I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15002

# Updates on Returns to Education in India: Analysis Using PLFS 2018-19 Data

Jie Chen
Sanghamitra Kanjilal-Bhaduri
Francesco Pastore

JANUARY 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15002

# Updates on Returns to Education in India: Analysis Using PLFS 2018-19 Data

**Jie Chen**
*Jiangsu University*

**Sanghamitra Kanjilal-Bhaduri**
*University of Algarve and STIGMA Foundation*

**Francesco Pastore**
*University of Campania "Luigi Vanvitelli" and IZA*

JANUARY 2022



ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## Updates on Returns to Education in India: Analysis Using PLFS 2018-19 Data

In this paper, we report returns to education in India using unit level data from the nationwide Periodic Labour Force Survey for 2018-19. OLS estimates from the classical Mincerian equation are presented. Various econometric techniques (e.g., conventional IV and heteroskedasticity-based IV models) are used to address endogeneity and sample selection issue. For regular workers, compared to those with no formal education, an additional year of literacy education increases yearly return by 2.3%, primary education by 3.4%, middle school education by 3.7%, secondary school education by 4.5%, higher secondary education by 5.8%, graduate and diploma by 9.8%, and postgraduate and above level of education by 8.2%. We also find a widening of the wage distribution, with striking differences across social groups, sectors, locations. First, returns to middle-school and above level of education are higher for women than for men; second, returns to graduate and above level of education are higher for urban than for rural workers; third, returns to workers in the public sector are higher than returns in the private or third sectors; fourth, returns to the scheduled tribe are the highest across all the castes. Over the last decade, returns to education have reduced. We provide evidence showing that this may be because more people hold higher levels of education qualifications, while the demand for skills remains quite stable. Overall, our policy suggestion is that in India, as in other low- middle-income countries, especially in rural areas, it is important to increase primary and secondary level of education in rural areas, and the tertiary level in urban areas and to equalize the life chances of some social groups.

**JEL Classification:** I26, J15, J16, J30, C20
**Keywords:** returns to education, endogeneity, sample selection, India

**Corresponding author:**
Francesco Pastore
University of Campania "Luigi Vanvitelli"
Corso Gran Priorato di Malta
I-81043, Capua (Caserta)
Italy
E-mail: francesco.pastore@unicampania.it

# 1 Introduction

An important indicator of the reward for education in the labour market is the rate of return to education. India has experienced an increase in educational attainment since the 1990s. In his study based on the National Sample Survey (NSS) data for 25 years (1983-2005), Fulford (2014) found that the number of years of education has increased steadily for men and women. This increase in education was made possible, among others, by the District Primary Education Program by Government of India with World Bank support during the 1990s.

According to OECD (2019), since then, the share of tertiary-educated adults has been further growing in India, albeit not satisfactorily and not equivocally among the entire adult population. In 2011, 11% of 25-64 year-olds had a tertiary education, and this improved to 14% only, that too exclusively among young adults (25-34 year-olds). Moreover, 17% of 25–34-year-old women had a tertiary qualification in 2019 compared to 22% of their male peers. By 2030, tertiary-educated adults from India are expected to make up more than one-fifth of the tertiary-educated population across OECD and G20 countries. India's contribution to the tertiary-educated population of these countries is projected to increase from 17.6% in 2015 to 20.8% in 2030. However, India has also by far the highest share of adults without a primary education among the G20 countries. In 2011, 46% of 25–64-year-olds had not completed primary education. The share of adults (25–64-year-olds) without upper secondary education is 71% in India, whereas among the 25-34-year-olds the share falls to 64%.

What would be the correct expectation about returns to education in a middle-income country like India? In their study of 142 economies for the period 1970-2014, Montenegro and Patrinos (2021) found that private returns to schooling are generally positive, higher in low- or middle-income economies, highest at the primary schooling level, higher for women and exhibiting modest declines over time. Based on their study we expect to find relatively high returns to education, as India is a low-middle income country and the share of tertiary educated is relatively small which would point to high returns to education. On the other hand, the dramatic increase in education attainment over the last decade would point to a slight reduction of returns to education as compared to the early 2010s, considering the relatively stable economic structure.

