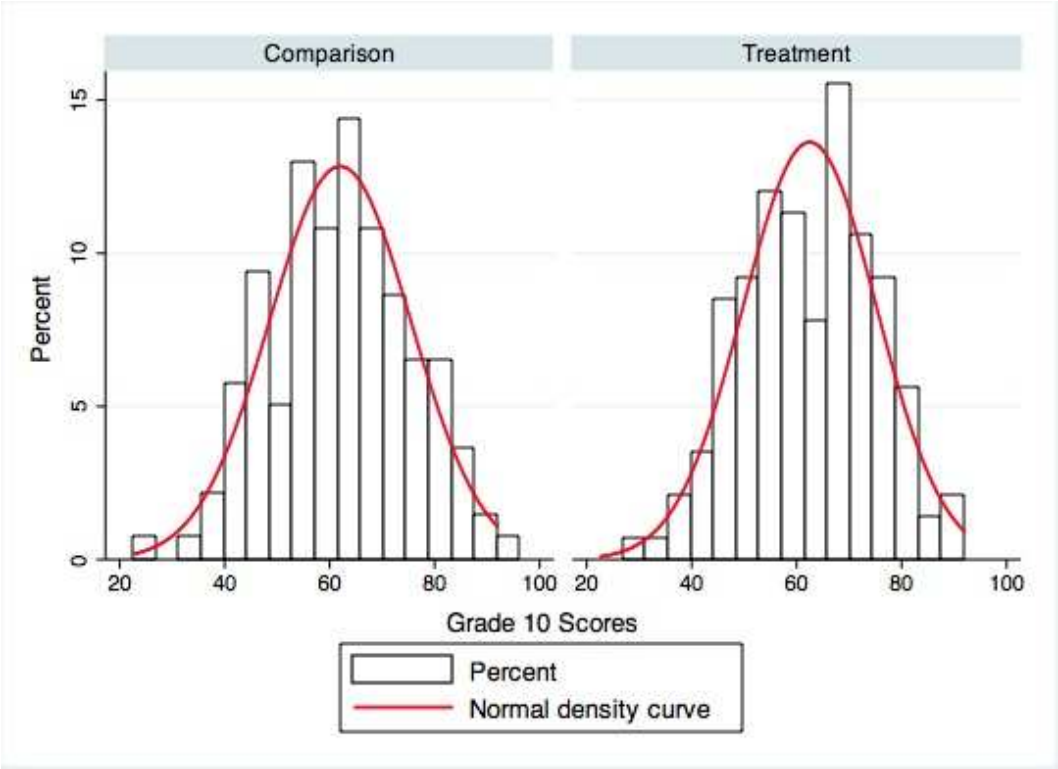

Figure 38. Distribution of scores on standardized test in grade 10 for treatment and comparison students in matched sample I. Scores on the grade 10 standardized test are available for 281 of the 314 students in the matched sample.

In terms of enrollment patterns after grade 10; about 48 percent students in matched sample I provided complete information on their current enrollment. For these students Figure 39 shows the proportion of students enrolled in various programs by treatment and control status. About 34 percent of the comparison group reported being enrolled in some type of vocational program. The corresponding proportion in the treatment group was 27 percent. The majority of students in the treatment group reported being enrolled in Junior college or at the higher secondary level and about 20 percent reported being enrolled in a professional degree program.

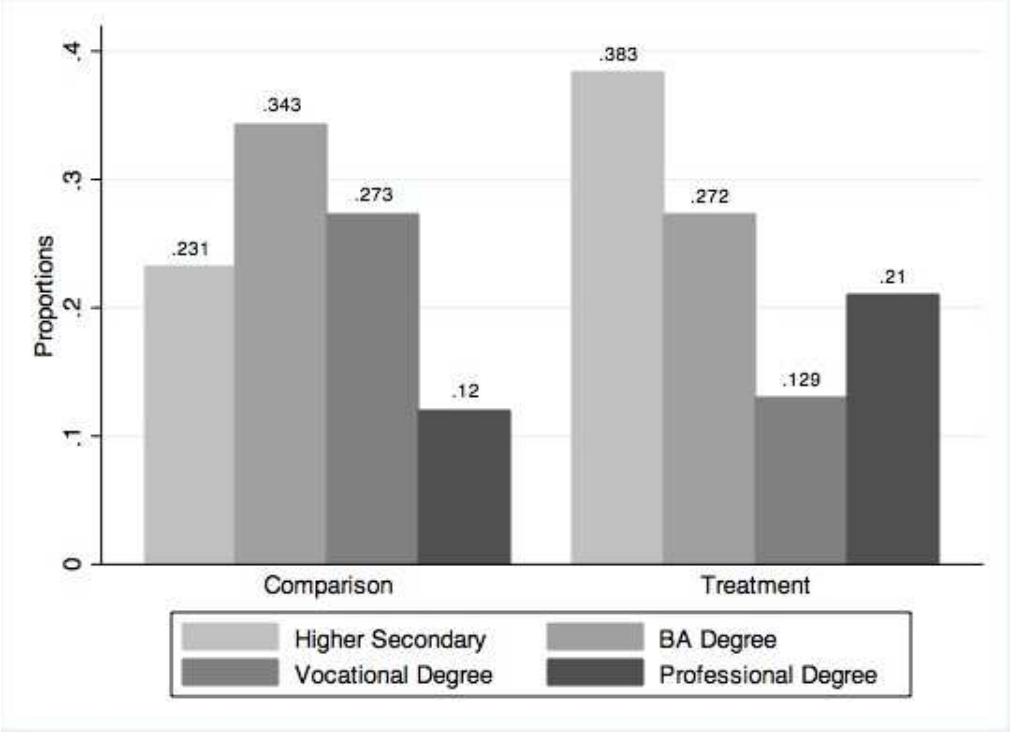

Figure 39. Proportion of students enrolled in various educational and training programs after grade 10. Data on program-wise enrollment is available for 150 of the 314 students in matched sample I.

6.4 Limitations

This study collected primary data to study the effect of secondary school-based vocational education and training on school completion and further education. Propensity score based techniques were used to create matched samples of treatment and comparison students. Missing data due to nonresponse and/or errors and limited common support between the treatment and comparison samples limit the generalizability of the findings from this study.

To address bias related to missing data, nonresponse weights were created using logistic regression (modeling the probability to respond and using the inverse of the predicted probabilities as weights). The probability of response was modeled as a function of demographic and location variables including district dummies. The treatment group showed slightly higher odds of postsecondary enrollment with the use of nonresponse weights.

Examining common support between the treatment and comparison samples led to discarding a few treatment cases in the matched sample. A comparison of the means for the treatment group in the matched and unmatched samples showed that removing these cases from the original treatment sample did not create any significant changes in mean values. Still, the magnitude of the effect found in this study can be considered a lower-bound estimate of the true effect of participating in IBT.

Chapter 7: Discussion and Conclusion

This chapter summarizes the findings from the preceding chapters (Chapters 4, 5 and 6) in the context of the current education and employment landscape in India. The implications of the findings in light of current technical and vocational education and training (TVET) policies in India are also discussed. The chapter concludes with directions for future research.

The working age group (15 to 64 years) comprises over 750 million people in India of which about 500 million are under 25 years of age. Over the next decade the working age group is projected to comprise nearly 66 percent of the country's population (Census of India, 2011). While an increase in the share of the working-age population can have a positive effect on GDP growth it also presents education, training and employment challenges (Mehrotra, Gandhi & Sahoo, 2013).

While there have been steady improvements in access to primary schooling in India (Planning Commission, 2008), learning levels at the primary level are abysmally low (Annual Status of Education Report, 2013) and the country faces severe challenges in transitions from the lower to upper secondary levels of school (Planning Commission, 2012). Research has shown that low educational participation rates along with low learning outcomes have serious implications for individual employment outcomes and for the productivity of the economy as a whole (Hanushek, 2008; Hanushek and Zhang, 2008).

In 2009-10, nearly 30 percent of those in the labor force in India had no formal schooling. Only 17 percent had completed high school. (See Table 23 for a gender-wise

breakdown of labor force participation by level of completed schooling in 2009-10). Those with some type of formal or informal vocational training comprised about 10 percent of labor force participants.[55] Educational and labor force participation by gender groups shows that women participate at significantly lower rates than men and have poorer outcomes (Klasen & Pieters, 2013). These statistics indicate serious skill shortages and gender-based inequities in the Indian economy.

Table 23

Education level of labor force participants in 2009-10 (Weighted percentages)

Level of education	All	Males	Females
Not literate	29.0	16.8	12.1
Literate without formal schooling	0.3	0.2	0.1
Non-formal education	0.1	0.1	0.0
Less than 5 years of schooling	9.4	7.2	2.2
5 years of formal schooling	14.5	11.5	3.0
7-8 years of formal schooling	17.6	14.9	2.7
10 years of formal schooling	12.2	10.8	1.4
12 years of formal schooling	6.7	5.9	0.8
Diploma or Certificate	1.4	1.1	0.3
Bachelor's degree	6.6	5.5	1.1
More than a bachelor's degree	2.2	1.7	0.5
N	*365,153,849*		

Source: Employment and Unemployment Survey of India, 2009-10.
Note. The table includes all those between 15-64 years of age in the labor force. Labor force participation is defined as per the "principal activity status" i.e. if an individual is employed for 180 or more days during the reference year.

In response to these skill shortages several reform efforts were undertaken to improve educational and employment outcomes (Planning Commission, 2008). These included the Right to Education Act, 2009 (focusing on elementary school completion), "vocationalisation" of secondary education (focusing on expansion of secondary-level

[55] See Appendix A for TVET participation rates among 15-59 year olds in 2009-10. See Mehrotra et al. (2013) for estimates of vocationally trained individuals by sector/industry.

TVET), the *Rashtriya Madhyamik Shiksha Abhiyan* (focusing on improvements at the secondary level), expansions in higher education, and the National Skill Development Policy, 2009 (addressing skill shortages). The most ambitious of these policies, the National Skill Development Policy, aims to provide skill training to 500 million Indians over the next decade by expanding and improving access to training services, upgrading the quality of existing services, improving female participation in TVET, and developing innovative models for delivering TVET (Visaria, 2013).

Although there is a growing body of literature on the effectiveness of TVET policies and programs (Adams, 2007), the TVET sector in India has been significantly under-researched. Findings from other countries show that TVET programs have heterogeneous effects and their success is closely linked to the specific objectives of the programs, their design and delivery and the macroeconomic context in which they operate (Gill, Fluitman & Dar, 2000). Because programs and policies have different effects in different contexts, a one-size-fits-all approach to designing and implementing TVET policies is discouraged (Adams et al., 2013; Gill et al., 2000). Research strongly advocates adopting a results-based policymaking approach where programs are pilot-tested, monitored and evaluated before being implemented on a large scale (Adams et al., 2013).

In an attempt to build the empirical research base on TVET in India and to enable the policy dialogue on addressing current educational, skill and employment challenges faced in the country, this dissertation asked three fundamental questions; what are the predictors of TVET participation in India, what are the consequences of participating in secondary school based TVET, and what are the consequences of participating in

postsecondary TVET? Preliminary findings suggest that participation patterns vary by type of TVET (formal and informal) and individual and household characteristics are significantly associated with TVET participation but act differently depending upon type of TVET. The consequences of TVET participation at the postsecondary level, as measured in terms of economic returns, and at the secondary level, measured in terms of postsecondary enrollment, are positive and significant.

7.1 Predictors of participation in TVET

Two rounds of the Employment and Unemployment Survey of India (2004-05 and 2009-10) were used to empirically examine TVET participation patterns for males and females. It was hypothesized that TVET participation is predicted by several individual, household and socio-cultural factors that are influenced by the individual's micro and macro-context. Due to data limitations the influence of only some of these factors on TVET participation was examined.

Preliminary analyses showed that among 15-29 year olds, TVET participation rates in 2009-10 were slightly lower than those in 2004-05 (7 percent versus 11 percent). These differences were largely driven by lower participation rates in informal TVET in 2009-10. Participation rates by gender indicate that females participated in TVET at lower rates than males, and the gender differences in participation were wider in the case of informal TVET than formal TVET. (See Figures 40 and 41).

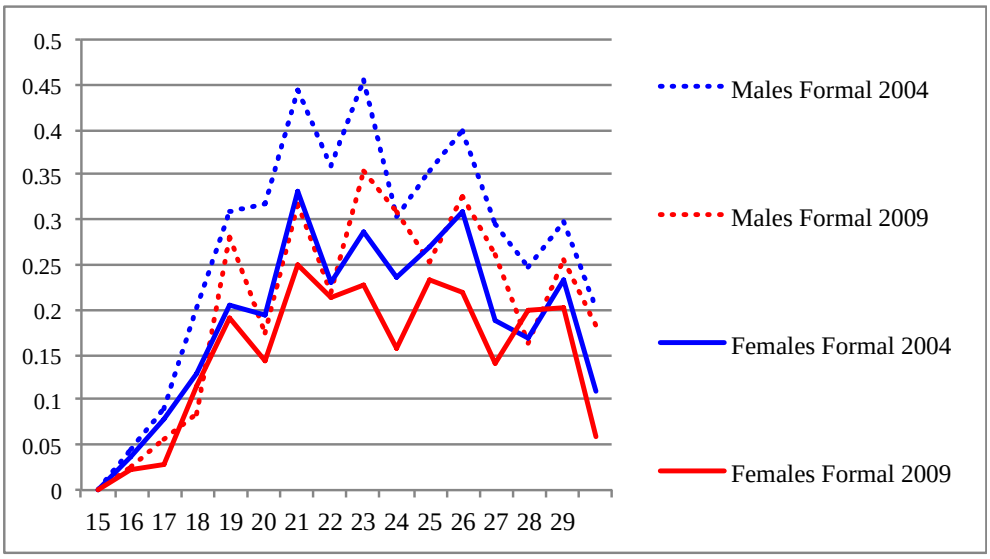

Figure 40. Participation in formal TVET among males and females between 15-29 years of age in 2004-05 and 2009-10.
Source: Employment and Unemployment Survey of India, 2004-05 & 2009-10.

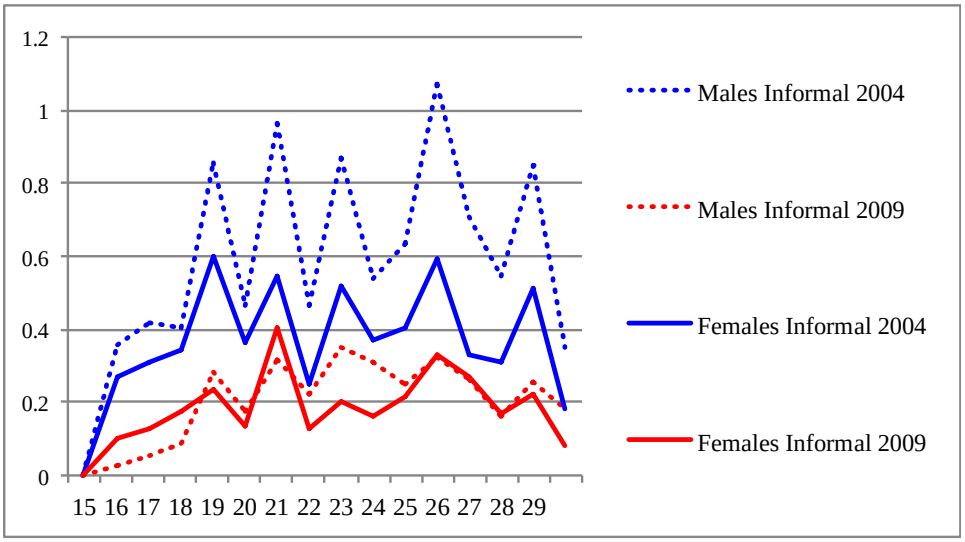

Figure 41. Participation in informal TVET among males and females between 15-29 years of age in 2004-05 and 2009-10.
Source: Employment and Unemployment Survey of India, 2004-05 & 2009-10.

These changes in participation patterns can be attributed to higher educational participation rates during the same period. In 2009-10, participation in all levels of general education from elementary to secondary and tertiary showed significant increases

over those in 2004-05. Further, labor force participation rates between 2004-05 and 2009-10 also showed significant increases in the proportion of males and females engaged in various types of wage work[56] and slight increases in the proportion of salaried workers.

In terms of the composition of formal and informal TVET groups by gender and location, urban males made up the majority of formal TVET participants. There were no remarkable changes in composition between 2004-05 and 2009-10 for formal TVET. In the case of informal TVET however, there was a shift in rural female and urban male participation. Compared to 2004-05, the number of urban males as a proportion of informal TVET participants increased significantly while the proportion of rural females decreased. While these shifts could be explained on the basis of changes in educational and labor force participation rates outlined above, they could also be related to changes in the way informal TVET participation data has been collected in 2004-05 and 2009-10. Two additional categories – "self-learning" and "on-the-job training" – were added to the definition of informal TVET in 2009-10 that were not included when gathering informal TVET participation data in 2004-05.

The descriptive analysis also showed clear differences in the average level of education among those who participated in formal and informal TVET. Formal TVET participants, on average, had a high school degree (12 years of formal schooling) while those participating in informal TVET averaged about 7 years of schooling. Formal TVET programs in India can be accessed by school dropouts, and at the secondary and postsecondary stages of education. But the relationship between education levels and

[56] The Mahatma Gandhi National Rural Guarantee Scheme, a public-works employment program, was rolled out between 2006 and 2008 and guaranteed rural households up to 100 days of employment that could be taken up any time during the year (cite).

formal TVET participation indicates that, for the most part, formal TVET programs offering basic certification programs that require fewer years of formal schooling for entry are not as popular as programs of longer duration offering training in technical fields and requiring at least 10 to 12 years of formal schooling. The former programs prepare youth for entry-level employment and semi-skilled jobs with a smaller proportion accessing advanced TVET options (Adams, 2007).

The role of education as a strong predictor of formal TVET participation was supported in the multivariate analysis. Completion of each level of schooling between primary to upper secondary significantly increased the likelihood of participation in formal TVET controlling for all other individual, household and contextual factors. Individuals with 12 years of schooling were nearly 5 times more likely to enroll in a formal TVET program than those who did not have a high school degree. This relationship was consistent across both time periods.

The reverse relationship was observed between education and participation in informal TVET. The odds of participating in TVET went down significantly with each level of education completed. Thus, while completing primary schooling increased the odds of participating in informal TVET by 1.5 times, those who completed secondary school had a 35 percent lower likelihood of enrolling in informal TVET than those who did not complete secondary school.

There are a couple of implications of this relationship between education and TVET participation in India. For one, it underscores the preferences of parents and students for higher levels of education and training, and for certain types of education and training over others. This is evidenced by the fact that in 2004-05, about 65 percent of all

male formal TVET participants sought training in engineering-related or technical fields. In 2009-10 this proportion rose to 75 percent. The corresponding proportion of women participating in formal TVET in technical fields was 33 percent and 50 percent in 2004-05 and 2009-10, respectively. While there are formal TVET participants in non-technical fields of study, they constitute a smaller proportion of formal TVET participants and have lower levels of educational attainment than formal TVET participants in technical fields.

The findings also imply that the benefits associated with formal TVET at the postsecondary level are perceived to be higher than the costs thus encouraging students from households above a certain income threshold to participate in formal TVET.

The predictive models estimated in Chapter 4 were not examined by field of study because of sample size limitations. Examining the variations in predictive patterns by field of study would further elucidate our understanding of the determinants of TVET in India.

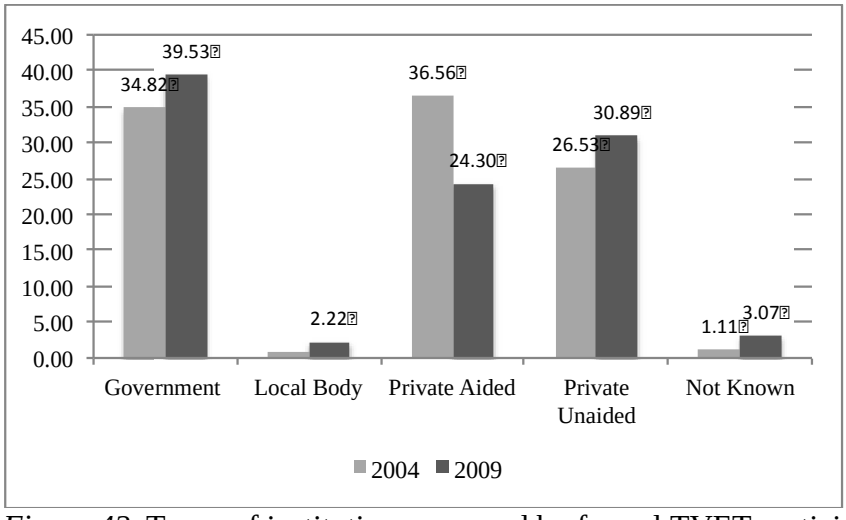

Figure 42. Types of institutions accessed by formal TVET participants in engineering-related fields in 2004-05 and 2009-10.
Source: Employment and Unemployment Survey of India, 2004-05 & 2009-10.

In addition to the role of education, the predictive models hypothesized that TVET participation was related to the characteristics of the household. Amongst the household-level characteristics examined, results indicate that household income (measured using household consumption expenditure as proxy for household income) significantly increases the odds of participation in formal TVET. This relationship was found to be stronger in the case of males than females and slightly larger in magnitude in 2004-05 than 2009-10.

Formal TVET programs can be accessed at public, private aided or private unaided institutions. The share of formal TVET participants accessing TVET programs at private unaided institutions has increased over time. (Figure 42 below shows the percent of formal TVET participants in engineering or technical courses enrolled at various types of institutions). While the cost of publicly provided formal TVET can be as low as Rupees 20 per month ($0.35 per month), the actual costs of offering TVET programs is much higher (Tilotia, n.d.). Formal TVET programs at private aided and private unaided institutions can cost anywhere upwards of Rupees 5,000 (about $80). Research on training participation in the Indian context has found that credit constraints are a significant barrier to participation and completion (Maitra & Mani, 2013). Thus, among households that overcome these budget constraints, participation in TVET is higher. Households that face significant financial constraints however, show lower levels of educational attainment and by extension, a lower likelihood of participating in formal TVET programs.

These findings lend support to current policy recommendations for subsidizing the costs of participating in TVET programs (Planning Commission, 2008; 2012), especially for those from disadvantaged and minority groups. To this end, the National Skill Development Corporation has recently launched the 'Standard Training Assessment Reward' (STAR) scheme that offers monetary incentives to youth and young adults for participating in TVET. The eligibility criteria include high school completion and enrolling in a TVET program pre-approved by the scheme.

While schemes such as STAR attempt to encourage TVET participation among high school completers in a way that meets the needs of the economy, there is also need to address the constraints and barriers that prevent access to and completion of secondary schooling. The predictive models showed that amongst the other household characteristics significantly related to TVET participation, the occupation of the household was a significant factor. Self-employed households significantly raised the likelihood of informal TVET participation among females. These findings are in keeping with what is known about female employment in non-agricultural work. Data from the NSSO surveys indicate that females in non-agricultural employment are usually employed in home-based work that is sub-contracted to them and of low-productivity (Planning Commission, 2012). As mentioned before, low educational attainment among this group implies that they are less likely to enroll in formal TVET programs, more likely to be employed in the informal sector, and therefore more likely to have lower employment outcomes.

The results for context characteristics showed that once individual, household and demographic variables are accounted for, context-level variables like the supply of TVET

institutions in the district do not explain any variation in TVET participation. The unemployment rate in the district showed very small but consistently positive effects on the odds of participation in formal TVET. This finding is supported by prior research in other contexts (Walstab, 2008).

The limited role of context-level variables was a surprising finding given the regional variation in the supply of TVET in the country (Ministry of Labor and Employment, 2011) and the variation in economic growth and labor force participation across different regions. One explanation for no significant association between TVET capacity and participation patterns is that the data used to measure the spread of TVET institutions could be dated or incomplete. The data were sourced from the website of the Directorate General of Education and Training (the apex body overseeing TVET in the country) and it is unclear when the data were last updated. It is also possible that alternate indicators of the macro-economic context could have better served the predictive models. Sector-wise job growth across districts/states, and responsiveness of state institutions in expanding educational and training opportunities might better explain the regional variation in TVET participation.

Educational participation in India shows significant variation across various demographic dimensions. Participation in formal and informal TVET was no different. The predictive models showed that after accounting for educational attainment, household characteristics and demographic controls, females were significantly less likely to participate in TVET. Between 2004-05 and 2009-10, the odds of females enrolling in formal TVET programs further reduced. In 2004-05, women were 15 percent less likely to enroll in formal TVET as compared to men, but in 2009-10 they were 20

percent less likely to participate as compared to males in the same age group. It has been shown that female educational and labor force participation in India goes down as household income increases (Klasen & Pieters, 2013). There is evidence that household consumption expenditure (assumed to be an indicator of household income) increased between 2004-05 and 2009-10. But these findings also lend empirical support to concerns regarding the gender imbalance in educational and training participation noted in policy discussions. Recommendations to improve female participation in TVET by improving access and budget constraints do not address the full scope of the problem. Research has shown that the effect of education on female labor force participation follows a U-shaped pattern. This indicates that labor force participation decisions for females in the middle of the education distribution are affected by factors other than their level of education. Klasen & Pieters (2013) find that women's own preferences for white-collar jobs and stigmatizing women's work outside of the public sector are related to low labor force participation among educated women. Thus, efforts to improve female participation rates in TVET, and in the labor market, require a cultural shift in attitudes regarding women's work.

The empirical results in Chapter 4 also found that in 2009-10, Adivasis had significantly higher probabilities of participation in informal TVET. This finding is related to lower educational outcomes for this group relative to other groups. Current policies focus on improving participation of these groups in skill development by removing credit constraints and improving access to skill development services. More emphasis should be placed on providing these disadvantaged and minority groups with basic literacy and second-chance education programs. Research has found that a good

foundation of formal education is an important variable influencing later skills acquisition and improves employment outcomes in the formal and informal sector (Adams, 2013).

7.2 Returns to postsecondary TVET

Empirical evidence on the benefits of TVET for the individual has been practically nonexistent in the Indian context. Previous research examining the economic returns to education and training has narrowly focused on general education and a small subset of technical education programs. Results from these studies have limited generalizability because of sample and methodological constraints.[57] This information gap has been addressed in Chapter 5 generating empirical evidence on the returns to TVET in India. Using nationally representative data from the first round of the Indian Human Development Survey (2004-05) the economic returns to TVET have been estimated for individuals across all income-generating activities and using methods that control for selection and endogeneity bias. Results indicate that the economic returns to TVET participation after grade 10 are significantly higher than the returns to a Bachelor's degree and are comparable to the returns to a Master's degree. Controlling for individual ability and various demographic characteristics, TVET participants earn nearly 19 percent more in annual wages than individuals not participating in TVET. The evidence

[57] Previous research on returns to education (Duraisamy, 2000 & Agrawal, 2011) has limited the sample to regular wage earners excluding those self-employed or those in casual work from the analytic sample. Further, these studies have not controlled for endogeneity bias (from including schooling on the right hand side) and omitted variable bias (from not including a measure of ability) in estimating returns.

also suggests that the returns estimates reported here are lower bound estimates, and the actual wages associated with postsecondary TVET participation could be higher.

Human capital theories suggest that when the marginal benefits of education and training exceed the marginal costs, individuals are more likely to participate in education and training (Becker, 1967). In the case of TVET in India, information regarding the actual benefits of TVET participation has until now been available. Given the low rates of participation in formal TVET in the country it can be inferred that youth and young adults perceive lower benefits accruing from TVET programs than general education programs. Surveys of youth have also shown that formal TVET programs occupy lower status as compared to general education programs contributing to the perception that these programs are associated with lower marginal benefits (Aggarwal et al., 2012).

Jensen's (2010) experiment found that in the absence of imperfect information regarding costs and benefits of schooling, students were more likely to make educational and training decisions in keeping with Becker's human capital theory. Thus when students were informed about the returns associated with higher levels of schooling, the proportion attending and completing secondary school significantly increased.

The results presented here have implications for meeting current policy targets for increasing TVET enrollments. Making public the findings on positive significant returns to TVET as compared to other programs could go a long way in reversing the low status perception of TVET in India. Future research, similar to Jensen's (2010) work in the Dominican Republic, could perhaps examine if information on TVET wages encourages participation in TVET programs. Research is also required to identify effective modalities

through which information regarding the benefits of TVET can be communicated to those who are most likely to benefit from these programs.

The findings presented in chapter 5 also show that English language fluency is associated with high positive returns after controlling for ability and educational achievement. These results replicate the findings in Azam et al. (2013) for the TVET population in India and find evidence that the labor market rewards English language skills. Current policies focusing on skill development in the country have underscored the need for "soft skills", including English language ability along with computer literacy, critical thinking and time management skills, in addition to broad educational and occupational requirements (Planning Commission, 2008). These findings lend support to current policies by providing empirical evidence that English language skills are associated with higher wages.[58]

The wage models estimated in Chapter 5 are limited in that the variation in returns by gender and urban-rural status could not be examined. Future research must address these limitations. Further, the returns to TVET were estimated for all those with 10 or more years of education. But as noted earlier, formal TVET programs are offered in various technical and non-technical fields requiring different levels of prior schooling. Empirical evidence on the returns associated with various programs having different entry requirements will be informative.

Finally, research examining the consequences of participating in TVET must expand it's scope to examine a more diverse set of indicators beyond wages. The

[58] A more nuanced discussion of the returns associated with English language skills and the complementarity between English ability and general education can be found in Azam, Chin & Prakash (2013).

indicators could include those associated with employment like duration of job search, or number of unemployment spells. The effect of TVET on health outcomes could also be studied.

7.3 Effect of secondary-level TVET

Vocational education at the secondary level has been offered in a small subset of secondary schools since 1988. Recent policy revisions have emphasized encouraging secondary school students to participate in vocational education in larger numbers with the dual objectives of preparing youth for the labor market and postsecondary vocational education while also improving retention and secondary school completion rates. States, in partnership civil society organizations, have undertaken innovative programs to pilot various models for delivering TVET at the secondary stage of schooling. While the results of these initiatives are not yet available, the literature on diversifying secondary education with a small number of vocational courses does not show positive results for employment (Lauglo & Maclean, 2005; Psacharopoulos, 1987). Studies have found that the payoff in terms of employment is substantial only when vocational courses form a major share of the curriculum and are closely linked to labor market needs. There is however evidence that shows that secondary school TVET is positively related to retention and high school graduation (Bishop & Mane, 2005).

An innovative secondary school-based TVET program in rural Maharashtra was evaluated to estimate the effect of participation on school completion, postsecondary enrollment and short-term employment outcomes. Propensity scores were used to create matched treatment and comparison groups. The results showed that participating in

TVET in grades 8, 9 and 10 significantly increased the probability of school completion and postsecondary enrollment. Treatment group students were twice more likely to enroll in an education program after completing grade 10 than students in the comparison group. The effect of secondary school TVET on short-term employment outcomes could not be tested.

These findings, although limited to one program in a few districts of rural Maharashtra, offer some preliminary support to using secondary school TVET to meet universal secondary education goals as articulated in the Twelfth Plan and the RMSA (Planning Commission, 2012). More rigorous evaluations of pilot programs are required to see if these programs work across urban-rural contexts in different states, and in different types of secondary schools. The TVET program evaluated here is offered in government-aided secondary schools that have a higher degree of autonomy and more resources than secondary schools that are managed and funded by the government. Further, the vocational curriculum offered as part of this program while applicable to the rural context, would not be relevant to students attending secondary schools in semi-urban and urban areas. The positive findings observed here therefore cannot be disassociated from the schools and context in which this program operates. The significant additional costs associated with offering vocational courses at the secondary school level also make it important to examine these programs from a cost-benefit perspective.

Future research on secondary school TVET should also examine the extent to which these programs address or exacerbate gaps in educational attainment between subgroups. As mentioned above, the program evaluated here is offered in government-

aided secondary schools. These schools are partially funded by the government and often have a higher fee structure than public schools. As a result, households belonging to disadvantaged and vulnerable groups cannot access these schools and programs. The lack of infrastructure, high teacher absenteeism and other problems of the public school system imply that students from disadvantaged groups attending the public school system are more likely to drop out because of lack of interest and financial constraints than students attending government-aided or private schools. Without basic education these students are less likely to participate in formal TVET and end up in low wage employment.

To address the skill development needs of students from disadvantaged backgrounds who face significant challenges accessing quality formal schooling, the government has set up a system of short-term training programs called 'modular employable skills'. The courses offered through this program are designed for participants with primary level education and can be delivered in a flexible manner to accommodate the needs of the learner. Female students and students from disadvantaged subgroups are eligible for subsidized fees.

7.4 Conclusion

Technical and vocational education and training has gained importance in India over the last decade. In response to skill shortages in the labor market, unemployment, and low educational outcomes amongst youth, and demographic shifts policymakers have focused on expanding TVET provision in the country. Vocational programs in India have traditionally been perceived as low status and the TVET sector has been under-utilized.

In light of current expansion plans and the lack of empirical research on the TVET sector in India this dissertation focused on addressing fundamental questions to enable the policy dialogue on TVET in India.

First, the dissertation focused on examining the predictors of participation in TVET. In order to improve participation in TVET it is important to understand the mechanisms underlying participation decisions. Information on the predictors of TVET participation also enables policymakers to design and target appropriate strategies for policy-relevant subgroups. Results from the predictive models showed that formal and informal TVET programs have slightly different underlying participation patterns. While participation in both types of TVET programs is significantly associated with individual and household characteristics, the magnitude and direction of these relationships differ by type of TVET. Notably, educational attainment and household income is associated with higher odds of formal TVET participation and lower odds of participation in informal TVET. As a result, the proportion of disadvantaged and minority subgroups among informal TVET participants is disproportionately higher than other groups. The data on labor force participation show that these marginalized groups are also disproportionately represented in the informal sector and have low-paying jobs.

Second, in order to examine the consequences of TVET participation on the individual, the returns to postsecondary TVET were estimated. The analysis controlled for selection bias and results showed that those who participated in TVET after completing at least 10 grades of education had yearly wages that were 18 percent higher than those who did not participate in TVET. The returns associated with formal TVET at

the higher secondary or postsecondary level were higher than the returns to a Bachelor's degree, and comparable to those associated with a Master's degree.

Third, a secondary school-based TVET program was evaluated to examine the effect of a diversified curriculum on school completion and postsecondary enrollment. The findings suggest that secondary school TVET can have positive effects on school completion and postsecondary enrollment. Results also show that the majority of students participating in TVET in the secondary schools included in the study enrolled in general education programs at the postsecondary level.

These findings indicate that a lack of high school education and inequitable access to quality formal schooling are significant obstacles in the way of achieving the skill development objectives outlined in the National Skill Development Policy. Further, efforts to encourage participation in TVET while focusing on issues around access, quality, and equity, also need to focus on changing the low status perception of TVET and making TVET more attractive to users. The role of information will be critical for raising awareness about current and future programs while also correcting misplaced perceptions regarding TVET options.

It is heartening to note that current policies and programs designed for the TVET sector align with several findings from the research reported here. The scale of the targets envisioned by current programs (for example, training 500 million individuals by 2020, setting up new training 3000 institutes through public-private partnerships) require extensive resources, a high degree of coordination, and implementation expertise that has been lacking in previous social sector programs.

Future research on TVET in India should focus on refining the preliminary evidence on determinants and benefits of TVET reported here. The focus should also be on closely monitoring current reform efforts to identify course corrections in a timely manner and identify successful strategies that can be replicated in other locations in the country.

Appendix A

Sampling Strategy (Employment and Unemployment Survey)

The sampling design involves a multi-stage stratified design (NSSO, 2006; 2011). The list of all 2001 Census villages formed the sampling frame for the rural areas while the Urban Frame Survey and individual towns constituted the sampling frame for the urban sector. Next, within each district of the country strata were created representing the urban and rural areas within each district. Rural strata were further divided into substrata – the first substrata included those villages where the proportion of child workers exceeded the average proportion for the state and the second substrata included the remaining villages. Probability proportion to size with replacement was used to select rural primary sampling units (PSUs) from each rural stratum and substratum and simple random sampling without replacement was applied for selecting urban PSUs. All households in these rural and urban PSUs were stratified into three second stage strata (SSS). The sample households from each SSS were selected by simple random sampling without replacement.

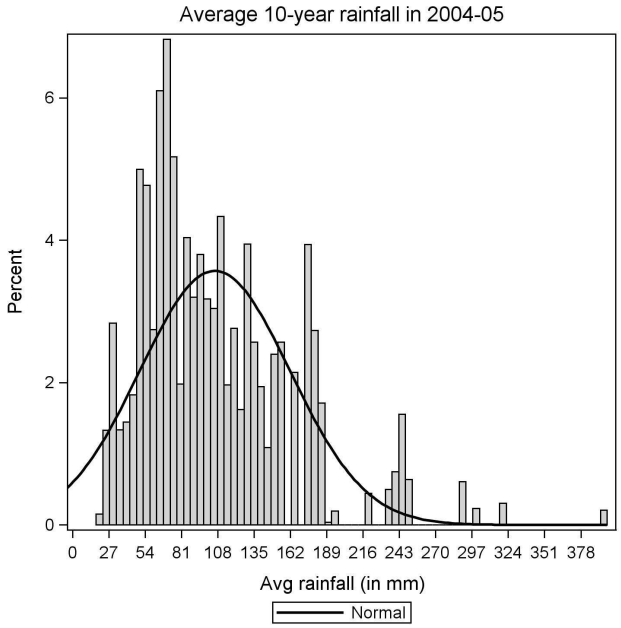

Figure A.1. Distribution of average rainfall at the district-level (2004-05)

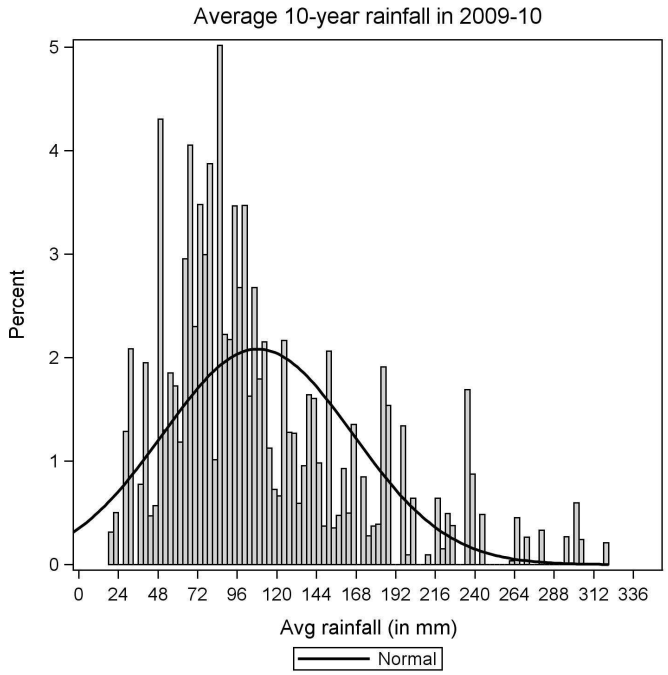

Figure A.2. Distribution of average rainfall at the district-level (2009-10)

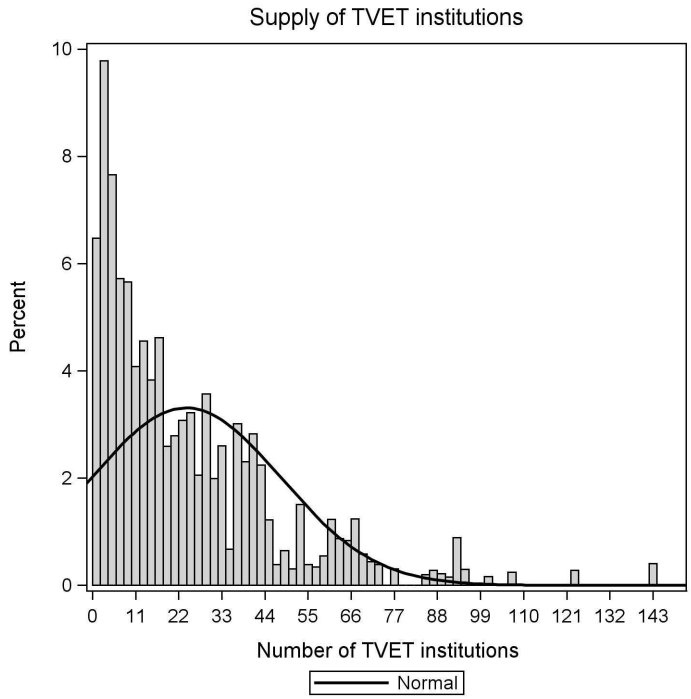

Figure A.3. Distribution of TVET institutions at the district-level (2004-05)

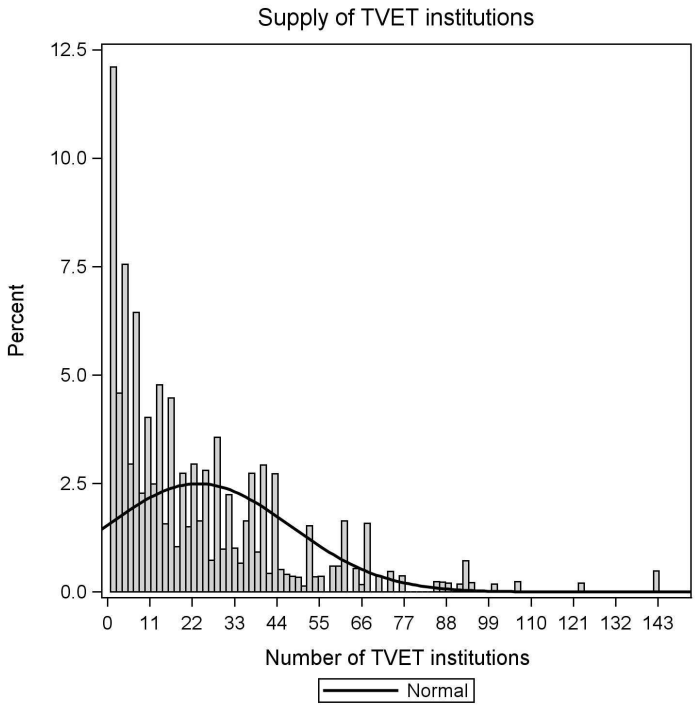

Figure A.4. Distribution of TVET institutions at the district-level (2009-10)