Our paper adds to the literature on returns to education in India (see, among others, Duraisamy, 2002; Vasudeva-Dutta, 2006; Kijima, 2006; Madheswaran & Attewell, 2007; Fulford, 2014) by using large-scale PLFS data of National Sample Survey Office (NSSO) for the most recent period available (2018-19)[1]. The previous papers covered up the period until the year 2011-12, and, even then, yearly returns to education were not provided for 2011-

[1] Although the third annual round of the PLFS data conducted during July 2019-June 2020 has been released recently, yet we do not consider it for our paper because there seems to be some discrepancies in estimates based on this release. This suggests that the data needs some settling before being ready for research purposes.

12, 68th Round NSS-EUS data by any previous study. Mendiratta and Gupt (2013) have provided estimates for the years 2004-05, 61st Round NSS-EUS data. Therefore, we are addressing an important research gap in education economics in India and producing an update after a period of almost two decades. It allows us to understand the most important changes that happened in the evolution of the market for skill. We find lower yearly returns to general education (5.5% for regular workers, 2.7% for self-employed workers, and 0.6% for casual workers) compared to the existing literature.

The conclusion remains even after we additionally control for any technical education. One reason of the difference is that the average years of completed education in our sample are significantly larger than those in previous studies, while we do not find a proportional increase in the sectors employing high skill workers (see Figure 2 below). As more people hold higher levels of education qualifications, while the demand remains quite stable, returns to education in general could decrease, as we observe here. In other words, our study points to a tendency of supply to win the race against demand for skill in the country over the last decade. This is in line with Asadullah (2006) claim that in low-middle-income countries especially in rural areas it is more important to increase the primary and secondary level of education rather than the tertiary level. Tertiary education is a goal to reduce inequalities within a complex society in favor of women and some social groups lagging behind.

In his study of changes in wage structure in urban India for the period 1983-2004, Azam (2012) mentioned that, while the returns to higher education (secondary and tertiary) both increased in the 1990s, they have become more heterogeneous across the distribution, with larger returns at the higher quantiles. However, previous studies did not delve into the heterogeneity of returns to education, per se. There seems to be a dearth of research focusing on this aspect and to the best of our knowledge, none using the PLFS data[2]. In this paper, we address the internal heterogeneity of returns to education to an extent that was never done previously in such a comprehensive way, as India is a varied society which has been experiencing rapid structural change since 1991. Estimates of the returns to education (general and technical education) in wage employment in India by gender, class, caste[3] (social-groups), religious groups, sector[4] (public, private, non-profit) and location (rural–urban) are provided in this study, making use of the ordinary least squares (OLS) method. As robustness checks for our results, we also use the instrumental variable (IV) method and the heteroskedasticity-based Lewbel method. To the best of our knowledge, there are no

[2] A recent paper by Bahl et. al (2021) while studying the process of school-to-work-transition in India, has estimated, using the PLFS 2018-19 data, whether investment in Vocational Education Training brings additional returns for workers across the age cohorts.

[3] The caste system is a social stratification, specific to India, which is more than 3000 years old. Caste is defined as a socially homogenous class and also an occupational grouping, membership of which is involuntary and hereditary. It divides individuals into rigid hierarchical groups based on their work and religion. See the Appendix for a detailed discussion on the Indian Caste System.

[4] PLFS enlists Enterprise Type as Propreitary, Partnership, Government/local Body, Public Sector Enterprise, Autonomous Bodies, Public/Private Limited Company, Co-operative Societies, Trust/other Non-Profit Institutions. We have created three sectors, viz. Public Sector, Private Sector and Non-Profit Sector (third sector) from this variable.

previous studies comparing OLS estimates with the Lewbel estimates for India and we are the first to do so. We provide evidence of an amazing heterogeneity of returns to education and premium for different groups.

Most of the previous studies on returns to education in India (Duraisamy, 2002; Vasudeva-Dutta, 2004; Kijima, 2006; Kanjilal-Bhaduri & Pastore, 2018) consider only the wage rate as dependent variable, which causes them not to include in their analysis the self-employed or informal workers and hence apply to a very selected section of the population. It is an important limitation in India, where a substantial proportion of the population either is self-employed or is working in the informal sector. Our study adds on the previous literature by taking into consideration the declared earnings for the self-employed individuals too, which were not available in previous survey data.

The rest of the paper is structured as follows: Section 2 presents the Motivation for such a study; Section 3 presents an exhaustive Literature Review; Section 4 presents the Methodology; Section 5 outlines the Data and Variables along with Descriptive Statistics; Section 6 discusses the Results; and, finally, Section 7 presents the Conclusion.

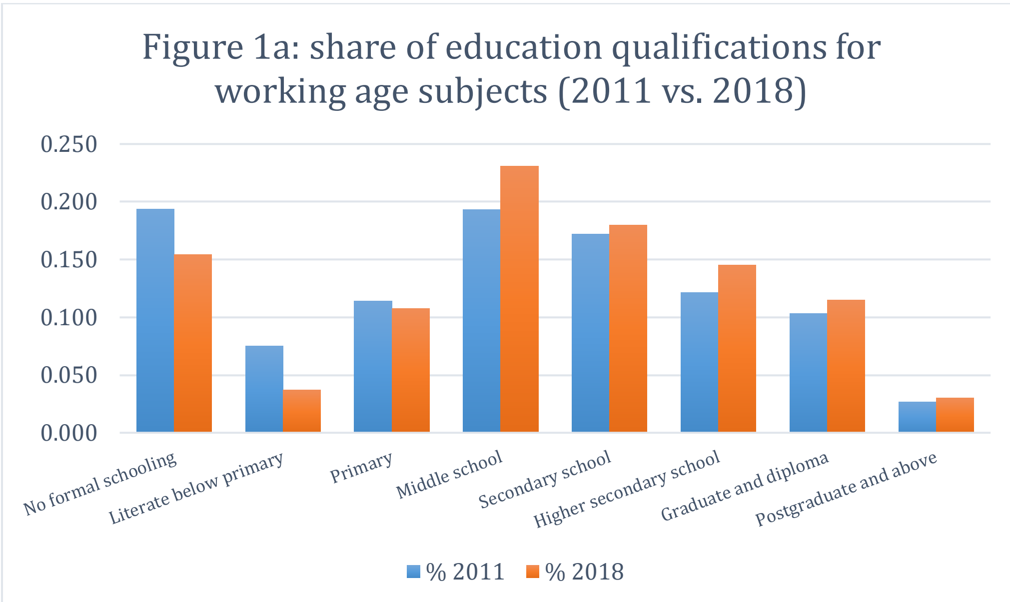
## 2. Motivation

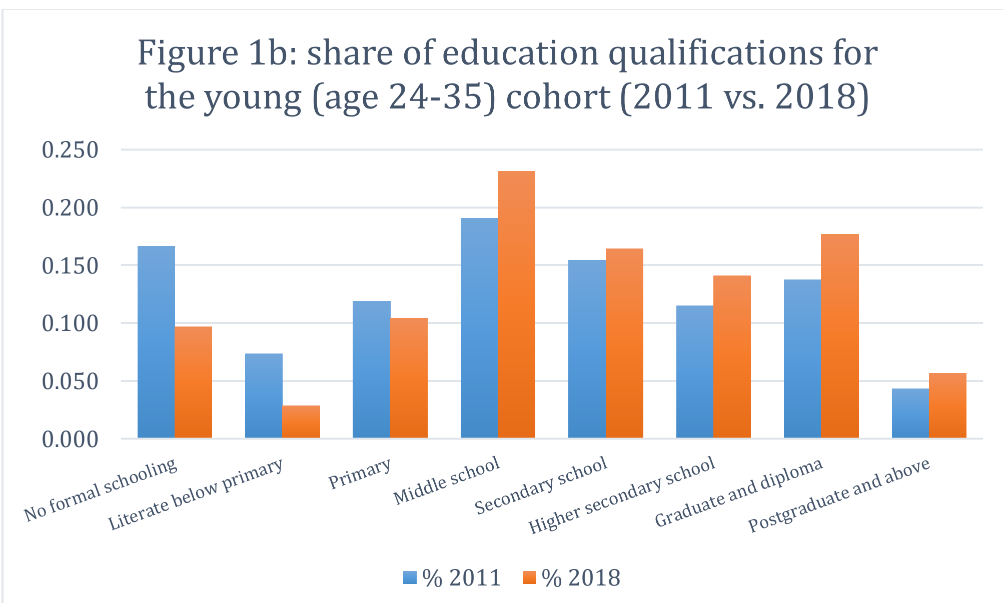
The last decade has been a period of deep structural change from both the supply and the demand side of the market for skills. India is one of the fastest growing economies in the world and has a high share of young population. The country boasts of an extensive tertiary education sector with a much higher percentage of the relevant age-cohort going to the universities (Dreze and Sen, 1991). As per Census 2011, youth (15-24 years) in India constitutes one-fifth (19.1%) of its total population, Education plays an important role in accruing the benefits of this demographic dividend (Lutz et al., 2019).