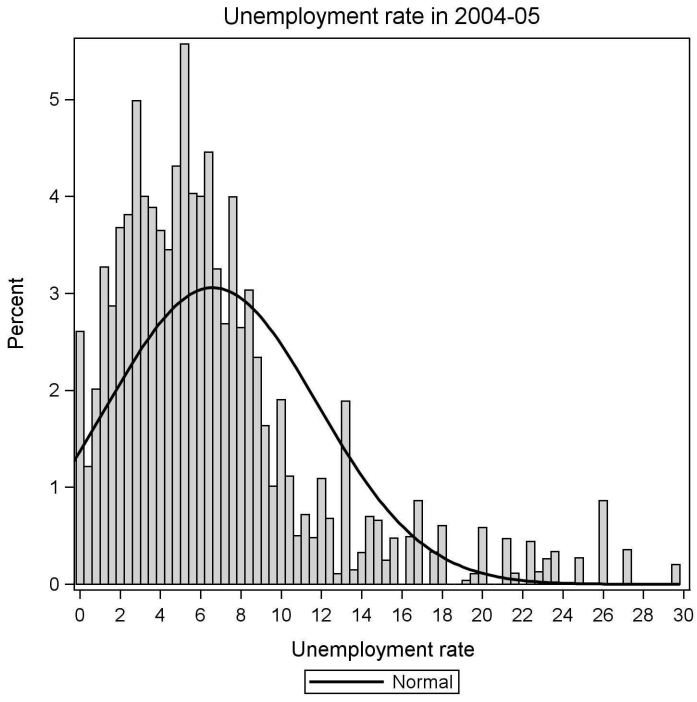

Figure A.5. Unemployment rate at the district-level (2004-05)

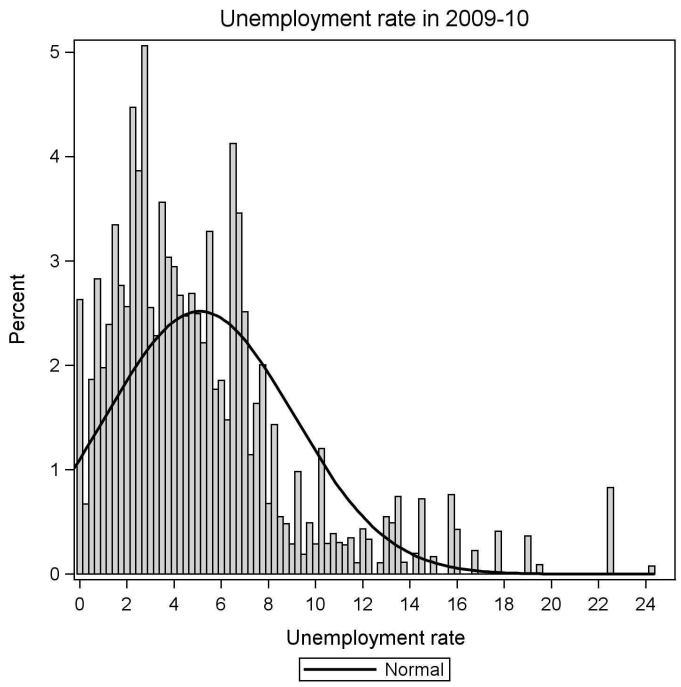

Figure A.6. Unemployment rate at the district-level (2009-10)

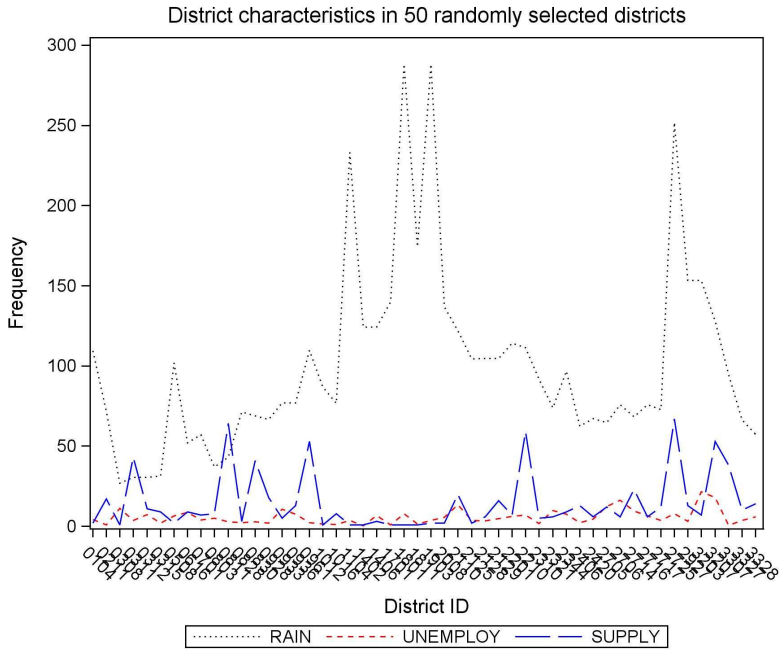

Figure A.7. District characteristics across 50 randomly selected districts

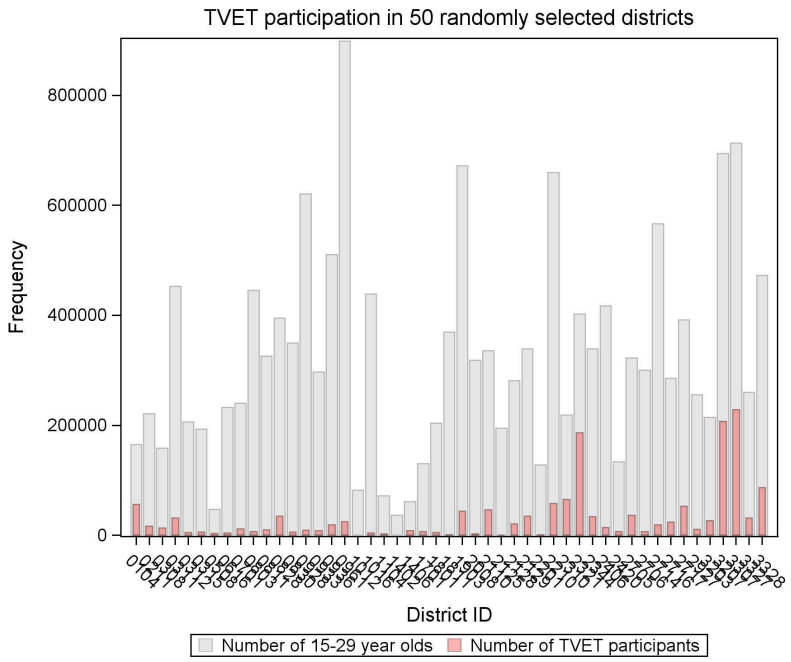

Figure A.8. Average TVET participation across 50 randomly selected districts

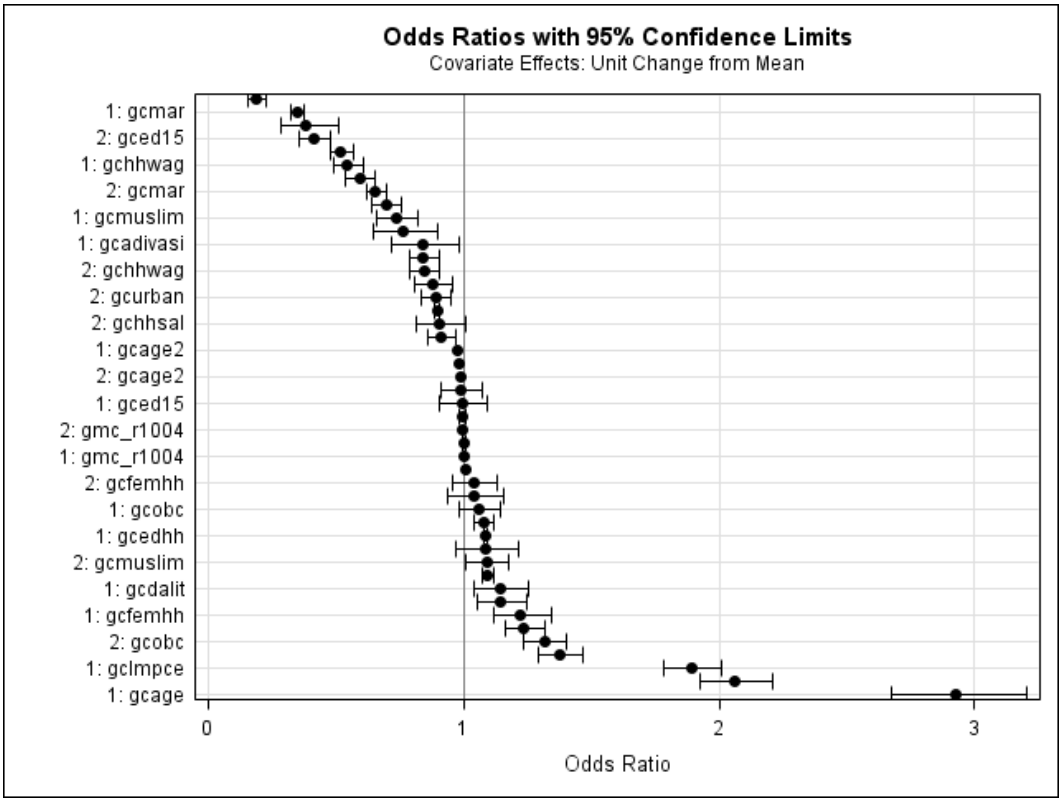

Figure A.9. Distribution of odds ratio estimates predicting participation in formal (1) and informal (2) TVET in 2004-05.

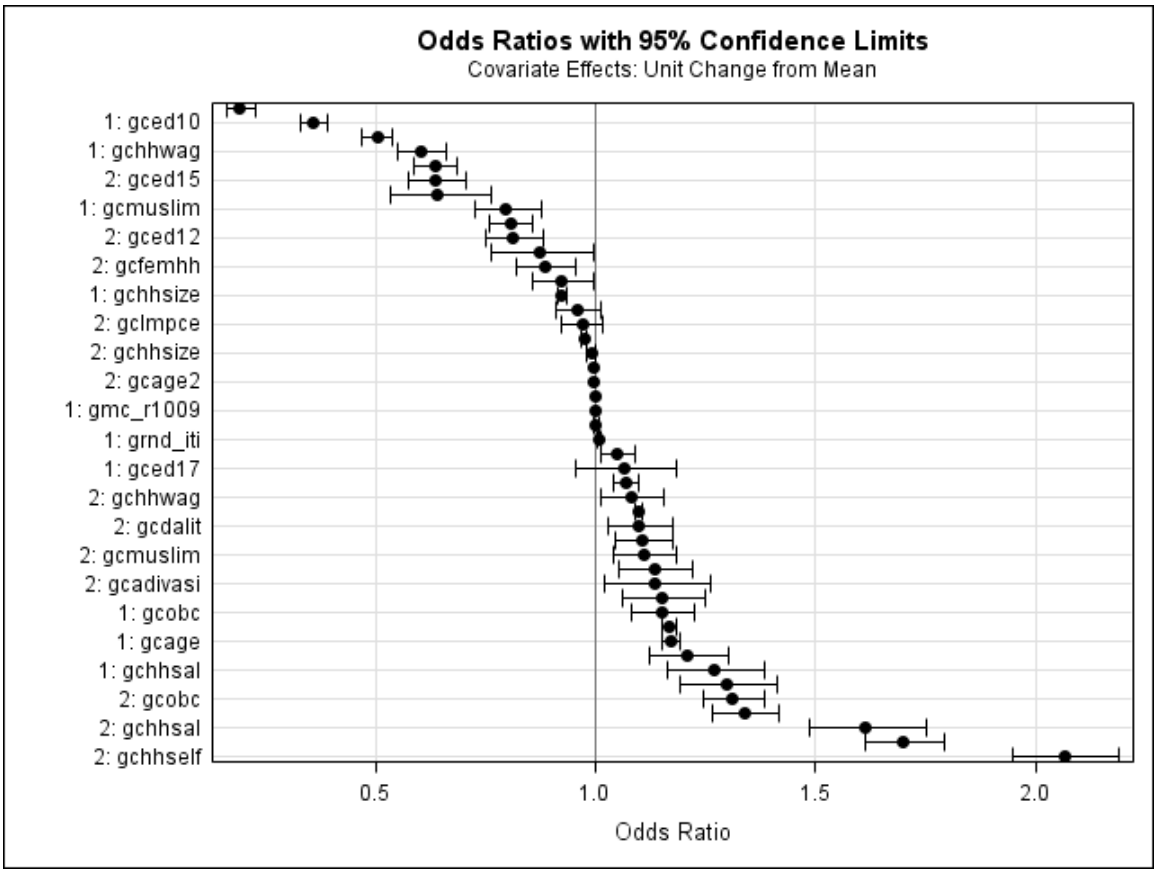

Figure A.10. Distribution of odds ratio estimates predicting participation in formal (1) and informal (2) TVET in 2009-10.

Table A.1
Weighted descriptive statistics (2009-10; 15-59 year olds)

Variables	Mean	SE (Mean)
Individual Characteristics		
TVET Participation (Binary)	0.078	0.002
TVET Participation (Multinomial)	0.128	0.003
Age	32.697	0.042
Age Squared	1213.490	2.969
Female	0.489	0.001
Urban	0.292	0.003
Years of Schooling	6.011	0.032
Monthly Household Expenditure	5721.767	44.720
Marital Status	0.695	0.002
Social Group - Dalit	0.197	0.003
Social Group - Adivasi	0.085	0.003
Social Group - OBC	0.411	0.004

Religious Group - Muslim	0.121	0.003
Head of the household's Education	5.338	0.040
Female-headed Household	0.083	0.002
N		
District Characteristics		
District TVET Capacity*	19.896	21.508
Average District Rainfall**	103.258	55.302
District Unemployment Rate	4.601	3.762
N	612	

Source: Employment and Unemployment Survey of India (2009-10)

Note. * N=508; ** N=559

Table A.2

Weighted descriptive statistics by gender and urbanicity (2009-10; 15-59 year olds)

	Rural Males		Rural Females	
	Mean	SE	Mean	SE
TVET Participation (Binary)	0.084	0.003	0.038	0.002
TVET Participation (Multinomial)	0.148	0.006	0.066	0.003
Age	32.556	0.072	32.870	0.067
Age Squared	1209.973	4.913	1222.204	4.755
Years of Schooling	6.092	0.042	3.884	0.036
Monthly Household Expenditure	4796.004	36.094	4750.837	35.334
Marital Status	0.659	0.003	0.762	0.003
Social Group - Dalit	0.219	0.005	0.218	0.005
Social Group - Adivasi	0.108	0.004	0.110	0.004
Social Group - OBC	0.416	0.006	0.424	0.006
Religious Group - Muslim	0.111	0.004	0.111	0.004
Head of the household's Education	4.288	0.049	4.166	0.043
Female-headed Household	0.053	0.002	0.106	0.003
N	*86653*		*85051*	

Source: Employment and Unemployment Survey of India (2009-10)

Table A.3

Weighted descriptive statistics by gender and urbanicity (2009-10; 15-59 year olds)

Variables	Urban Males		Urban Females	
	Mean	SE	Mean	SE
TVET Participation (Binary)	0.161	0.004	0.072	0.003
TVET Participation (Multinomial)	0.249	0.007	0.104	0.005
Age	32.588	0.080	32.747	0.075

Age Squared	1204.930	5.610	1210.042	5.296
Years of Schooling	9.112	0.059	7.737	0.063
Monthly Household Expenditure	7913.810	122.275	8129.226	121.907
Marital Status	0.615	0.004	0.708	0.004
Social Group - Dalit	0.146	0.005	0.143	0.005
Social Group - Adivasi	0.028	0.002	0.029	0.002
Social Group - OBC	0.388	0.007	0.389	0.007
Religious Group - Muslim	0.146	0.006	0.147	0.006
Head of the household's Education	8.030	0.075	8.021	0.074
Female-headed Household	0.063	0.002	0.122	0.004
N	59472		56103	

Source: Employment and Unemployment Survey of India (2009-10)

Table A.4

Odds ratio estimates of factors predicting participation in any TVET among 15-59 year olds, by gender (2009-10)

	Full Sample	Males	Females
Demographic Controls:			
Age	1.15***	1.17***	1.14***
Age Squared	1.00***	1.00***	1.00***
Female (Dummy; Ref: Male)	0.43***		
Urban (Dummy; Ref: Rural)	0.97	0.97	0.95
Marital Status (Dummy; Ref: Unmarried)	0.84***	0.89***	0.67***
OBC (Dummy; Ref: Other)	1.08***	0.94***	1.40**
Dalit (Dummy; Ref: Other)	1.29***	1.30	1.27***
Adivasi (Dummy; Ref: Other)	1.19	1.13	1.30**
Muslim (Dummy; Ref: Other)	1.10***	1.15***	0.96
Individual Characteristics:			
5 years of schooling (Dummy)	1.40***	1.30*	1.27***
10 years of schooling (Dummy)	0.93**	0.83***	1.13
12 years of schooling (Dummy)	2.56***	2.51***	2.62***
Undergraduate Degree (Dummy)	0.73***	0.63***	0.95
Masters Degree (Dummy)	1.08	0.94	1.30
Household Characteristics:			
Log of Consumption Expenditure	1.13***	1.16***	1.13**
Household Head's Schooling	1.57**	1.69	1.33
Female Household Head (Dummy)	0.94*	0.93	0.93*
Household Size	0.99***	1.00***	0.99***
Household Occupation: Self-employment	1.13***	1.06***	1.23**
Household Occupation: Salaried	0.96***	0.96***	0.96*

Household Occupation: Wage Work	1.56*	1.68*	1.29
Context Characteristics:			
Number of TVET institutions	1.00*	1.01	1.00
Unemployment Rate	1.06**	1.06*	1.06
Average 10-year rainfall	1.00	1.00**	1.00
N	*235331*	*119571*	*115760*

Source: Employment and Unemployment Survey of India (2009-10)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5

Odds ratio estimates of factors predicting participation in formal and informal TVET among 15-59 year olds (2009-10)

	Formal	Informal
Demographic Controls:		
Age	1.14***	1.16***
Age Squared	1.00***	1.00***
Female (Dummy; Ref: Male)	0.65***	0.33***
Urban (Dummy; Ref: Rural)	1.06*	0.96
Marital Status (Dummy; Ref: Unmarried)	0.68***	0.98***
Dalit (Dummy; Ref: Other)	1.31***	1.11
Adivasi (Dummy; Ref: Other)	1.07	1.16**
OBC (Dummy; Ref: Other)	1.23***	1.32***
Muslim (Dummy; Ref: Other)	0.98	1.12***
Individual Characteristics:		
5 years of schooling (Dummy)	3.26***	1.47***
10 years of schooling (Dummy)	2.39***	0.79***
12 years of schooling (Dummy)	6.48***	0.73***
Undergraduate Degree (Dummy)	0.74***	0.75***
Masters Degree (Dummy)	1.01	1.00
Household Characteristics:		
Log of Consumption Expenditure	1.32***	0.99
Household Size	0.93***	0.99**
Household Occupation: Self-employment	0.98*	2.04***
Household Occupation: Salaried	1.37*	1.60***
Household Occupation: Wage Work	0.91	1.06
Household Head's Schooling	1.01*	0.98***
Female Household Head (Dummy)	1.19*	1.13
Context Characteristics:		
Number of TVET institutions	1.01***	1.00
Unemployment Rate	1.08***	1.05
Average 10-year rainfall	1.00	1.00*

N	*7770*	*11549*

Source: Employment and Unemployment Survey of India (2009-10)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6

Odds ratio estimates of factors predicting participation in formal and informal TVET among 15-59 year olds, by gender (2009-10)

	Males		Females	
	Formal	Informal	Formal	Informal
Demographic Controls:				
Age	1.86***	2.54***	2.18***	1.91***
Age Squared	0.99***	0.98***	0.99***	0.99***
Urban (Dummy; Ref: Rural)	1.08	0.96	1.17*	1.04
Marital Status (Dummy; Ref: Unmarried)	0.78**	0.87*	0.45***	0.74***
Dalit (Dummy; Ref: Other)	1.18*	0.95	1.49***	1.30**
Adivasi (Dummy; Ref: Other)	1.01	1.05	1.05	1.84***
OBC (Dummy; Ref: Other)	1.27***	1.19**	1.06	1.31***
Muslim (Dummy; Ref: Other)	1.07	1.27***	1.00	1.07
Individual Characteristics:				
5 years of schooling (Dummy)	1.03	1.31***	2.25***	1.41***
10 years of schooling (Dummy)	3.46	0.66***	1.89***	0.81**
12 years of schooling (Dummy)	5.70***	0.54***	4.27***	0.73**
Undergraduate Degree (Dummy)	0.79***	0.67***	0.89	0.79
Masters Degree (Dummy)	0.95	0.90	1.11	1.08
Household Characteristics:				
Log of Consumption Expenditure	1.35***	1.06	1.17**	1.10
Household Size	0.93***	0.98	0.94***	0.97*
Household Occupation: Self-employment	0.89	2.23***	0.95	1.52***
Household Occupation: Salaried	1.21*	1.76***	1.14	1.16
Household Occupation: Wage Work	0.80**	1.11	1.05	0.92
Household Head's Schooling	1.01	0.98**	1.01	0.97***
Female Household Head (Dummy)	1.15	0.95	1.14	1.13
Context Characteristics:				
Number of TVET institutions	1.01***	1.00	1.01*	1.00
Unemployment Rate	1.05***	1.03	1.06***	1.04
Average 10-year rainfall	1.00	1.00*	1.00	1.00
N	*2524*	*3188*	*1698*	*1372*

Appendix B

Sampling Strategy (Indian Human Development Survey)

The primary sampling units (PSUs) in the IHDS are urban blocks and villages selected using different designs. The probability proportional to size technique was used to sample urban blocks (Desai et al., 2010). A more complex design was adopted for selection of rural PSUs. Half of the rural households included in the IHDS were randomly selected from the sample of a previous survey – the Human Development Profile of India (HDPI) conducted in 1993-1994 covering 16 major states, 195 districts and 1,765 villages. About 80% of the households randomly selected from the HDPI sample could be contacted for "re-interview" for the IHDS. Those households that could not be contacted for "re-interview" were replaced with other randomly selected households within the same district.

The other half of the rural households sampled in the IHDS included a random selection from districts excluded in the HDPI sample.

In "re-interview" districts, two additional villages were randomly selected based on probability proportional to size. Representativeness checks determined that there were no differences between the "fresh" and "re-interviewed" samples on key demographic and economic outcomes (Desai et al., 2010).[59]

Occupational Categories of Wage Workers

It should be noted that within the sample of those currently employed in paid work, individuals were further classified on the basis of primary occupation. The following occupational categories are available in the IHDS data - agricultural wage work, nonagricultural wage work, salaried work, and self-employment. There was considerable overlap between these categories. For example, 22% respondents who

[59] See Desai et al, (2010, pp. 214-222) for a detailed explanation of sampling methods and tests of robustness, including comparisons with other nationally representative surveys and the Census of India.

claimed that their primary occupation was self-employment also claimed salaried work as their primary employment. Similarly 9.4% respondents who were primarily engaged in agricultural wage work also claimed nonagricultural wage work as their primary occupation. In order to create a neat classification of “salaried”, “self-employed”, and “informal sector workers”, the following procedure was adopted –

1) Cross-tabs of pairs of occupational categories were created to identify cases that claimed primary appointment in more than one occupational category. Table 10 shows these cross-tabs.
2) For each pair of occupational categories examined, the larger proportion of cases were classified as the primary occupation (>=240 hours). See cells in bold font in table 10. This was true in the case of all pairs except agricultural and nonagricultural wage labor where the difference in proportion of agricultural laborers who also identified as nonagricultural labor was very similar.
3) Cases in the majority occupational category (see bolded cells in table 10) were retained as belonging to that category.
4) The agricultural and nonagricultural wage workers together formed a separate category that represents “informal sector workers”. See table 11a and table 11b for sample sizes of each occupational category.

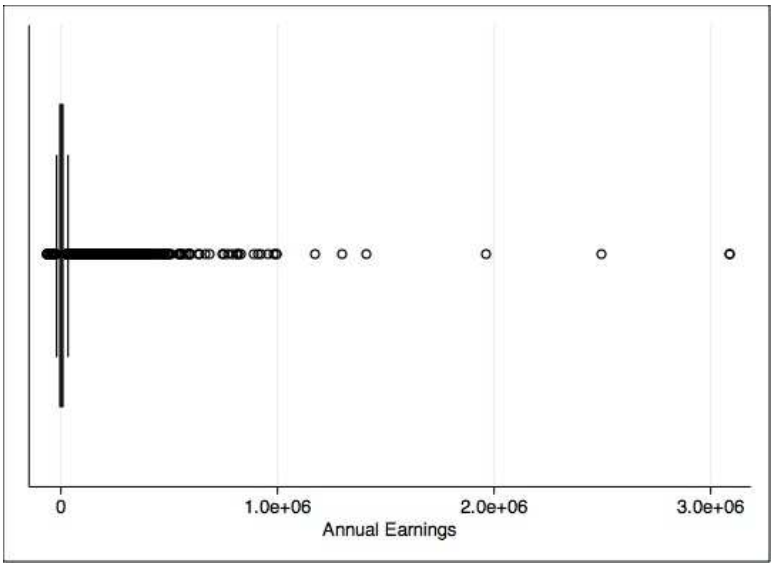

Figure B.1. Distribution of annual earnings (in Indian Rupees) with outlying values (untrimmed sample for returns to general education)

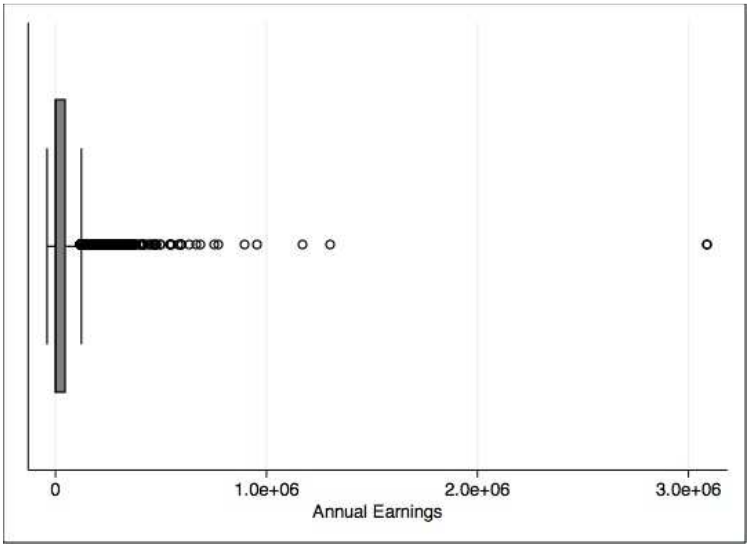

Figure B.2. Distribution of annual earnings (in Indian Rupees) in the untrimmed TVET sample

Table B.1

First stage results (Predicting labor force participation – For returns to schooling)

	Coef.	SE
Years of schooling	-0.015***	0.002
Age-Between 15-21 years	-0.564***	0.038
Age-Between 22-28 years	0.213***	0.035
Age-Between 29-39 years	0.410***	0.031
Female (Ref: Male)	-0.571***	0.039
Age Group 1*Female	0.012	0.043
Age Group 2*Female	-0.464***	0.039
Age Group 3*Female	-0.253***	0.037
Urban (Ref: Rural)	-0.425***	0.022
Marital Status (Ref: Unmarried)	0.853***	0.028
Marital Status*Female	-0.734***	0.036
Social Group - OBC (Ref: Non-SCST)	0.058*	0.025
Social Group - Dalit (Ref: Non-SCST)	-0.028	0.027
Social Group - Adivasi (Ref: Non-SCST)	0.212***	0.044
Religious Group - Muslim (Ref: Others)	-0.169***	0.031

Ability (>60% in grade 10)	0.058	0.038
Ability (<60% in grade 10)	0.075***	0.023
English Fluency	-0.050	0.028
Exclusion Restrictions:		
Household size	-0.040***	0.005
Number of children in the household	0.059***	0.008
Household Assets	-0.041***	0.003
Wald Chi-2	*142.26***	
rho	0.231***	0.019
sigma	1.094***	0.010
lambda	0.253***	0.021

Table B.2

First stage results (Predicting completed years of schooling)

	Coef.	SE
Age-Between 15-21 years	1.669***	0.088
Age-Between 22-28 years	1.727***	0.058
Age-Between 29-39 years	0.980***	0.040
Female (Ref: Male)	-1.771***	0.097
Age Group 1*Female	0.796***	0.127
Age Group 2*Female	0.140	0.098
Age Group 3*Female	-0.115	0.070
Urban (Ref: Rural)	0.332***	0.049
Marital Status (Ref: Unmarried)	-0.445***	0.068
Marital Status*Female	-0.546***	0.093
Social Group - OBC (Ref: Non-SCST)	-0.353***	0.053
Social Group - Dalit (Ref: Non-SCST)	-0.773***	0.062
Social Group - Adivasi (Ref: Non-SCST)	-0.820***	0.083
Religious Group - Muslim (Ref: Others)	-0.932***	0.067
Ability (>60% in grade 10)	3.649***	0.087
Ability (<60% in grade 10)	3.465***	0.061
English Fluency	2.654***	0.058
Instrument:		
Household Head's Education	0.426***	0.005

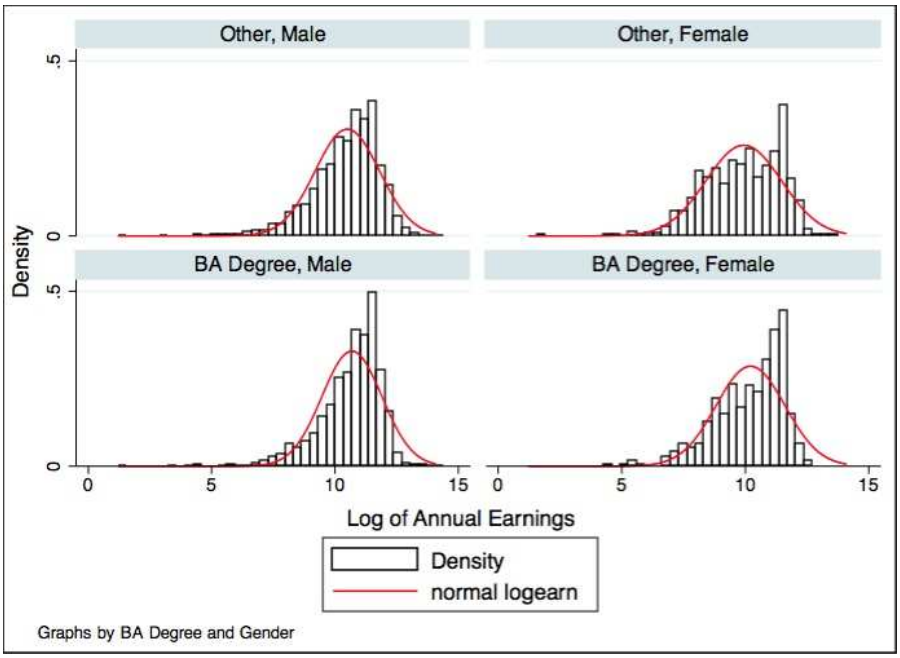

Figure B.3. Distribution of log annual earnings by gender and Bachelor's degree attainment

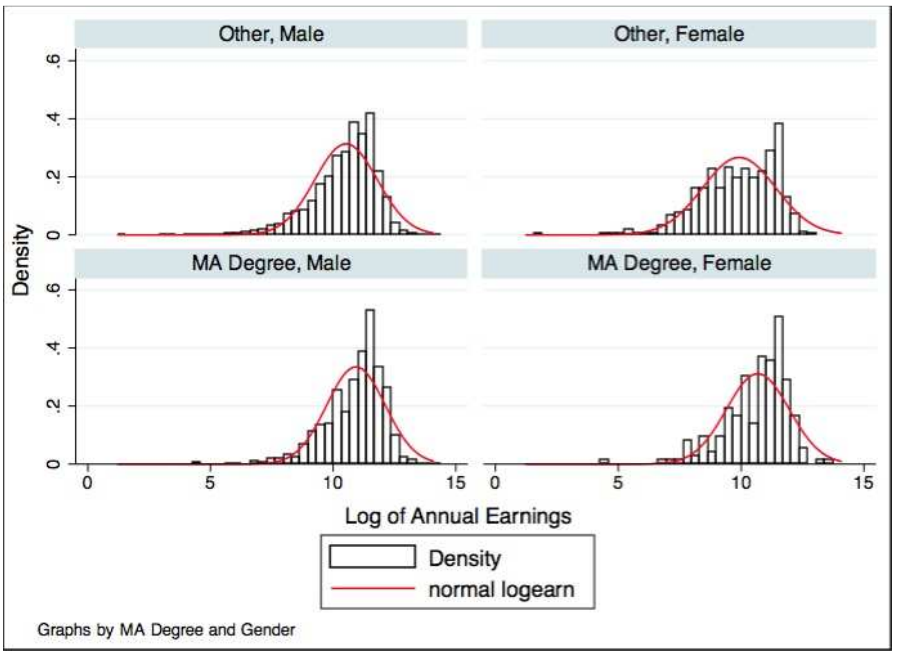

Figure B.4. Distribution of log annual earnings by gender and Master's degree attainment

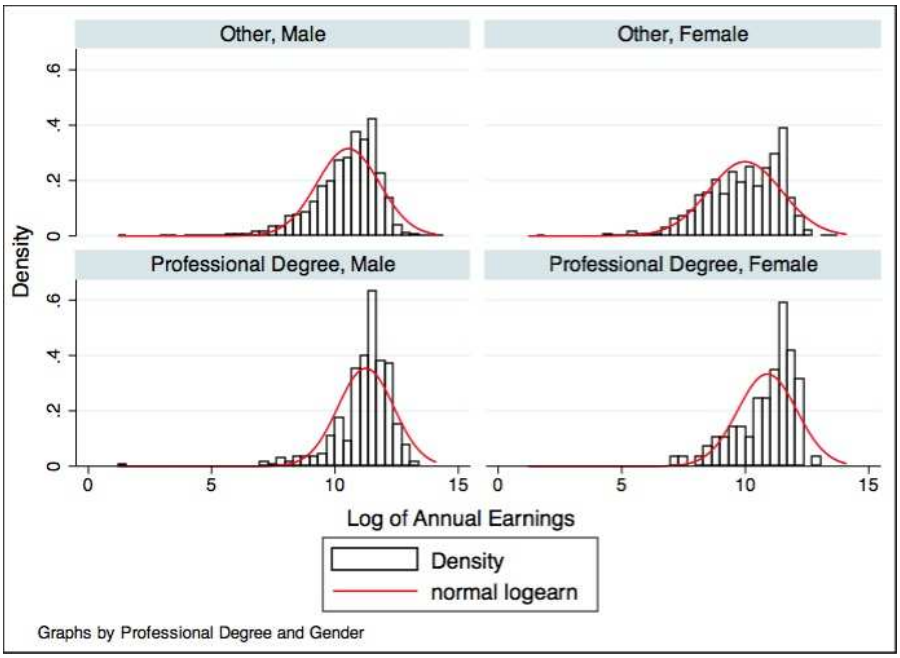

Figure B.5. Distribution of log annual earnings by gender and Professional degree attainment

Table B.3

First stage results (Predicting labor force participation – For returns to TVET)

	Coef.	SE
BA Degree (Ref: Other)	0.053	0.046
MA Degree (Ref: Other)	0.221***	0.059
Professional Degree (Ref: Other)	0.272**	0.087
TVET (Ref: Other)	0.016	0.073
Age-Between 15-21 years	-1.288***	0.097
Age-Between 22-28 years	-0.300***	0.077
Age-Between 29-39 years	0.187**	0.064
Female (Ref: Male)	-0.473***	0.104
Age Group 1*Female	0.313*	0.132
Age Group 2*Female	-0.200*	0.094
Age Group 3*Female	-0.311***	0.084
Urban (Ref: Rural)	-0.191***	0.051
Marital Status (Ref: Unmarried)	0.706***	0.067
Marital Status*Female	-1.033***	0.099
Social Group - OBC (Ref: Non-SCST)	-0.008	0.045
Social Group - Dalit (Ref: Non-SCST)	0.012	0.058
Social Group - Adivasi (Ref: Non-SCST)	0.114	0.091
Religious Group - Muslim (Ref: Others)	-0.041	0.071

Ability (>60% in grade 10)	0.055	0.056
Ability (<60% in grade 10)	0.072	0.047
English Fluency	-0.005	0.045
Exclusion Restrictions:		
Household size	-0.038***	0.009
Number of children in the household	0.109***	0.018
Household Assets	-0.033***	0.006
rho	*0.344****	*0.049*
sigma	*1.068****	*0.017*
lambda	*0.368****	*0.055*
Wald Chi-2	*40.84*	
Unweighted N		*7293*
Weighted N		*30000000*

Table B.4

First stage results (Predicting completed years of schooling – TVET sample)

	Coef.	SE
Age-Between 15-21 years	-0.867***	0.164
Age-Between 22-28 years	0.216***	0.095
Age-Between 29-39 years	0.242***	0.059
Female (Ref: Male)	0.253***	0.160
Age Group 1*Female	-0.126***	0.255
Age Group 2*Female	-0.134***	0.165
Age Group 3*Female	-0.218***	0.139
Urban (Ref: Rural)	0.147***	0.067
Marital Status (Ref: Unmarried)	-0.219***	0.104
Marital Status*Female	-0.307***	0.151
Social Group - OBC (Ref: Non-SCST)	-0.137***	0.065
Social Group - Dalit (Ref: Non-SCST)	-0.193***	0.102
Social Group - Adivasi (Ref: Non-SCST)	-0.141***	0.158
Religious Group - Muslim (Ref: Others)	-0.177***	0.106
Ability (>60% in grade 10)	1.217***	0.083
Ability (<60% in grade 10)	0.764***	0.083
English Fluency	1.283***	0.086
Household Head's Education	0.083***	0.008
Household Head's TVET Participation	-1.204***	0.161

Table B.5

First stage results (Predicting TVET participation – TVET sample)

	Coef.	SE
Age-Between 15-21 years	0.101	0.058
Age-Between 22-28 years	0.011	0.016
Age-Between 29-39 years	0.003	0.007
Female (Ref: Male)	-0.059*	0.026
Age Group 1*Female	-0.083	0.066
Age Group 2*Female	-0.014	0.025
Age Group 3*Female	0.010	0.018
Urban (Ref: Rural)	0.005	0.006
Marital Status (Ref: Unmarried)	-0.030	0.022
Marital Status*Female	0.051*	0.025
Social Group - OBC (Ref: Non-SCST)	0.010	0.007
Social Group - Dalit (Ref: Non-SCST)	0.028	0.018
Social Group - Adivasi (Ref: Non-SCST)	0.018	0.021
Religious Group - Muslim (Ref: Others)	-0.002	0.010
Ability (>60% in grade 10)	0.026**	0.009
Ability (<60% in grade 10)	0.021*	0.009
English Fluency	-0.012	0.011
Household Head's Education	-0.004**	0.001
Household Head's TVET Participation	0.813***	0.025

Table B.6

OLS estimates of returns to TVET – with PSU-level fixed effects

	Coef.	SE
BA Degree (Ref: Other)	0.076*	0.037
MA Degree (Ref: Other)	0.175***	0.050
Professional Degree (Ref: Other)	0.344***	0.067
TVET (Ref: Other)	0.233***	0.060
Age-Between 15-21 years	-0.972***	0.088
Age-Between 22-28 years	-0.542***	0.046
Age-Between 29-39 years	-0.198***	0.030
Female (Ref: Male)	-0.033	0.098
Age Group 1*Female	-0.383**	0.150

Age Group 2*Female	-0.363***	0.098
Age Group 3*Female	-0.207**	0.081
Urban (Ref: Rural)	(omitted)	
Marital Status (Ref: Unmarried)	0.282***	0.045
Marital Status*Female	-0.312***	0.090
Social Group - OBC (Ref: Non-SCST)	-0.093**	0.038
Social Group - Dalit (Ref: Non-SCST)	-0.044	0.053
Social Group - Adivasi (Ref: Non-SCST)	-0.075	0.091
Religious Group - Muslim (Ref: Others)	0.100	0.079
Ability (>60% in grade 10)	0.376***	0.049
Ability (<60% in grade 10)	0.125**	0.042
English Fluency	0.243***	0.042
Intercept	10.177***	0.066
Unweighted N	*7877*	

Appendix C

Submit this document on the Trust/Society letterhead, duly signed and stamped by the President and Secretary.

Date:

To,
President,
Lend-A-Hand India,
9, Jeevan Vihar Society, Near Pride Panorama,
Off. Senapati Bapat Road, Pune 411 016

Sub: Resolution passed to start Project Swadheen in our school: (Name) _____
and Commitments by the Trust/Society and Lend-A-Hand India

Dear Sir,

Our Trust/Society and the concerned school Head Master have learned about Project Swadheen under which a very beneficial curriculum for students on 8th to 10th grade is taught. Our Society/Trust is therefore interested to start this program in our above school.