In 2011-2012 ($68^{th}$ Round NSS-EUS data), 19.4% of the entire sample (15-59 years old workers) had no formal education; 7.5% were literate but had below primary level of education; 11.4% had primary level; 19.3% had studied up to middle level; 17.2% had secondary level education; 12.1% had higher secondary education; 10.3% were graduates/diploma holders and 2.7% were educated post-graduate and above. We compare the shares using two different datasets because the PLFS data is not available in 2011-2012; however, it is an extension of the NSS-EUS data (which was discontinued after 2011-12 and converted into PLFS since 2017-18) and hence similar to it.

Our dataset shows that, in 2018-2019, 15.4% of the entire sample (15-59 years old workers) had no formal education; 3.7% were literate but had below primary level of education; 10.8% had primary level; 23.1% had studied up to middle level; 18.0% had secondary level education; 14.5% had higher secondary education; 11.5% were graduates/diploma holders and 3.0% were educated post-graduate and above level.

Comparing the two datasets it can be concluded that from 2011 to 2018 there has been a decline in below primary and primary level of education, but an increase in middle, secondary, higher secondary, graduate and post-graduate level education, as can be seen from Figure 1a below. The trend further stands out if we zoom into the young cohort aged 24-35, as shown in Figure 1b.




On the demand side, the Indian economy opened up to world trade in 1991, which resulted in an increased demand for highly skilled labour. Foster and Rosenzweig (1996) found in their study that rapid technical progress brought about by such reforms have caused an increase in returns to schooling. Over the years, other studies concluded that as liberalization increased the demand for higher educated individuals, so this led to an increased rate of return corresponding to higher levels of education (Duraisamy, 2002; Vasudeva-Dutta, 2006; Kijima, 2006; Madheswaran & Attewell, 2007; Agrawal, 2011; Geetha Rani, 2014).

Yet ironically, official statistics (Statista, 2019) show that, with a share of 16.3 percent, Graduates made up the highest unemployment rate in 2019, followed by individuals with a post graduate degree or above with a share of 14.2 percent. The unemployment rate in the country was higher among youth with higher educational qualifications, indicating a poor rate of returns to education, suggesting that the supply of education is increasing beyond the demand.

Indeed, the evidence seems to suggest that the 1990s' push in favor of a higher demand for skills has come to a standstill in more recent years. To see this, Figure 2a and Figure 2b may be of some help. Each dot in the figures represent an industrial sector. On the horizontal axis we measure the share of employed individuals holding tertiary education (2a) and high secondary education (2b) in each sector in 2011. On the vertical axis, we measure the change in the employment rate in each sector from 2011 to 2018. The figures clearly show that the expanding sectors are those occupying a low to middle share of tertiary educated. Only a minor fraction of sectors with a low share of graduates has experienced an employment reduction over the period considered. The industrial sectors occupying the largest share of individuals with tertiary education lie very close to the 0-line of no increase in employment. Overall, this is clear evidence that the demand for skills has not increased much over the last decade or so.

Figure 2a: share of university-degree holders vs. change in employment between 2011-2018

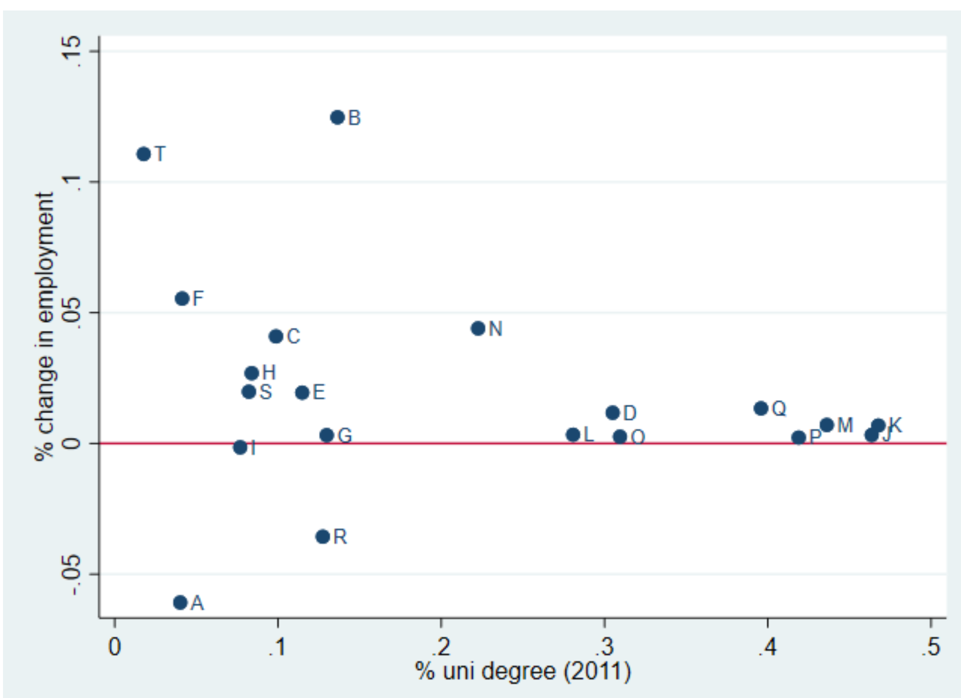
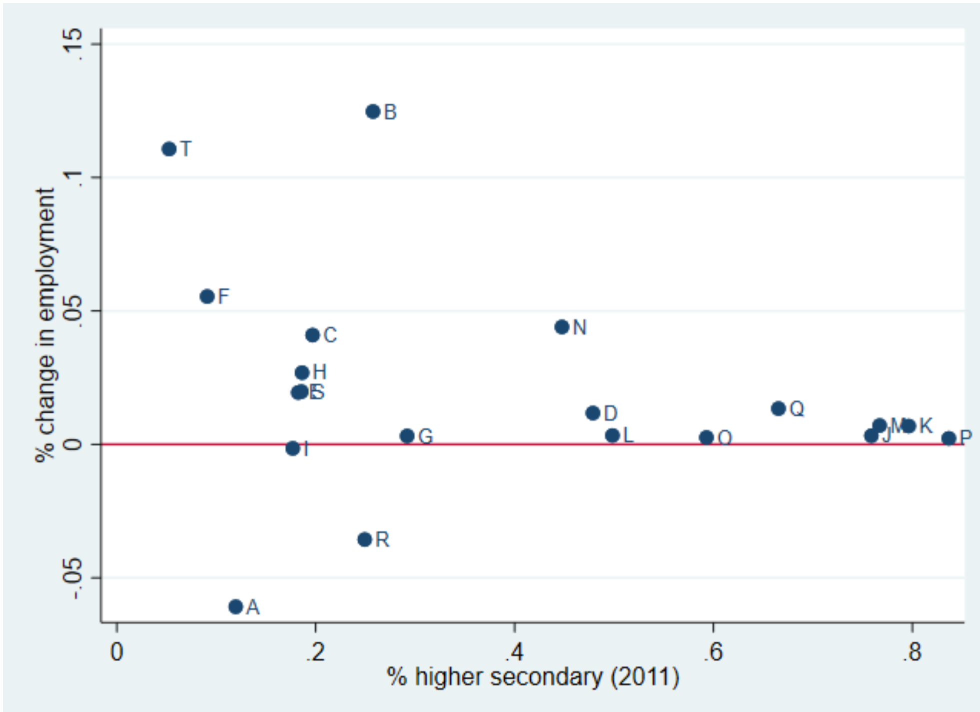


Figure 2b: share of higher secondary degree holders vs. change in employment between 2011-2018




Notes: A - Agriculture, forestry and fishing; B - Mining and quarrying; C – Manufacturing; D - Electricity, gas, steam and air conditioning supply; E – Water supply; sewerage, waste management and remediation activities; F – Construction; G - Wholesale and retail trade; repair of motor vehicles and motorcycles; H - Transportation and storage; I - Accommodation and Food service activities; J - Information and communication; K- Financial and insurance activities; L – Real estate activities; M - Professional, scientific and technical activities; N - Administrative and support service activities; O - Public administration and defence; compulsory social security; P – Education; Q – Human health and social work activities; R - Arts, entertainment and recreation; S - Other service activities; T – Activities of households as employers; undifferentiated goods- and services producing activities of households for own use. Data are for working age individuals only.