The Trust/Society has passed the following resolution in its meeting held on _____.

"Resolved that project Swadheen be started in (School name) :_____ for the academic year June 2012. **It is further resolved that the Trust/Society and the school is committed to continue the program for three years minimum and further.** It was further resolved that the Trust/Society will open a separate bank account to be operated by the Head Master and another authorized signatory from the school's permanent staff to receive and disburse the funds under project Swadheen. The Trust/Society assures to extend all cooperation and assistance to make this program a grand success."

The Trust/Society and the undersigned assure that the school will take all the preparatory steps in order to launch the project from June _____ as under:

1. The school will assign TEN school periods per week for the teaching of the program.
2. Appropriate room will be provided to house the workshop as required under the program.
3. The school will make provision of at least half an acre of land, belonging either to the school, or on lease for carrying out the practicals in agriculture/horticulture.
4. The workshop will be equipped with the necessary tools and equipments as per the list provided in the project implementation document.
5. The school will identify and will jointly select FOUR instructors, who are ideally local trade practitioners in the relevant subjects for each section of the program.
6. The science and / or mathematics teacher in the school will be made available to take the theory sessions under the curriculum including costing and drawing.
7. The science and / mathematics teacher or any other appropriate person will act as coordinator of Swadheen program and would maintain the necessary records, reports and accounts.
8. The school and its management are committed to make the program sustainable after completing THIRD year.
9. The school will charge a suitable tuition fee to the students undergoing this program to meet part of the program expenses.
10. The Head Master will act as a program leader and will be responsible for proper accounting of the funds, timely payments to the staff under the program, and deposits of funds generated from collection of student fees, community work, and occasional donations received for the program, etc.
11. The program coordinator, under the guidance of the Head Master, will be responsible for making up of the lost time / teaching due to unplanned holidays or any other unforeseen reasons.
12. The Head Master will ensure that the instructors and the coordinator attend the training offered under the program, maintain the necessary documentation and reports.
13. The Head Master and Coordinator will attend pre-planned orientation / training and feedback meetings etc. at their cost. LAHI will bear the training and other incidental expenses of such meetings. Instructors are paid honorarium for 12 months and therefore they are also expected to attend the training, meetings at their cost.
14. The school will facilitate the visit Lend-A-Hand India's field officer who will visit the school at least once in a month for guidance, on the job training, data collection, and troubleshooting, if any. The LAHI field officer will ordinarily spend two days at the school and wherever the schools are remotely located the school will facilitate his/her stay and food arrangements at cost.
15. In case there is some delay in disbursement due to unavoidable circumstances, the school will ensure timely purchase of material and payment to instructors.
16. In case the coordinator or instructors do not perform satisfactorily, the management committee and the Head Master will reallocate the task to another suitable person.

We understand that Lend-A-Hand India, under project Swadheen, is committed to extend following assistance:

- Provide partial financial assistance for THREE years to run the program (removed 80% word)
- Provide all the necessary assistance to implement and monitor the program
- Provide sufficient and necessary training to the instructors who will deliver the program to the students
- Provide monitoring assistance and timely reinforcements to the program from time to time with the help of a field officer.

We have read and accept the contents of this undertaking and look forward to hearing from you and our mutually beneficial long-term association.

Secretary President

(Trust/Society Rubber stamp)

Figure C.1. Template of Resolution from the School's Management Committee to implement IBT

Table C.1

Suggested Timetable for IBT Schools

Standard 8th						
Period	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1	Theory + IT	Marathi	Marathi	Marathi	Marathi	Maths
2	Drawing+Costing	Hindi	Drawing	Hindi	Hindi	Hindi
3	Drawing+Costing	Geography	Hindi	Drawing	English	Marathi
4	IBT Practical	English	English	English	History	English
5	IBT Practical	Science	Science	Science	Science	History
6	IBT Practical	Maths	Geography	Maths	Maths	
7	IBT Practical	English	Maths	Hindi	History	
8	IBT Practical	Science	History	English	Maths	
9	IBT Practical	Drawing	PT	PT	PT	
10	IBT Practical					
Standard 9th						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1	Science-II	Theory + IT	Hindi	Science-II	Science-II	English
2	English	Drawing+Costing	Science-II	English	English	English
3	Geometry	Drawing+Costing	Marathi	Geometry	History	Hindi
4	Algebra	IBT Practical	Geometry	Algebra	Algebra	Geography
5	Economics	IBT Practical	Science-I	Science-I	English	Marathi
6	Science-II	IBT Practical	English	English	Geography	Gen.Sci.
7	Marathi	IBT Practical	Marathi	Hindi	Algebra	
8	Hindi	IBT Practical	History	NCC	Marathi	
9	PT	IBT Practical	PT	NCC/PT	PT	
10		IBT Practical				
Standard 10th						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
1	Marathi	Marathi	Theory + IT	Algebra	Marathi	English
2	Science-I	Algebra	Drawing+Costing	Geography	Geography	Hindi
3	Geography	Science-II	Drawing+Costing	History	History	Geometry
4	Science-II	Geometry	IBT Practical	Science-II	Science-II	Algebra
5	Marathi	Marathi	IBT Practical	English	English	Marathi
6	Algebra	History	IBT Practical	Geometry	Hindi	Gen. Sci.
7	Marathi	Marathi	IBT Practical	Hindi	Science-I	
8	Hindi	English	IBT Practical	Science-I	English	
9	PT	PT	IBT Practical	PT	PT	
10			IBT Practical			

Table C.2

Response rates by district

	Surveyed		Not surveyed	
	Treatment	Comparison	Treatment	Comparison
Ahmadnagar	22	273	28	12
Jalgaon	107	439	76	107
Pune	104	576	51	335
Raigad	45	191	3	83
Thane	27	416	92	222

Table C.3

Analytic sample by school (Treatment group)

Treatment School ID	Percent
27031402101	13.12
27031403401	9.38
27031405601	11.88
27211301202	13.12
27211809301	1.88
27240807201	5.62
27250100201	11.88
27250502801	1.88
27251008401	6.25
27251101301	14.38
27261203601	10.62
N	*160*

Table C.4

Analytic sample by school (Comparison group)

Comparison School ID	Percent
27031402101	4.39
27031404101	1.89
27031407502	5.34
27031407601	0.34
27031408307	4.74
27031409301	4.31

27031410301	7.24
27211300701	1.64
27211301202	3.96
27211302501	6.12
27211806101	1.72
27211807101	1.81
27211807901	1.46
27211809301	3.45
27240804901	0.86
27240811301	4.74
27240812101	2.67
27250103503	0.69
27250107301	3.88
27250107801	3.53
27250506001	3.62
27250510301	3.62
27251009701	0.34
27251011301	0.34
27251015001	0.26
27251100501	3.45
27251101901	1.46
27251108501	2.5
27261203901	9.91
27261205501	0.43
27261205509	0.86
27261206601	5.86
27261216904	2.58
N	*1161*

Table C.5

Proportion of missing data on relevant variables

Variable	Number missing	Percent missing
Outcome:		
Postsecondary enrollment (Dummy)	204	9.6
Pre-intervention academic achievement:		
Grade 7 Performance (Z scores)	472	22.22
Controls:		
Female (Dummy)	0	0
Religious Group – Hindu (Dummy)	44	2.07

Social Group – Dalit (Dummy)	65	3.06
Social Group – Adivasi (Dummy)	65	3.06
Social Group – OBC (Dummy)	65	3.06
Household size	66	3.11
High income household (Dummy)	92	4.33
Medium income household (Dummy)	92	4.33
Low income household (Dummy)	92	4.33
Household head in agricultural work (Dummy)	206	9.7
Household head self-employed (Dummy)	0	0
Household head in informal work (Dummy)	222	10.45
Household head in salaried work (Dummy)	222	10.45
Distance to nearest town/city	0	0
Distance to bus station	0	0
Access to public transport (Dummy)	0	0
Village size	0	0
N	*2124*	

Table C.6

Log odds estimates of participation in IBT

	Analytic Sample 1		Analytic Sample 2	
	Log Odds	SE	Log Odds	SE
Female (Dummy)	-1.027	0.556	-0.370	0.683
Hindu (Dummy)	-1.855	0.376	-2.217	0.447
Social Group – Dalit (Dummy)	-0.201	0.442	-0.581	0.544
Social Group – Adivasi (Dummy)	-0.918	0.363	-1.479	0.458
Social Group – OBC (Dummy)	-0.778	0.368	-0.818	0.435
Household Size	0.247	0.080	0.375	0.103
High income household (Dummy)	0.552	0.291	-0.137	0.381
Medium income household (Dummy)	0.636	0.233	0.449	0.300
Household head in agriculture (Dummy)	1.424	0.642	2.549	1.128
Household head in self-employment (Dummy)	0.996	0.664	2.100	1.159
Household head in salaried work (Dummy)	0.907	0.666	1.642	1.165
Grade 7 performance (Z scores)	-0.018	0.020	0.013	0.026
Distance to nearest town/city	0.077	0.011	0.031	0.018
Distance to bus station	-0.076	0.020	-0.353	0.058
Access to public transport (Dummy)	0.501	0.641	1.076	0.677
Village size	0.000	0.000	0.000	0.000
Female*Hindu	0.626	0.542	0.049	0.674

Female*Dalit	-0.866	0.895	-0.977	1.177
Female*Adivasi	0.423	0.470	-0.838	0.626
Female*OBC	-0.600	0.534	-1.427	0.691
Female*Grade 7 performance	0.027	0.020	-0.004	0.024
Distance to nearest town/city*Grade 7 performance	0.001	0.001	-0.001	0.002
N	*1321*		*1139*	

Note. Pseudo R^2 (Analytic Sample 1) = 0.22; Pseudo R^2 (Analytic Sample 2) = 0.32

Table C.7

Means on select indicators for the treatment groups from the unmatched and matched samples

	Unmatched sample II		Matched sample II	
	Mean	SD	Mean	SD
Postsecondary enrollment (Dummy)	0.72	0.45	0.73	0.44
Age	20.07	1.36	20.13	1.34
Female (Dummy)	0.37	0.49	0.41	0.49
Hindu (Dummy)	0.71	0.45	0.73	0.44
Social Group – Dalit (Dummy)	0.07	0.26	0.07	0.25
Social Group – Adivasi (Dummy)	0.26	0.44	0.27	0.44
Social Group – OBC (Dummy)	0.17	0.38	0.15	0.36
Household Size	4.74	1.11	4.70	1.03
High income household (Dummy)	0.17	0.38	0.18	0.39
Medium income household (Dummy)	0.53	0.50	0.52	0.50
Low income household (Dummy)	0.30	0.46	0.30	0.46
Household head in agriculture (Dummy)	0.70	0.46	0.72	0.45
Household head in self-employment (Dummy)	0.17	0.37	0.13	0.34
Household head in informal work (Dummy)	0.01	0.09	0.01	0.10
Household head in salaried work (Dummy)	0.13	0.34	0.13	0.34
Grade 7 Performance	64.59	10.29	64.71	10.50
Grade 7 performance (Z scores)	-0.12	8.87	-0.03	9.02
Distance to nearest town/city	12.78	8.55	12.58	8.40
Access to public transport (Dummy)	0.97	0.16	1.51	1.99
Distance to bus station	1.57	2.24	0.98	0.14
Village size	7928.70	8583.02	6860.00	7127.98
N	*115*		*105*	

Table C.8

Means on select indicators for the comparison groups from the unmatched and matched samples

	Unmatched sample II		Matched sample II	
	Mean	SD	Mean	SD
Postsecondary enrollment (Dummy)	0.50	0.50	0.46	0.50
Age	19.87	1.30	19.77	1.16
Female (Dummy)	0.47	0.49	0.26	0.49
Hindu (Dummy)	0.96	0.21	0.82	0.39
Social Group – Dalit (Dummy)	0.09	0.29	0.06	0.23
Social Group – Adivasi (Dummy)	0.21	0.41	0.19	0.39
Social Group – OBC (Dummy)	0.29	0.45	0.13	0.34
Household Size	4.36	1.08	4.73	1.53
High income household (Dummy)	0.21	0.41	0.22	0.42
Medium income household (Dummy)	0.39	0.49	0.56	0.50
Low income household (Dummy)	0.40	0.49	0.22	0.42
Household head in agriculture (Dummy)	0.55	0.50	0.68	0.47
Household head in self-employment (Dummy)	0.22	0.41	0.15	0.36
Household head in informal work (Dummy)	0.05	0.21	0.03	0.17
Household head in salaried work (Dummy)	0.19	0.39	0.14	0.35
Grade 7 Performance	57.74	11.17	58.64	11.52
Grade 7 performance (Z scores)	0.52	10.40	0.94	10.86
Distance to nearest town/city	9.36	5.62	12.15	8.92
Access to public transport (Dummy)	0.92	0.27	1.59	2.47
Distance to bus station	4.40	4.06	0.98	0.14
Village size	5260.83	3717.36	5658.84	3401.95
N	*1024*		*105*	

Table C.9

Mean school characteristics for all treatment and comparison schools

	All Schools		Treatment schools		Comparison schools	
	Mean	SD	Mean	SD	Mean	SD
Recognized-Aided (Dummy)	0.81	0.39	0.82	0.40	0.81	0.40
Marathi Medium (Dummy)	0.88	0.33	1.00	0.00	0.84	0.37
Principal's Education	1.10	0.65	1.18	0.60	1.08	0.67
Principal's Salary	42555.73	10640.10	39081.09	10752.11	43588.73	10531.15
Number of teachers	16.27	11.73	13.00	8.12	17.21	12.52
Number of staff	23.20	13.28	22.00	9.21	23.55	14.33
Number of grades	6.10	1.75	5.45	2.30	6.29	1.54
School size	513.57	392.65	534.82	525.62	507.42	353.79
Grade size	80.93	48.17	86.44	63.25	79.33	43.80
Number of classrooms	219.82	1429.21	16.09	9.39	278.79	1622.96
School Infrastructure Index*	5.76	1.27	6.45	0.93	5.55	1.29
Grade 10 completion rate (2009)	0.84	0.17	0.88	0.10	0.83	0.18
Grade 9 retention rate (2009)	0.95	0.15	0.99	0.02	0.94	0.17
Grade 10 retention rate (2009)	0.87	0.52	0.72	0.56	0.91	0.51
N	*44*		*11*		*33*	

Note. * School infrastructure index includes availability of cultivable land, water, electricity, playground, computer laboratory, and audiovisual equipment at the school.

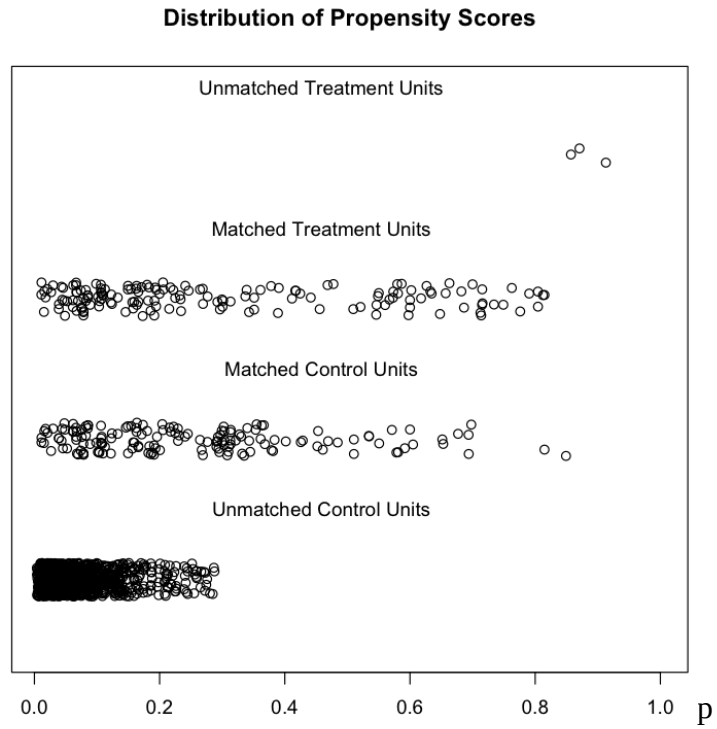

Figure C.2. Distribution of propensity scores in analytic sample I using 1:1 nearest neighbor matching

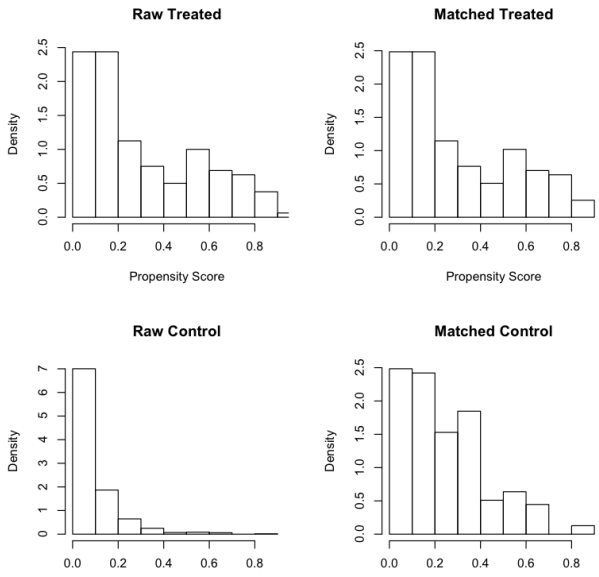

Figure C.3. Comparison of propensity score distributions in the original and matched data (For analytic sample I – using1:1 nearest neighbor matching)

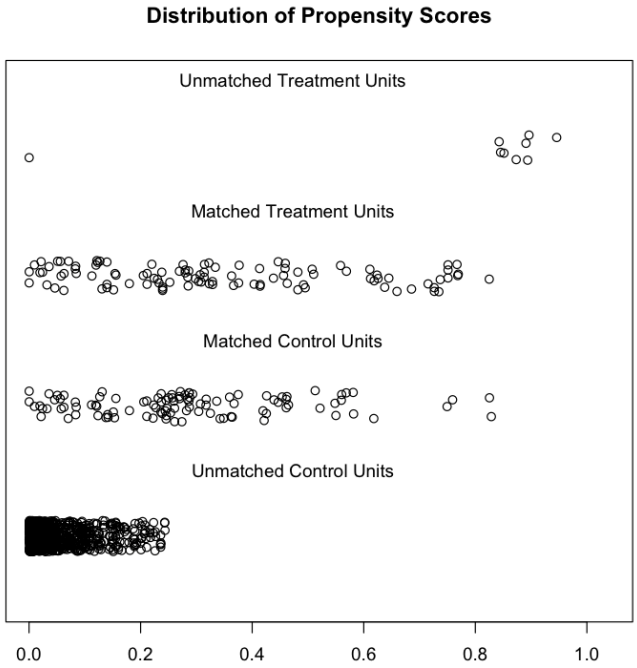

Figure C.4. Distribution of propensity scores in analytic sample II using 1:1 nearest neighbor matching

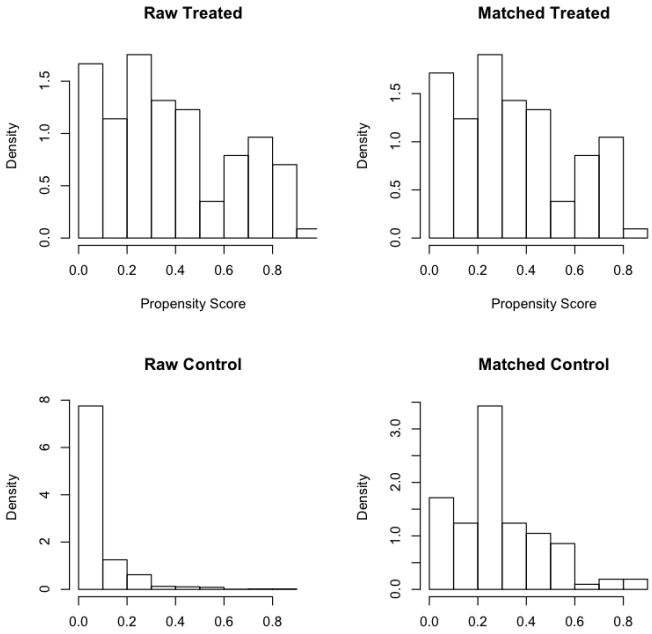

Figure C.5. Comparison of propensity score distributions in the original and matched data (For analytic sample II – using1:1 nearest neighbor matching)

Table C.10

List of treatment and potential non-treatment schools with key selection indicators

Distance from treatment	Village Name	Sub-District	Grades	Minority	Language	Electricity	School Management	Hilly Area	Tribal Area	School Type	Size
Treatment	**Gawadewadi**	**Ambegaon**	**8th-10th**	**0**	**Marathi**	**1**	**1**	**1**	**0**	**1**	**0**
4	Awsari Khurd	Ambegaon	8th-10th	0	Marathi	1	1	1	0	1	0
15	Chas	Ambegaon	5th-10th	0	Marathi	1	1	1	0	1	0
25	Chincholi	Ambegaon	5th-10th	0	Marathi	1	1	1	0	1	0
close	Dimbhe Khurd	Ambegaon	5th-10th	1	Marathi	1	1	1	1	1	1
Treatment	**Hingangaon**	**Haveli**	**8th-10th**	**1**	**Marathi**	**1**	**1**	**0**	**0**	**1**	**0**
-	Ambegaon Budruk	Haveli	1st-10th	1	Marathi	0	1	0	0	1	0
-	Shivane	Haveli	5th-12th	1	Marathi, English	1	1	0	0	1	0
10	Pimpri Sandas	Haveli	5th-10th	0	Marathi	1	1	0	0	1	0
-	Alandi Mahtoba	Haveli	5th-10th	0	Marathi	1	1	0	0	1	0
Treatment	**Asade**	**Mulshi**	**5th-10th**	**0**	**Marathi, English**	**1**	**1**	**1**	**0**	**1**	**0**
12	Shere	Mulshi	5th-10th	0	Marathi	1	1	0	0	1	0
15	Pirangut	Mulshi	5th-10th	0	Marathi, English	1	1	1	0	1	0
32	Mutha	Mulshi	5th-10th	0	Marathi	1	1	0	0	1	1
15	Khechare	Mulshi	8th-10th	0	Marathi	0	1	1	0	1	0

Treatment	**Bhivadi**	**Purandar**	**5th-10th**	**0**	**Marathi**	**1**	**1**	**1**	**0**	**1**	**0**
10	Ketkavle	Purandar	5th-10th	0	Marathi	1	1	0	0	1	0
15	Chambali	Purandar	5th-10th	0	Marathi	1	1	1	0	1	0
19	Pargaon	Purandar	5th-10th	0	Marathi	1	1	1	0	1	0
Treatment	**Lohara**	**Rawer**	**8th-10th**	**0**	**Marathi**	**1**	**1**	**0**	**1**	**1**	**0**
3.7	Kusumba	Rawer	5th-10th	0	Marathi	1	1	0	1	1	0
18	Dasnur	Rawer	5th-10th	0	Marathi	1	1	0	0	1	0
20	Balwadi	Rawer	5th-10th	0	Marathi	1	1	0	0	1	0
10	Abhode	Rawer	5th-10th	0	Marathi	1	1	0	1	1	1
5	Utkhede	Rawer	5th-10th	0	Marathi	1	1	0	0	1	1
20	Nimbol	Rawer	5th-10th	0	Marathi	1	1	0	1	1	1
11	Rasalpur	Rawer	8th-10th	0	Marathi	1	1	0	0	1	1
12	Vivare	Rawer									
18	Nimbhora	Rawer	5th-10th	0	Marathi	1	1	0	0	1	1
5	Kumbharkheda	Rawer	5th-10th	0	Marathi	1	1	0	0	1	1
9	Lalmati	Rawer	8th-10th	0	Marathi	1	2	1	1	1	0
8	Chinawel	Rawer	5^{th}-10^{th}	0	Urdu	1	1	0	0	1	0
11	Rasalpur	Rawer	5^{th}-10^{th}	0	Marathi, Hindi, English	1	1	0	0	1	0
Treatment	**Khiroda**	**Rawer**	**5^{th}-12^{th}**	**0**	**Marathi**	**1**	**1**	**0**	**0**	**1**	**0**
2	Rozoda	Rawer	5^{th}-10^{th}	0	Marathi	1	1	0	0	1	0
11	Maskawad	Rawer	5^{th}-12^{th}	0	Marathi	1	1	0	0	1	0

12	Thorgavahan	Rawer	5^{th}-12^{th}	0	Marathi	1	1	0	0	1	0
15	Udali	Rawer	5^{th}-10^{th}	0	Marathi	1	1	0	0	1	0
Treatment	**Pal**	**Rawer**	**8^{th}-12^{th}**	**0**	**Marathi, English, Hindi**	**1**	**1**	**1**	**1**	**1**	**0**
22	Rasalpur	Rawer	5^{th}-10^{th}	0	Marathi, Hindi, English	1	1	0	0	1	0
25	Raver	Rawer	5^{th}-12^{th}	0	Marathi	1	1	0	0	1	0
16	Abhode	Rawer	5^{th}-10^{th}	0	Marathi	1	1	0	1	1	1
16	Kusumba	Rawer	5^{th}-10^{th}	0	Marathi	1	1	0	1	1	0
24	Rozoda	Rawer	5^{th}-10^{th}	0	Marathi	1	1	0	0	1	0
22	Rasalpur	Rawer	8^{th}-10^{th}	0	Marathi	1	1	0	0	1	1
close	Pal	Rawer	5^{th}-10^{th}	0	Marathi, Hindi, English	1	1	1	1	0	0
close	Pal	Rawer	8^{th}-10^{th}	0	Urdu	1	1	1	1	0	1
Treatment	**Vikramgad**	**Vikramgad**	**5^{th}-12^{th}**	**1**	**Marathi**	**1**	**1**	**1**	**1**	**1**	**0**
10	Sakhare	Vikramgad	5^{th}-12^{th}	1	Marathi	1	1	1	1	1	0
	Kurze	Vikramgad	5^{th}-10^{th}	1	Marathi	1	1	1	1	1	0
close	Talwada	Vikramgad	8^{th}-12^{th}	1	Marathi	1	1	0	1	1	99
9	Alonde	Vikramgad	5^{th}-12^{th}	0	Marathi	1	1	1	1	1	0
Treatment	**Talasari**	**Talasari**	**5^{th}-10^{th}**	**0**	**Marathi**	**1**	**1**	**0**	**1**	**0**	**0**
13	Girgaon	Talasari	8^{th}-10^{th}	0	Marathi	1	1	0	1	0	1

26	Vewaji	Talasari	1st-12th	0	Hindi, Marathi, English	1	1	0	1	0	0
26	Vewaji	Talasari	1st-10th	0	English, Marathi, Hindi	0	1	0	1	0	1
close	Zari	Talasari	8th-10th	0	Marathi	1	1	0	1	1	0
Treatment	**Manchihill**	**Sangamner**	**5th-12th**	**1**	**Marathi**	**1**	1	**0**	**0**	**0**	**0**
-	Ozar	Sangamner	8th-10th	1	-	1	1	1	0	0	1
12	Vadgaon Pan	Sangamner	5th-12th	1	Marathi	1	1	1	0	1	0
25	Ghulewadi	Sangamner	5th-10th	0	Marathi	1	1	1	0	0	0
	Sakur	Sangamner	5th-12th	0	Marathi	1	1	1	0	0	1
12	Ashvi Khurd	Sangamner	5th-12th	0	Marathi	1	1	1	0	0	1
Treatment	**Shivkar**	**Panvel**	**5th-10th**	**0**	**Marathi**	**1**	**1**	**0**	**0**	**1**	**0**
8	Vakadi	Panvel	1st-10th	0	Marathi	1	1	0	0	1	0
10	Palaspe	Panvel	5th-10th	0	Marathi	1	1	0	0	1	0
10	Vaje	Panvel	5th-10th	0	Marathi	1	1	0	0	1	0
10	Poyanje	Panvel	5th-10th	0	Marathi	1	1	0	0	1	0

Source: Secondary Education Management Information System (http://semisonline.net/)
Note: School Management = Private (1), Public (0), Other (2); School Type = Recognized Aided (1), Recognized Unaided (0); Size = <40 students (1), >40 students (0)

Principal / School Survey Instrument

Informed Consent Agreement for the Vocational Education School Survey

I would like to invite you to participate in the present research study. The purpose of this study is to estimate the impacts of vocational education at the secondary school level on school completion, postsecondary enrollment, and employment outcomes among participants. Fifty-two (52) schools across the state of Maharashtra have been invited to participate in this study. Students who attended these schools will also be included in the study.

Your participation will entail completing a survey that gathers information about your school, the school staff, average performance of students in the school, and your perceptions of vocational education. This survey will take approximately 40 minutes to complete. Your participation in the survey is voluntary. If I ask any question you do not wish to answer, let me know and I will go to the next question. There are no consequences if you do not to answer the survey.

The data gathered as part of this research process will be de-identified once it has been converted to electronic format. All names and other identifiers will be removed from the data and will not be used or appear in any analysis or research report.

There are no known risks associated with your participation in this research. Your participation will help the research since your views are important to help us understand school-based vocational education programs.

If there is anything about the study or your participation that you do not understand, or if you wish to speak with someone about the study, you may contact, Namrata Tognatta at phone: **020-25884180** or email namratat@gse.upenn.edu .
At this time, do you have any questions about the survey? Do you agree to participate in this survey?

☐ Yes ☐ No

_____
Signature of Respondent and Date

_____
Signature of Witness and Date

Taluka _____, District _____, Maharashtra

Pre-printed school information

Block A. Identification

(Verify that this is complete before you leave for the field.)

A.1.School ID: | | | | | | | | | | | |

A.1. (a) Name: .....

A.1. (b) Address: Village/Post/Taluka/District

.....

.....

Pin Code | | | | | | | |

A.1. (c) Landline number: | | | | | | | | |

A.2. Date/Dates of survey: | | | | | | | | |

(DD) (MM) (YY)

A.3. Signature of surveyor: .....

A.4. Signature of supervisor: .....

A.5. Code for interviewer's result: | | iew conducted - 1; Refused - 2; No interview conducted – 3)

A. 6. If 3 in A.5: specify reason for no interview: ☐ (School closed - 1; Headmaster & Vice Principal not available - 2; Specify other reason here – 3: _____)

[Prior to visiting the school call the headmaster to get an appointment for the interview. If the Headmaster is not available, try to obtain information on his/her availability. The Vice Principal can be interviewed if the investigator feels that it is not going to be possible to interview the Headmaster. This decision must be discussed with the field supervisor. Use the space below to make notes]

Protocol: If option 3 in A.6, then this questionnaire goes to a field supervisor. The field supervisor must decide with the investigator if the Headmaster can be replaced with the Vice Principal at the school.

Investigator's Notes:

Notes for interviewer: The respondent must be the school principal, or vice-principal.

Block B: School Information

SCHOOL CHARACTERISTICS

B.1. When was the school established? ☐☐ ☐☐☐☐

MM YYYY

B.2. Type of school: ☐ (Grant-in-aid – 1; Non-grant – 2; Other – 3 (specify) _____)

B.3. Primary medium of instruction: ☐ glish – 1; Marathi – 2; Hindi – 3; Urdu – 4; Other – 5 (specify) _____)

[Circle multiple options if applicable]

B.4. Is there a Parents-Teachers Association (PTA) in the school? ☐ (Yes -1; No – 2) ***[Note: If NO, skip to B.7.]***

B.5. How many times a year does the association meet? ☐

B.6. What is the average attendance in these PTA meetings? ☐

B.7. Is the school managed by a School Management Committee (SMC)? ☐ (Yes -1; No – 2)

B.8. Is it a local management? ☐ (Yes -1; No – 2)

B.9. Name of the society/trust: .....

B.10. Name of the President: .....

B.11. Contact (Phone number): .....

B.12. Does the school provide IBT training? ☐ (Yes -1; No – 2)

[Note: If NO, skip to B.15.]

[Note: At Non-IBT schools, specify that IBT - Introduction to Basic Technology – is a skill training program]

B.13. When was the program established? ☐☐ (MM) ☐☐☐☐ (YYYY)

B.14. How was the school approached for launching the IBT program? ☐ (School approached LAHI on the HM's recommendation – 1; School approached LAHI on the school society's recommendation – 2; LAHI approached the school – 3; Other – 4; specify _____)

HEAD MASTER'S/RESPONDENT'S PERSONAL CHARACTERISTICS

B.15. Headmaster's/Headmistress's Full Name: .....

B.16. HM's Address: .....

B.17. HM's Phone number: (STD Code-Number) .....

B.18. HM's Email address .....

B.19. Religion: ☐ (Hindu – 1; Muslim – 2; Christian - 3; Sikh - 4; Other-5 (specify).....; Don't know-DK)

B.20. Caste: ☐ (SC – 1; ST – 2; OBC – 3; Brahmin-4; Other-5 (specify).....; Don't know-DK)

B.21. Highest education level completed: ☐ (Class Number: 1 – 12; Started College but did not complete – 13; Bachelor Degree – 14; Professional degree or diploma (including B.Ed., MBA, etc)– 15; Masters degree or higher – 16; Other – 17; No School – 18)

B.22. How many years have you been a HeadMaster (at this or any other school? ☐ (1-3 years – 1; 4-6 years – 2; 7-10 years – 3; More than 10 years - 4)

[More than one option can be selected, if relevant. Professional courses include B.Ed., MBA, etc. and Vocational courses include Diplomas or Certificates at ITI, ITC, Polytechnic, etc]

B.23. Monthly Income/Salary: ☐

B.24. If respondent is not the Headmaster, what is the respondent's position at the school? ……

Investigator's Notes:

Block C: School Staff & Teachers

C.1. Number of full-time teachers in the school:	Male	☐	Female	☐
C.2. Number of part-time teachers in the school:	Male	☐	Female	☐
C.3. Number of non-teaching staff:	Male	☐	Female	☐
C.4. Number of support staff:	Male	☐	Female	☐
C.5. Number of IBT Instructors:	Male	☐	Female	☐

C. Please provide educational details of any FOUR regular Teachers present at the school today:
[Note: You must ensure that the IBT coordinator is listed below]
[TID = Teacher ID]

C.6. TID	C.7. Teacher's Full Name	C.8. Address & Phone number	C.9. Is this teacher the IBT coordinator? (Yes – 1, No – 2, NA – 3) [For Non-IBT schools, choose NA option]	C.10. Highest class completed (Class Number: 1 – 12; Bachelor Degree – 13; Started College but did not complete – 14; Any other college/professional degree or diploma (including Masters)– 15; Other – 18; No School – 19)	C.11. If ever enrolled in college, what degree? (BA(pass) – 1; BA (honors) – 2; B.Sc. (pass) – 3; B.Sc. (honors) – 4; B. Com (pass) – 5; B.Com (honors) – 6; Masters (specify course)- 7; Vocational course – 8; Professional Course- 9; Other – 10; Not Applicable- NA)
01					
02					
03					
04					

C. Please provide employment details of the FOUR regular Teachers listed above:

C.12. TID	C.13. Teacher's Initials	Primary Job				Prior Employment (before IBT instruction)			
		C.14. Sector/ Function (Appendix A)	C.15. Start Date (mm/yy)	C.16. Hours worked per week	C.17. Monthly earnings	C.18. Sector/ Industry	C.19. Number of years of work experience	C.20. Hours worked per week	C.21. Monthly earnings
01									
02									
03									
04									

Block D: School Size

D.1. Number of classes/grades: [ ] to [ ] (For example, 1 to Jr. College/12)

D.2. Total number of students in the school: Male [ ] D.3. Female [ ]

D. Total number of students in the following classes:

	D.a. Class/Grade	D.b. Number of Divisions	D.c. Number of male students	D.d. Number of female students
D.4.	VII			
D.5.	VIII			
D.6.	IX			
D.7.	X			
D.8.	XI			
D.9.	XII			

D. Total number of students in IBT classes: ***[Note: For Non-IBT schools use NA option]***

	D.a. Class/Grade	D.b. Number of male students	D.c. Number of female students
D.10.	VIII		
D.11.	IX		
D.12.	X		

Block E: School Infrastructure

E.1. Number of classrooms in school building: …….

E.2 Number of staff and admin rooms in the school building: ……

E.3. Number of common or extra rooms in the school building: …….

E.4. Number of toilets: Only Girls ☐ Only Boys ☐ Boys & Girls ☐ Teachers/Staff ☐

E.5. Condition of toilet/s: for girls ☐ (Good – 1; Satisfactory – 2; Unsatisfactory – 3; Broken – 4; NA – 5)

E.6. Condition of toilet/s: for boys ☐ (Good – 1; Satisfactory – 2; Unsatisfactory – 3; Broken – 4; NA – 5)

E.7. Condition of toilet/s: common ☐ (Good – 1; Satisfactory – 2; Unsatisfactory – 3; Broken – 4; NA – 5)

E.8. Does the school have a computer laboratory? ☐ (Yes -1; No – 2) ***[Note: If NO, skip to E.10]***

E.9. How many working computers are there in the laboratory? ……. working computers.

E.10. Does the school have a public address system (loudspeaker)? ☐ (Yes -1; No – 2)

E.11. Does the school have other audio-visual equipment? (For example; LCD projector, sound system, etc) ☐ (Yes -1; No – 2)

E.12. Does the school have a playground? ☐ (Yes -1; No – 2)

E.13. Is there three-phase electricity available at the school? (Yes -1; No – 2)

E.14. Is water available at the school for growing plants, crops or maintaining a garden? ☐ (Yes -1; No – 2)
[Note: If No, skip to E.16.]

E.15. Specify the source of water: ☐ (Pipe – 1; Well – 2; Tubewell – 3; Canal – 4; River – 5; Other – 6)

E.16. Is drinking water available at the school? ☐ (Yes -1; No – 2)

E.17. Does the school own any cultivable land? ☐ (Yes -1; No – 2) ***[Note: If NO, skip to E.19]***

E.18. How many acres of cultivable land does the school own? ……. acres.

E.19. What is the probability of the village or community providing the school with one acre of cultivable land for student projects?

☐ (High – 1; Medium – 2; Low – 3; Don't Know – DK)

E.20. How far is the school from the main village? …… kms.

E.21. Does the village hold a weekly bazaar? ☐ (Yes -1; No – 2)

E.22.a. Does the school offer students any extra-curricular training or skill training program? ☐ (Yes -1; No – 2)

[Note: If NO, skip to Block F]

E.22.b. What is the name of the program? ……

E.23. Which students does the program target? ☐ (Dropouts – 1; Absentee/Irregular students – 2; Disadvantaged students - 3; Academically-behind students – 4; Regular students - 5; Other - 6)

E.24. Which grades/classes does the program target? ☐ to ☐

E.25. Which option best describes this program? ☐ (Remedial Education – 1; Supplementary academic training – 2; Skill training – 3; Life Skills – 4; Other – 5)

E.26. Who provides the training or teaches the students as part of this program? ☐

(Regular teachers at the school – 1; Principal and/or Vice Principal – 2; Para teachers at the school – 3; Community members hired/trained for this purpose – 4; Other individuals hired/trained for this purpose – 5; Other – 6: specify ……..)

E.27. When are students trained or taught as part of this program? ☐ (During the school day – 1; After the school day – 2; Before the school day – 3; On weekends/holidays – 4; Other – 5)

E. 28. Are the students tested as part of this program? ☐ (Yes -1; No – 2)

E.29. Does the school have any equipment/tools/materials to conduct this program? ☐ (Yes -1; No – 2)

E.30. List equipment available: ……

Block F: Attendance & Academic Achievement

F. What was the average performance on class X board exams for the following years:

	a. Year	b. Total number of male students enrolled in class X	c. Total number of female students enrolled in class X	d. Number of male students who took the SSC exam	e. Number of female students who took the SSC exam	f. Number of male students who passed the SSC exam	g. Number of female students who passed the SSC exam	h. Average Percentage Marks
F.1	2003							
F.2	2004							
F.3	2005							
F.4	2006							
F.5	2007							
F.6	2008							
F.7	2009							
F.8	2010							

F. Please provide information on dropouts for the following classes/grades in the year 2005:

	a. Class	b. Number of boys enrolled at the beginning of the year	c. Number of girls enrolled at the beginning of the year	d. Total number of boys who took the end-of-year exam/board exam	e. Total number of girls who took the end-of-year exam/board exam
F.9	VII				
F.10	VIII				
F.11	IX				
F.12	X				

F. Please provide information on dropouts for the following classes/grades in the year 2006:

	a. Class	b. Number of boys enrolled at the beginning of the year	c. Number of girls enrolled at the beginning of the year	d. Total number of boys who took the end-of-year exam/board exam	e. Total number of girls who took the end-of-year exam/board exam
F.13	VII				
F.14	VIII				
F.15	IX				
F.16	X				

F. Please provide information on dropouts for the following classes/grades in the year 2009:

	a. Class	b. Number of boys enrolled at the beginning of the year	c. Number of girls enrolled at the beginning of the year	d. Total number of boys who took the end-of-year exam/board exam	e. Total number of girls who took the end-of-year exam/board exam
F.17	VII				
F.18	VIII				
F.19	IX				
F.20	X				

F. Please provide information on dropouts for the following classes/grades in the year 2012:

	a. Class	b. Number of boys enrolled at the beginning of the year	c. Number of girls enrolled at the beginning of the year	d. Total number of boys who took the end-of-year exam/board exam	e. Total number of girls who took the end-of-year exam/board exam
F.21	VII				
F.22	VIII				
F.23	IX				
F.24	X				

Block G: School Community

G.1. Population of the village: ……

G.2. Major occupation in the village: ☐ (Agriculture – 1; Fisheries – 2; Animal Husbandry – 3; Other – 4, specify _____)

[Note, If not Agriculture, then skip to G.4]

G.3. List the major crops grown: ……

G.4. Name of major town/city closest to the village: ……

G.5. Distance to closest major town/city: ……. Kms.

G.6. Is the town/village accessible by public transport? ☐ (Yes -1; No – 2)
[Note: If NO, skip to G.8.]

G.7. Specify what type of public transport is available: ☐ (ST Bus – 1; Rail – 2; Air -3; Other – 4, Specify _____)

G.8. How far is the nearest bus stand/railway station from the village? …… kms.

G.9. How many villages does the school serve? ……..

G.10. List the names of the villages: ……

Block H: Principal Perceptions

H.1. Since when have you been the head of this school? ☐☐ MM ☐☐☐☐ YYYY

[Note: Skip H.2. for Non-IBT schools – USE 'NA']

[Note: In IBT schools if the respondent became Headmaster after 2006, skip this question and use 'NA' code]

H.2. How did you feel about the IBT program when it was first proposed? ☐ (Choose the option that best describes how you felt at that time: I was eager to launch it – 1; I was concerned about costs/other operational aspects of the program – 2; I was not interested – 3; NA – 4)

H. For the following statements, select 'Agree' 'Disagree' or 'Maybe': (Agree – 1; Disagree – 2; Maybe – 3)

H.4. Skill training is for students who are not interested in or cannot pursue education beyond high school. ☐

H.5. Skill training should be introduced at the secondary school level. ☐

H.6. Skill training should be introduced at the postsecondary level. ☐

H.7. Skill training provides the same returns as a graduate degree. ☐

H.8. Skill training provides better returns when coupled with a graduate degree. ☐

[Note: Provide the following explanations to Headmasters at NON-IBT schools]

"IBT or Introduction to Basic Technology is a vocational/skill training program offered at some secondary schools in your district. The program uses theory and practical classes to give students basic skills and knowledge in 4 areas – Agriculture, Energy & Environment, Home & Health, and Welding. Local practitioners or entrepreneurs from the community who have these skills are trained to teach students at the school. Students are expected to complete projects for the school and community as part of their training. The program is 3 years long and begins in the 8th standard. The one-time cost of setting up the program is Rs.40,000 and it costs Rs.4,000 per year to maintain."

H.9. A program like IBT is a good example of experiential learning. ☐

H.10. A program like IBT can motivate participating students to attend school regularly. ☐

H.11. A program like IBT can improve engagement in school and school activities among participating students. ☐

H.12. A program like IBT is not worth the cost. ☐

---------- End of survey ----------

Student Survey Instrument

Informed Consent Agreement for the Vocational Education Student Survey

I would like to invite you to participate in the present research study. The purpose of this study is to estimate the impacts of vocational education at the secondary school level on school completion, postsecondary enrollment, and employment outcomes among participants. Fifty two (52) schools across the state of Maharashtra and students who attended these schools have been invited to participate in this study.

Your participation will entail answering a series of questions. This will include questions on your background, the education and employment details of your family members, your educational history, and your employment experiences. This survey will take approximately 45 minutes to complete. Participation in the survey is completely voluntary. If I ask any question you do not wish to answer let me know and I will go to the next question; or you can choose to discontinue the interview. There will be no consequences to refusing to answer any question.

The data gathered as part of this research process will be de-identified once it has been converted to electronic format. All names and other identifiers will be removed from the data and will not be used or appear in any analysis or research report.

There are no known risks associated with your participation in this research beyond those of everyday life. You will/not be compensated for participating in this interview. Your participation will help the research since your views are important to help us understand school-based vocational education programs.

If there is anything about the study or your participation that is unclear or that you do not understand, or if you wish to report a research–related problem, you may contact Namrata Tognatta at phone: **020-25884180** or email namratat@gse.upenn.edu.

At this time, do you have any questions about the survey? Do you agree to participate in this survey?

☐ Yes ☐ No

_____

Signature of Respondent and Date

_____

Signature of Witness and Date

Taluka _____, District _____, Maharashtra

Individual and Household Questionnaire

Block A. Identification

(Verify that this is complete before you leave for the field.)

A.1.Respondent Id: | | | | | | | | | | | | | | |

A.2. Name: .....

A.3. Address: .....
.....
.....

Pin Code | | | | | | | |

A.4. Mobile phone no: | | | | | | | | | | |

A.5. Landline number:

A.6. Date of survey: | | | | | | | | |

(DD) (MM) (YYYY)

A.7. Time of survey: Start _____ End _____ (Hours/Minutes)

A.8. Signature of surveyor: .....

A.9. Signature of supervisor: .....

A.10. Code for interviewer's result: | | (Interview conducted - 1; Refused - 2; No interview conducted – 3)

A.11. If 3 in A.6: specify reason for no interview: ☐ (Cannot locate residence - 1; Moved to a new village/town - 2; Currently not at home - 3; Other – 4: specify _____)

[If the respondent has moved to a different house within or outside the village, try to obtain information on the respondent's new location. If the respondent is currently not at home, call the respondent to get an appointment for the interview. Use the space below to make notes]

Protocol: If option 3 in A.9, then this questionnaire goes to a field supervisor

Notes for interviewer: The "respondent" must be the person identified for interview purposes.
Household *is a person or group of persons who occupy a part of or an entire building and who usually live together and eat from the same kitchen. A* ***Householder*** *is anyone who usually lives in the household, whether she/he is at home during the survey or is temporarily absent. A householder who has been away for 6 or more months is not regarded as a householder. A guest who has stayed in the household for 6 or more months (continuously in the last 6 months) is regarded as a householder. The* ***head of the household*** *is a person who is regarded/assigned as the head of the household.*

Notes:

Block B: Household information ***[Note for interviewer: respondent should be the participant whose name you have in your list]***

B.1. Individual's Full Name: .....

B.2. Father's/Husband's Full Name: .....

B.3. Religion: ☐ (Hindu – 1; Muslim – 2; Christian - 3; Sikh - 4; Other-5 (specify) .....; Don't know-DK)

B.4. Caste ☐ (Dalit – 1; Adivasi – 2; OBC – 3; Brahmin-4; Other-5 (specify).....; Don't know-DK)

B.5. Father's/Husband's occupation: ☐☐ **Sector | Function (Refer to Key)**

Block C: Household Characteristics

Please answer the following questions about all the members of your household. Members of your household include all individuals who live here on a regular basis.

<u>MAKE SURE</u> THAT PID NUMBER 01 IS FOR THE RESPONDENT I.E. THE STUDENT WHO'S NAME APPEARS IN YOUR LIST

C.1. PID Number	**C.2. Full Name**	**C.3. Father's Full Name**	**C.4. Mother's Full Name**	**C.5. Age in completed years** *(0 if less than 1. DK if Don't know)*	**C.6. Gender** Male – 1; Female – 2	**C.7. Relationship with Head of Household Head** – 1; Wife/Husband of Head – 2; Son/daughter– 3; Grandson/granddaughter – 4; Father or Mother – 5 ; Sister or Brother – 6; Niece or Nephew – 7; Son/Daughter-in-Law – 8; Brother/Sister-in-Law – 9; Father/Mother-in-Law – 10; Other Family Member/Relative–11; Servant/Servant's Relative – 12; Tenant/Tenant's Relative – 13; Other Person not related – 14.	**C.8. Marital Status** Married – 1; Divorced – 2; Separated – 3; Widow/Widower – 4; Never Married – 5;	**C.9 Currently enrolled in School or college?** School – 1; College – 2; Neither – 3; Don't know – 4.	**C.10. Highest Class Completed/Class currently enrolled in** Class Number: 1 – 12; Bachelor Degree – 13; Started College but did not complete – 14 Any other college/professional degree or diploma (including Masters)– 15; Other – 18; No School – 19;	**C.11. If ever enrolled in college, what degree? ***** BA – 1; B.Sc. – 2; B.Com – 3; Masters (specify course)-4 Vocational course (ITI, ITC, Polytechnic, etc) – 5; Professional Course – 6; Other – 7; Not Applicable- NA *(more than one option can be selected, if relevant)*
01										
02										
03										
04										
05										
06										
07										
08										

[Note: Ensure that all members listed as Fathers and Mothers are included as household members unless not living in the household
*** Examples of Professional course: B.Ed., CA, CS, CFA, MBA
Examples of Vocational Courses: Sewing, Carpentry, Diploma or Certificate courses, etc

For C.15. If an individual is employed as a manager in a hospital then the sector is 02 and function is MNGR. If he is a tuition teacher then his sector is 03 and function is OWNR.
For C.16. Include part-time work as well as work done on weekends. Only that work for which payment is made must be considered.

C.12. PID Number *(Copy from the previous page)*	**C. 13. What was the medium of instruction in your school/college?** Marathi – 1; English – 2; Hindi – 3; Urdu – 4; Other language (Specify) – 5;	**C.14. Did [....] work in the last 1 month? (Only consider work for which a payment was made)** Yes – 1; No – 2; Don't know -DK. *[If Yes, continue to C.15, else skip to next section]*	**C.15. What is the primary occupation of [....]?: Sector | Function (Refer to Appendix A)**	**C.16. How many hours did [..] spend working as C.15 in the past one week?**	**C.17. How much did [....] earn during the last one month?** *[in Rupees]*	**C.18. In 2012 how many months did [....] work?** < 1 month - 1; One to 3 months – 2; Three to 6 months – 3; More than 6 months in a year – 4; Don't know – DK.
01 **(Respondent)**						
02						
03						
04						
05						
06						
07						
08						
09						

[Note: The next set of questions is for female, married respondents only – ASSIGN THE SAME PID AS THAT USED ABOVE]

Please answer the following questions regarding your spouse:

C.19. PID Number	**C.20. Full Name**	**C.21. Age in completed years** *(0 if less than 1. DK if Don't know)*	**C.22. Currently enrolled in School or college?** School – 1; College – 2; Neither – 3; Don't know – 4.	**C.23. Highest Class Completed/Class currently enrolled in** Class Number: 1 – 12; Bachelor Degree – 13; Started College but did not complete – 14 Any other college/professional degree or diploma (including Masters)– 15; Other – 18; No School – 19;	**C.24. If ever enrolled in college, what degree?** BA – 1; B.Sc. – 2; B.Com – 3; Masters (specify course)-4 Vocational course (ITI, ITC, Polytechnic, etc) – 5; Professional Course – 6; Other – 7; Not Applicable- NA *(more than one option can be selected, if relevant)*	**C.25. Did [....] work in the last 1 month? (Only consider work for which a payment was made)** Yes – 1; No – 2; Don't know -DK. */If Yes, continue to C.27, else skip to Block D.]*	**C.26. What is the primary occupation of [....]? Sector | Function (Refer to Appendix A)**	**C.27. How much did [....] earn during the last one month?** *[in Rupees]*	**C.28. In 2012 how many months did [....] work?** < 1 month - 1; One to 3 months – 2; Three to 6 months – 3; More than 6 months in a year – 4; Don't know – DK.
(Respondent's spouse)									

Block D: House/Dwelling

D.1 For how many years have you lived in this house?.....

D.2 How many rooms in your house? (Excluding bathroom, toilet and verandah).....

D.3. Is the house *kuchha* or *pukka*? ☐ (Pukka – 1; Kuchha – 2; Don't know - DK)

D.4. Connected to electricity? ☐ (Yes – 1; No – 2; Don't know - DK)

[Note: If Yes, continue to D.5, else skip to D.6]

D.5. On average how many hours did you not have electricity during the last week? …… hours

D.6. Toilet in House? ☐ (Yes – 1; No – 2)

D.7. Main Source of Drinking Water? ☐ (Pipe – 1; Well – 2; Tube well – 3; Other (Specify) – 4; Don't know – DK .....)

D.8. Separate Kitchen? ☐ (Yes – 1; No – 2)

NOTES

Block E: Consumer Durables

Does the household contain of the following consumer durables?
[Note: Go through the list below with the respondent and obtain details for each item.]

E.a. Identification Number	**E.b, Description**	**E.c. Quantity owned** *[Write zero if not owned. If zero, skip to next item.]*	**E.d. Did you purchase during the last 6 months?** Yes – 1; No – 2; Don't know – DK. *[Write NA if zero in the previous column.]*	**E.e. Value of the good purchased (in Rupees)** *[Write NA if not purchased in the last 6 months. DK if Don't know.]*
1.	Radio/Tape Recorder			
2.	TV/ Cable TV/Satellite TV/Dish TV			
3.	VCR/VCD/DVD Player			
4.	Computer/Laptop			
5.	Cycle			
6.	Motor Cycle/Moped/Scooter			
7.	Car/Jeep/Truck/Other 4 wheeler			
8.	Refrigerator			
9.	Fan			
10.	Cooler/ Air conditioner			
11.	Kerosene Stove/Gas Stove			
12.	Kerosene Lamp			
13.	Landline Telephone			
14.	Mobile telephone			
15.	Sewing Machine			

Block F: Educational History

F.1. Did your secondary school (Class 8th and above) provide any supplementary academic or skill training? ☐ (Yes – 1; No – 2)
[Note: If No, go to F.7.]

F.2. Which option best describes the training offered? ☐ (Remedial Education – 1; Supplementary academic training – 2; Skill training – 3; Other – 4, specify _____)

F.3. What was the training program called? _____

F.4. Did you participate in this training in class VIII? ☐ (Yes – 1; No – 2)

F.5. Did you participate in this training in class IX? ☐ (Yes – 1; No – 2)

F.6. Did you participate in this training in class X? ☐ (Yes – 1; No – 2)

[Note: For questions F.7-F.10, use letter grades (A, B, C, etc) if students indicate that they received a letter grade instead of overall percentage. Or if student received overall percentage and letter grade, fill in both in each box.]

F.7. What was your overall score in the 5th grade? ☐ (Above 85% - 1; 71-85% - 2; 61-70% - 3; 51-60% - 4; Below 50% - 5; Don't know – DK)

F.8. What was your overall score in the 6th grade? (Above 85% - 1; 71-85% - 2; 61-70% - 3; 51-60% - 4; Below 50% - 5; Don't know – DK)

F.9. What was your overall score in the 7th grade? ☐ (Above 85% - 1; 71-85% - 2; 61-70% - 3; 51-60% - 4; Below 50% - 5; Don't know – DK)

F.10. What was your overall score in the 8th grade? ☐ (Above 85% - 1; 71-85% - 2; 61-70% - 3; 51-60% - 4; Below 50% - 5; Don't know – DK)

F.11. What percentage marks did you score on the SSC board exam? ……. ***[Note: If respondent has not taken the SSC Board exam, mark as 'NA']***

F.12. What percentage marks did you score on the HSC board exam? ……. ***[Note: If respondent has not taken the HSC Board exam, mark as 'NA']***

F.13. Did you ever fail a grade/class while in school? ☐ (Yes – 1; No – 2) ***[Note: If NO, skip F.14]***

F.14. What grade/class did you fail and how many times did you repeat that grade/class? ☐☐ (Grade/No. of times repeated)

F.15. Are you currently enrolled in a program of study? ☐ (Yes – 1; No – 2) ***[Note: If No, skip to F.25.]***

F.16. Are you studying full-time or part-time? ☐ (Full-time – 1; Part-time – 2)

F.17. When did you enroll in your current program of study? ☐☐ ☐☐ (MM/YY)

F.18. What is your current program of study? ☐ (Higher Secondary School/Junior College in Arts – 1; Higher Secondary School/Junior College in Commerce – 2; Higher Secondary School/Junior College in Science – 3; Vocational Course – 4; B.A. – 5; B.Com – 6; B.Sc – 7; Professional Course – 8; Other – 9, specify _____)

[Note, Items F.19 –22 are for those who respond (4) Vocational course in F.18]

F.19. Are you enrolled at a private or government vocational or technical institute? ☐ (Government – 1; Private – 2; Don't Know – DK; Other – 4, specify _____ _____)

F.20. What type of institute do you currently attend? ☐ (Industrial Training Institutes (ITIs)/Industrial Training centers (ITCs) -01, School offering vocational courses (Secondary, Higher Secondary level) -02, UGC (first degree level) -03, Polytechnics -04, Community Polytechnics/ Janshiksha Sanstha -05, National Open School -06, Hotel Management Institutes -07, Food craft & Catering Institutes -08, Small Industries Service Institutes/District Industries Centres/Tool Room Centres -09, Fashion Technology Institutes -10, Tailoring, Embroidery and Stitch Craft Institutes -11, Nursing Institutes -12, Rehabilation/ Physiotheraphy /Ophthalmic and Dental Institutes -13, Institutes giving Diploma in Pharmacy -14, Hospital and Medical Training Institutes -15, Nursery Teachers' Training Institutes -16, Institutes offering training for Agricultural Extension -17, Training provided by Carpet Weaving Centers -18, Handloom/ Handicraft Design Training Centers/ KVIC -19, Recognised Motor Driving Schools -20, Institute for Secretariat Practices -21, Recognised Beautician Schools -22, Institutes run by Companies/ Corporations -23, Institutes for Journalism and Mass Communication -24, other institutes -99)

F.21. What is the duration of your program (in months)? ……

F.22. What is your field of training? ☐ (Mechanical engineering trades -01, electrical and electronic engineering trades -02, computer trades -03, civil engineering and building construction related works -04, chemical engineering trades -05, leather related work -06, textile related work -07, catering, nutrition, hotels and restaurant related work -08,

artisan/ craftsman/handicraft and cottage based production work -09, creative arts/ artists -10, agriculture and crop production related skills and food preservation related work -11, non-crop based agricultural and other related activities -12, health and paramedical services related work -13, office and business related work -14, driving and motor mechanic work -15, beautician, hairdressing & related work -16, work related to tour operators/travel managers -17, photography and related work -18, work related to childcare, nutrition, pre-schools and crèche -19, journalism, mass communication and media related work -20, printing technology related work -21, other -99)

F.23. How many effective hours did you attend your current educational institution last week or the last week the institute was in session? ☐☐

F.24. Are you working while attending your current institution? ☐ (Yes – 1; No – 2)

F.25. What is your reason for not being enrolled in a program of study? ☐ (Not interested in studying – 1; Illness – 2; Sought/Seeking employment – 3; Marriage – 4; Low marks/percentage – 5; Could not afford fees – 6; Parents/household members did not approve – 7; Distance was too far – 8; Did not want to put in hard work – 9; Household responsibilities – 10; Other – 11, specify _____ _____)

F.26. What was the last level of education you were enrolled in? ☐ (Lower secondary (VIII-X) – 1; Higher Secondary (XI-XII) – 2; College/Diploma – 3; Other – 4)

F.27. Did you complete the last level of education you were enrolled in? ☐ (Yes – 1; No – 2) ***[Note: If YES, skip F.28]***

F.28. Why did you leave this level of education? ☐ (Not interested in studying – 1; Illness – 2; Sought/Seeking employment – 3; Marriage – 4; Low marks/percentage – 5; Could not afford fees – 6; Parents/household members did not approve – 7; Distance was too far – 8; Did not want to put in hard work – 9; Household responsibilities – 10; Other – 11, specify _____)

Block G: Employment History

G. Were you engaged in any of these activities in the past week?

G.1. Working/Trying to work: ☐ (Yes – 1; No – 2)

G.2. Job Search: ☐ (Yes – 1; No – 2)

G.3. Housekeeping: NA (Yes – 1; No – 2)

G.4. Have you had casual employment in the last month? ☐ (Yes – 1; No – 2)
[Note: If No in G.4, continue to G.14]

G.5. How many jobs? Describe each job. ***[Note: Request respondent to provide details on what he/she does as part of his/her job]***
a. Description of job 1: ……

b. Description for job 2: ……

c. Description of job 3: ……

List occupation code for each job: ☐☐ ☐☐ ☐☐ **Sector | Function (Refer to Appendix A)**

d.Job 1 e.Job 2 f.Job 3

Ask for the 3 most important jobs if more than one job.

		a. Job 1 ( )	**b. Job 2 ( )**	**c. Job 3 ( )**
G.6.	When did you secure this job? (dd/mm/yy)			
G.7.	How many days were you engaged in job [...] in the past 30 days? (Number of days)			
G.8.	Are you still engaged in job […] (Yes – 1; No – 2)			
G.9.	How were you paid for your job? [Cash – 1; Kind – 2]			
G.10.	If you were paid in cash: how much did you get in cash per day? (Rs. per day)			
G.11.	If you were paid in kind what did you receive, describe?			
G.12.	Quantity and units of what you received in G.11.			
G.13.	Approximate Value (in Rs) of the items received in G.11			

[Note: If the individual has the same job in 3 different locations (for example, a household helper working as tutor in 3 households), it should be coded as only ONE job.]

G.14. Have you been a permanent employee in the last 30 days? ☐ (Yes – 1; No – 2) ***[Note: If No in G.7 skip to G.25.]***

G.15. How many weeks have you been a permanent employee? .....

G.16. How many jobs? Describe each job. ***[Note: Request respondent to provide some details on what he/she does as part of his/her job]***

a. Description of job 1: ……

b. Description for job 2: ……

c. Description of job 3: ……

List occupation code for each: **Sector | Function (Refer to Appendix A)**

☐☐ d.Job1 ☐☐ e.Job2 ☐☐ f.Job3

Ask for the 3 most important jobs if more than one job.

		a. Job 1 ( )	**b. Job 2 ( )**	**c. Job 3 ( )**
G.17.	When did you secure this job? (dd/mm/yy)			
G.18.	How many days were you engaged in job [...]in the past 30 days? (Days)			
G.19.	Are you still engaged in job […] (Yes – 1; No – 2)			
G.20.	How were you paid for your job? [Cash – 1; Kind – 2]			
G.21.	If you were paid in cash: how much did you get in cash per day? (Rs per day)			
G.22.	If you were paid in kind what did you receive, describe?			
G.23.	Quantity and units of what you received in G.22.			
G.24.	Approximate Value (in Rs) of the items received in G.22			

G.25. Have you earned any income from self-employment (for example business) in the past 30 days? ☐ (Yes – 1; No – 2) ***[Note: If No, skip to G.30.]***

G.26. How many months have you been self-employed?.....(Note: if less than 1 month, write down no. of weeks since self-employed)

G.27. Type of Business? | | | **Sector | Function (Refer to Appendix A)**

G.28. Gross Revenue in the last 30 days? .....

G.29. Net Revenue in the last 30 days? .....

G.30. Total costs incurred towards your business in the last 30 days? ……

G.31. Net Profit gained in the last 30 days after taking out all business costs? ……

We would like to get some more details on your job and employment

G.32. How many hours did you spend working last week? .....

G.33. Did you look for more work in the last week? | | (Yes – 1; No – 2) ***[Note: If Yes, skip to G.33.]***

G.34. Why did you <u>not</u> look for work? | | (Already sufficiently employed – 1; No jobs or work available – 2; Physically/Mentally disabled – 3; Housewife/child rearing – 4; Student – 5; No skills to get a job – 6; Preparing for competitive exams-8; Other – 9, specify _____)

G.35. How many hours did you spend looking for a job in the last week? .....

G.36. How did you search for jobs? | | (Knocking on doors – 1; Looking up advertisements – 2; Calling friends and relatives – 3; Internet based job sites – 4; Employment Exchange – 5; Other – 5, specify _____)

G.37. How much money did you spend looking for a job in the last week? .....

G.38. What kinds of a job were you looking for? | | | **Sector | Function (Refer to Appendix A)**

---------- END OF SURVEY ----------

Sector of Employment	**Professional Function**
Government – 01	PROF- Professional Technical and Kindred Workers
Healthcare/Hospital – 02	MNGR- Administrative, Executive and Managerial
Education – 03	CLER- Sales and Clerical
Banking/Insurance/Finance - 04	CRAFT- Craft and Kindred Workers
Administrative and professional - 05	OPER- Production Workers and Transport Operatives
Hospitality, Tourism & Restaurant - 06	SERV- Service Workers and Labourer
Retail – 07	OWNR - Owner/Proprietor/Self-employed
Construction and Real Estate - 08	UNEM – Unemployed
Electronics/IT/ITES/Telecom - 09	OTHER- (specify everywhere)
Energy, Manufacturing, Production & Operations – 10	
Transportation – 11 Other-12 (specify everywhere)	

(Note: *Make sure to specify what "other" is in all parts where question related to job sector and function was asked*)

References

Adams, A.V. (2007). *The role of youth skills development in the transition to work: A global review*. Report No. 5, Human Development Network Children and Youth Department. Washington DC: The World Bank.

Adams, A.V.; Johansson de Silva, S. & Razmara, S. (2013). *Improving skills development in the informal sector: Strategies for Sub-Saharan Africa*. Directions in Development – Human Development, Washington DC: The World Bank.

Agarwal, B. (1994). *A Field of One's Own: Gender and Land Rights in South Asia*. Cambridge, UK: Cambridge University Press.

Agodini, R.; Uhl, S.; Novak, T. (2004). *Factors that influence participation in secondary vocational education*. MPR Reference No. 8879-400. NJ: Mathematic Policy Research. Retrieved from http://www.mathematica-mpr.com/Publications/PDFs/vocationalfactors.pdf

Ainley, J. & Corrigan, M. (2005). Participation in and progress through new apprenticeships. Longitudinal Surveys of Australian Youth, Research Report, No. 44. Camberwell, VI: Australian Council for Educational Research. Retrieved from http://research.acer.au/lsay_research/48

Allison, P.D. (1999). Multiple Regress: A Primer. Thousand Oaks, CA: Pine Forge Press.

Aggarwal, M.; Kapur, D.; Tognatta, N. (2011). The Skills they Want: Aspirations of Students in Emerging India. CASI Working Paper Series, n. 12-03. Philadelphia, PA: Center for the Advanced Study of India, University of Pennsylvania. Retrieved from http://casi.sas.upenn.edu/system/files/The+Skills+They+Want,+Aspirations+of+Students+in+Emerging+India+-+Aggarwal,+Kapur,+Tognatta_0.pdf

Agrawal, T. (2012). Vocational education and training in India: Challenges, status and labour market outcomes. *Journal of Vocational Education and Training, 64*(4), 453-474.

Amin, M. (2008). *Competition and Demographics*. Policy Research Working Paper, Enterprise Analysis Unit. Washington DC: The World Bank.

Aypay, A. (2003). The tough choice at high school door: An investigation of the factors that lead students to general or vocational schools. *International Journal of Educational Development, 23*, 517-527.

Azam, M.; Chin, A.; Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change, 61*(2), 335-367.

Basu, A.M. (1992). Culture, the Status of Women, and Demographic Behaviour. Oxford, UK: Clarendon Press.

Behrman, J.R. (1999). Schooling in Asia: Selected microevidence on determinants, effects and policy implications. *Journal of Asian Economics, 10*, 147-194.

Behrman, J.R. (2010). Investment in education: Inputs and incentives. *Handbook of Development Economics, 5*, 4483-4975.

Behrman, J.R. & Knowles, J.C. (1997). *How Strongly is Child Schooling Associated with Household Income*. PIER Working Paper, n.97-022. Philadelphia, PA: Penn Institute for Economic Research, University of Pennsylvania.

Behrman, J.R. & Knowles, J.C. (1999). Household income and child schooling in Vietnam. *The World Bank Economic Review, 13*(2), 211-256.

Behrman, J.R. & Rosenweig, M.R. (2002). Does increasing women's schooling raise the schooling of the next generation? *American Economic Review, 92*(1), 323-334.

Behrman, J.R. & Wolfe, B.L. (1987). Investments in schooling in two generations in pre-revolutionary Nicaragua: The roles of family background and school supply. *Journal of Development Economics, 27*, 395-419.

Benavot, A. (1983). The rise and decline of vocational education. *Sociology of Education, 56*(2), 63-76.

Becker, G.S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy 70(*5), 9–49.

Becker, G.S. (1993). *Human capital: A theoretical and empirical analysis with special reference to education* (3rd edition). Chicago: University of Chicago Press.

Birdsall, N. (1982). *Child Schooling and the Measurement of Living Standards*. LSMS Working Paper, n.14; Washington D.C: The World Bank.

Birdsall, N. (1985). Public inputs and child schooling in Brazil. *Journal of Development Economics, 18(*1), 67-86.

Bishop, J. H. & Mane. F. (2005). Economic returns to vocational courses in U.S. high schools," in J. Lauglo and R. Maclean (Eds.) *Vocationalisation of Secondary Education Revisited*. Netherlands, Springer, 329-362 p.

Boissiere, M. (2004). *Determinants of Primary Education Outcomes in Developing*

Countries. World Bank Operations Evaluation Department. Washington DC: The World Bank. Retrieved from http://www.worldbank.org/oed

Borghans, L.; De Grip, A.; Heijke, H. (1996). Labor market information and the choice of vocational specialization. *Economics of Education Review, 15*(1), 59-74.

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp 241–258). Westport, CT: Greenwood Press.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica, 69*(5), 1127-60.

Chandrashekhar, S. & Mukhopadhyay, A. (2006). Primary education as a fundamental right: Cost implications. *Economic and Political Weekly, 41*(35), 3797-3804.

Chappell, C. (2003). Researching vocational education and training: Where to from here? *Journal of Vocational Education and Training, 55*(1), 21-32.

Chaudhury, N., Christiansen, L. & Asadullah, M. (2006). *School, Household Risks and Gender: Determinants of Child Schooling in Ethiopia*. World Bank Working Paper, Washington DC: The World Bank.

Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94* (Issue Supplement), S95–S120.

Connelly, R. & Zheng, Z. (2003). Determinants of school enrollment and completion of 10 to 18 year olds in China. *Economics of Education Review, 22*(4), 379-388.

Corman, H. & Davidson, P. (1984). Economic aspects of postsecondary schooling decisions. *Economics of Education Review, 3*(1), 131-139.

Curtis, D.D. (2008). *VET pathways taken by school leavers.* LSAY Research Report No. 52. Melbourne: ACER

Dika, S.L., and Singh, K. (2002). Applications of social capital in educational literature: A critical synthesis. *Review of Educational Research 72*(), 31–60.

Deninger, K. (2003). Does cost of schooling affect enrollment by the poor? Universal primary education in Uganda. *Economics of Education Review, 22*(3), 291-305.

Deolalikar, A.B. (1993). Differences in the returns to schooling and in school enrollment rates in Indonesia. *Journal of Human Resources, 28*(4), 899-932.

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

Drèze, J. & Kingdon, G. (2001). School participation in rural India. *Review of Development Studies, 5*(1), 1-24.

Drèze, J. & Sen, A. (1995). India: Economic Development and Social Opportunity. Oxford: Oxford University Press.

Duraisamy, P. (2000). Changes in returns to education in India, 1983-1994: By gender, age-cohort and location. *Economics of Education Review, 21*(6), 609-622.

Filmer, D. & Pritchett, L.H. (2001). Estimating wealth effects without expenditure data or tears: An application to educational enrollments in states of India. *Demography, 38*(1), 115-132.

Foley, P. (2007). *The socio-economic status of vocational education and training students in Australia*. Adelaide, SA: National Center for Vocational Education Research.

Fullarton, S. (2001). *VET in schools: Participation and pathways*. Longitudinal Surveys of Australian Youth, Research Report No. 21. Camberwell, VI: Australian Council for Educational Research.

Glewwe, P. & Jacoby, H. (1994). Student achievement and school choice in low-income countries: Evidence from Ghana. Journal of Human Resources, 29(3), 843-864.

Glewwe, P. & Jacoby, H. (1995). An economic analysis of delayed primary school enrollment in a low income country: The role of early childhood nutrition. *The Review of Economics and Statistics, 77,* 156-169.

Grubb, W.N. (1988). Vocationalizing higher education: The causes of enrollment and completion in public two-year colleges, 1970-1980. *Economics of Education Review, 7*(3), 301-319.

Grubb, W.N. (2006). *Vocational Education and Training: Issues for a Thematic Review*. Paris: OECD.

Grubb, W.N. & Lazerson, M. (2004). *The Education Gospel: The Economic Value of Schooling*. Cambridge: Harvard University Press.

Grubb, W.N. & Sweet, R. (2004). *Alternatives to Universities Reconsidered*. Education Policy Analysis, 2004. Paris: OECD.

Hansen, B.B. (2004). Full matching in an observational study of coaching for the SAT. *Journal of the American Statistical Association, 99*, 609-18.

Hanushek, E.A. (2008). *The economic value of education and cognitive skills*. Retrieved on May 17, 2012 from http://psychologyandeducation.edublogs.org/files/2010/11/3_AERA_Hanushek_Proof-mkgp2r.pdf

Hanushek, E.A. (1995). Interpreting recent research on schooling in developing countries. *The World Bank Research Observer, 10*(2), 227-246.

Hanushek, E.A., Woessmann, L. & Zhang, L. (2011). *General education, vocational education and labor market outcomes over the life-cycle*. Working paper 17504, Cambridge, MA: National Bureau of Economic Research.

Hanushek, E.A. & Zhang, L. (2008). *Quality-consistent estimates of international returns to skill [mimeo]*. Stanford, CA: Stanford University, Hoover Institution

Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153-61.

Ho, D.; Imai, K; King, G.; Stuart, E.A. (2006). Matchit: Nonparametric Preprocessing for Parametric Casual Inference (Version 2.2-11) [Software]. Available at http://gking.harvard.edu/matchit

Ho, D.; Imai, K; King, G.; Stuart, E.A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*(3), 199-236.

Horvat, E.M. (2001). Understanding equity and access in higher education: The potential contribution of Pierre Bourdieu. In J.C. Smart (ed.), *Higher Education: Handbook of Theory and Research, 16*, 158–171). New York: Agathon Press.

Hossler, D., Schmit, J., and Vesper, N. (1999). *Going to College: How Social, Economic, and Educational Factors Influence the Decisions Students Make*. Baltimore: Johns Hopkins University Press.

International Labour Organization (ILO). (2003). *Industrial Training Institutes of India: The Efficiency Study Report*. Geneva: ILO.

Iyengar, R. (2012). Social capital as the catalyst for school participation. *Compare: A Journal of Comparative and International Education, 42*(6), 839-862.

Jeffrey, R. & Jeffrey, P. (1996). Don't marry me to a plowman: Women's everyday lives in rural north India. New Delhi: Sage Publications.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics, 125*(2), 515-548.

Kambhampati, U.S. & Pal, S. (2001). Role of parental literacy in explaining gender difference: Evidence from child schooling in India. *European Journal of Development Research, 13*(2), 97-119.

Karmel, T. (2011). *As Clear as Mud: Defining Vocational Education and Training*. Adelaide: National Centre for Vocational Education Research.

Kilpatrick, S. & Allen, K. (2001). Review of research: Factors influencing demand for vocational education and training courses. Adelaide: National Center for Vocational Education Research. Retrieved from http://www.ncver.edu.au/publications/610.html

King, G. & Zeng, L. (2007). When can history be our guide? The pitfalls of counterfactual inference. *International Studies Quarterly* (March). Retrieved from http://gking.harvard.edu/files/abs/counterf-abs.shtml

King, G. & Zeng, L. (2006). The dangers of extreme counterfactuals. *Political Analysis, 14*, 131-59.

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

Kingdon, G. (2002). The gender gap in educational attainment in India: How much can be explained? *Journal of Development Studies*, *39*(2), 25-53.

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

Kingdon, G.G. & Monk, C. (2010). Health, nutrition and academic achievement: New evidence from India. Economics Series Working Papers, CSAE WPS/2010-14. University of Oxford, Department of Economics.

Klasen, S. & Pieters, J. (2013). *What explains the stagnation of female labor force participation in urban India?* Discussion Paper No. 7597. Bonn, Germany: IZA, Institute for the Study of Labor.

Kremer, M.R. (1995). Research on schooling: What we know and what we don't – A comment on Hanushek. *The World Bank Research Observer, 10*(2), 247-254.

Kremer, M.R.; Miguel, E.A.; Thornton, R.L. (2004). Incentives to learn. The Review of Economics and Statistics, 91(3), 437-456.

Lamb, S., Long, M. & Malley, J. (1998). Vocational education and training in Australia: *An analysis of participation and outcomes using longitudinal survey data*. Camberwell VIC: Australian Council for Educational Research. Retrieved from http://www.voced.edu.au/content/ngv19147

Lauglo, J. & Maclean, R. (2005). *Vocationalisation of Secondary Education Revisited*. Netherlands, Springer.

Lillard, L. & Willis, R,J. (1994). Intergenerational educational mobility: Effects of family and state in Malaysia. *The Journal of Human Resources, 29*(4), 1126-1166.

Lutz, W.; Goujon, A.; Doblhammer-Reiter, G. (1998). Demographic dimensions in forecasting: Adding education to age and sex. Population and Development Review, 24, 42-58.

Mani, S.; Hoddinott, J. & Strauss, J. (2009). *Determinants of schooling outcomes: Empirical evidence from rural Ethiopia*. Fordham Economics Discussion Paper Series, No. 2009-03. Department of Economics, Fordham University.

Marks, G.; Fleming, N.; Long, M.; McMillan, J. (2000). Patterns of participation in year 12 and higher education in Australia: Trends and Issues. Longitudinal Surveys of Australian Youth, Research Report No. 66, Camberwell, VI: ACER.

McDonough, P.M. (1997). *Choosing Colleges: How Social Class and Schools Structure Opportunity*. Albany: State University of New York Press.

Moenjak, T. & Worswick, C. (2003). Vocational education in Thailand: A study of choice and returns. *Economics of Education Review, 22*, 99-107.

Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from Madagascar. Unpublished Paper, Massachusetts Institute of Technology.

Pal, S. (2004). How much of the gender difference in child school enrolment can be explained? Evidence from rural India. *Bulletin of Economic Research, 56*(2), 0307-3378.

Perna, L.W., and Titus, M. (2005). The relationship between parental involvement as social capital and college enrollment: An examination of racial/ethnic group differences. *The Journal of Higher Education 76*(5), 485–518.

Perna, L.W. (2006). Studying college access and choice: A proposed conceptual model. In J.C. Smart (Ed.) *Higher Education: Handbook of Theory and Research, 21,* 99-157.

Psacharapoulos, G. (1987). To vocationalize or not to vocationalize: That is the curriculum question. *International Review of Education, 33*, 187-211.

Psacharapoulos, G. (1988). Vocational Education. *International Review of Education (Special Issue), 34*(2), 149-172.

Psacharapoulos, G. (1997). Vocational education and training today: Challenges and responses. *Journal of Vocational Education and Training, 49*(3), 285-393.

Planning Commission, Government of India. (2008). Eleventh Five Year Plan (2007-2012). New Delhi: Sage Publications.

Planning Commission, Government of India (2013). Twelfth Five Year Plan (2012-2017). New Delhi: Sage Publications.

Rosenbaum, P.R. (2001). *Observational Studies*. 2nd ed. New York: Springer.

Rothman, S. (2008). Participation in vocational education and training to age 24. Longitudinal Surveys of Australian Youth, Briefing Report No. 17. Camberwell, VI: Australian Council for Educational Research.

Rubin (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

Ryan, P. (2001). The school-to-work transition: A cross-national perspective. *Journal of Economic Literature, XXXIX*, 34-92.

Sandefur, G.D.; Meier, A.M.; Campbell, M.E. (2006). Family resources, social capital, and college attendance. *Social Science Research, 35*, 525-553.

Schultz, T.W. (1961). Investment in human capital. *American Economic Review, 51*, 1–17.

Schultz, P.T. (2004). School subsidies for the poor: Evaluating the Mexican Progresa poverty program. *Journal of Development Economics, 74*(1), 199-250.

Sharma, Y.P. (2010). *Skill Development Programmes in India*. Study Visit on Skills and TVET: India and China, November 2010.

Siddhu, G. (2011). Who makes it to secondary school? Determinants of transition to secondary schools in India. *Journal of Educational Development, 31*(3), 394-401.

Stanton-Salazar, R. D. (2001). *Manufacturing hope and despair: The school and kin support networks of U.S.-Mexican youth*. New York: Teachers College Press.

Stuart, E.A. & Green, K.M. (2008). Using full matching to estimate causal effects in nonexperimental studies: Examining the relationship between adolescent marijuana use and adult outcomes. *Developmental Psychology*, *44*(2), 395-406.

Sulock, J. (1982). The demand for community college education. *Economics of Education Review*, *2*(3), 351-361.

Suryadarma, D., Suryahadi, A., Sumarto, S. (2006). *Causes of low secondary school enrollment in Indonesia*. Jakarta, Indonesia: SMERU Research Institute. Retrieved from http://www.eaber.org/intranet/documents/42/1635/SMERU_Suryadarma_2006.pdf

Tansel, A. (2002). Determinants of school attainment of boys and girls in Turkey: Individual, household and community factors. *Economics of Education Review*, *21*(4), 455-470.

Teese, R. & Walstab, A. (2008). *Social area differences in vocational education and training participation*. Adelaide, SA: National Center for Vocational Education Research.

Tilak, J.B.G. (2002). *Determinants of household expenditure on education in rural India*. National Council of Applied Economic Research, Working Paper Series No. 88. New Delhi: NCAER.

UNESCO. (2001). *Revised Recommendation Concerning Technical and Vocational Education and Training.* Paris: UNESCO.

UNESCO Institute of Statistics (UIS). (2006). *Participation in Formal Technical and Vocational Education and Training Programmes Worldwide: An Initial Statistical Study*. Bonn: UNESCO-UNEVOC.

Walstab, A. & Lamb, S. (2008). *Participation in vocational education and training across Australia: A regional analysis*. Adelaide: National Center for Vocational Education Research. Retrieved from http://www.ncver.edu.au/publications/1998.html

Wooldridge, J.M. (2010). Econometric Analysis of Cross Section and Panel Data. 2nd ed. Cambridge: Massachusetts Institute of Technology.

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

World Bank (2009). *Secondary education in India: Universalizing opportunity*. Human Development Unit, South Asia Region. Washington D.C: World Bank.

World Bank (2006). *Skill Development in India: The Vocational Education and Training System*. Human Development Unit, South Asia Region, World Bank.

Xiao, J. & Tsang, M.C. (2004). Determinants of participation and nonparticipation in job-related education and training in Shenzhen, China. *Human Resource Development Quarterly, 13*(4), 389-420.