Also, the Indian Education System has yet to achieve caste, class and regional equality. Thorat and Newman (2007), among others, mention that even after 60 years of effort by the Indian government to make the education system inclusive and empower the weaker sections of society, participation of the lower castes in higher education still does not match their share in the total population.

The abovementioned stylized facts motivated us to look deeper into the following points and to emphasize on the heterogeneity analysis:

a) It is not fully clear what is the return to education in the recent years as the last available estimates are more than a decade old;

b) We know very little not only about returns to education in terms of wages, but also of employment opportunities in recent years;

c) We know very little about determinants of self-employment and informal sector earnings from previous studies;

d) The Indian society is very complex and articulated with different castes, religions and other factors that alter the market mechanisms, although there is little evidence regarding the actual differences in returns to education by castes and religion.

## 3. Literature review

Previous studies provide evidence that individuals with higher levels of education, better skills, and greater experience have higher incomes after correcting for individual, household, and other differences (Psacharopoulos, 1985; Mankiw et al., 1992; Sianesi and Van Reenen, 2000; Card, 2001). The pioneering works of Human Capital Theory by Schultz (1960) and Becker (1964), embed the calculation of rate of return, thereby providing it an important position in development literature (Mitra, 2019). Other studies, (among others, Psacharopoulos and Patrinos, 2002; Fasih et al., 2012) have highlighted the importance of the contribution of human capital towards an increase in an individual's earning capacity. Studies on wage returns to education in developed and developing countries show that, internationally, one additional year of education adds approximately 10% to a person's wage, at the mean of the distribution (Psacharopoulos and Patrinos, 2004; Montenegro and Patrinos, 2021).

Although the literature on returns to education is one of the most extensive in labour economics, yet, quite surprisingly the literature on returns to education for India is quite outdated, being based on data about a decade old, while there is dearth of more recent studies. Studies, which used nationally representative surveys (Duraisamy, 2002; Vasudeva-Dutta, 2006;; Kijima, 2006; Madheswaran & Attewell, 2007; Agrawal, 2011; Azam, 2012; Geetha Rani, 2014), found that rates of return increased with higher levels of education because of liberalisation that, in turn, increased the demand for higher educated individuals. For women, Kanjilal-Bhaduri & Pastore (2018), while studying NSS-EUS 68$^{th}$ round (2011-12) data, found U-shaped returns to education.

Duraisamy (2002) uses the employment rounds of NSS in 1983 and 1993-94 to estimate a Mincerian regression of log wages on education for individuals and finds that the returns to secondary education are the highest (17.3%), compared to returns to all the other education qualifications (7.9% for primary, 7.4% for middle, 9.3% for higher secondary, and 11.7% for college). His study finds that returns per year of schooling are, in general, significantly lower in rural areas than in the urban areas. However, the returns to lower-level education (e.g., primary, secondary) in rural areas are significantly higher than those in urban areas (69% for men and 32% for women). The author notes that the rural-urban status is the place of current residence, instead of the place of schooling or the place of birth.

Thus, the reversing rural-urban difference at lower education levels most likely indicates that individuals with higher education attainment self-select into urban areas, rather than indicating the actual rural-urban difference in returns to education.

Vasudeva-Dutta (2004) looks at the adult male regular and casual wage earners, taking into account the presence of a dual labour market. Her study finds that the returns to education for casual workers are constant whereas *for* regular workers the relationship between returns and education levels is U-shaped. Specifically, the returns to middle school (2.02%) are the lowest, compared to primary school (2.62%), secondary school (4.72%), and graduate school (9.48%)[5]. Her study also finds evidence of a widening wage gap between regular workers with graduate and primary education which she explains could possibly be a consequence of trade liberalization and other reforms pursued during the 1990s.

Azam (2012) examines changes in the wage structure in urban India over three employment rounds of NSS, during 1983-2004, and suggests that the returns to secondary and tertiary education increased during the 1990s. Using quantile regression method, the study presents the changes in returns to different levels of education over time (1983-2004) for wage earners across the entire wage distribution. It finds there is not much change in returns to different levels of education during 1983–93, but the returns to tertiary education increased nearly by 18% across the entire distribution, from 1993 to 2004. Kijima (2006) while examining changes in the wage inequality in urban India during 1983-99 found out increasing returns at the tertiary education levels after 1991.

Fulford (2014) found a positive relation between education and wages for men (but not for women) for the years 1983-2004: men had a 4.6% return to an added year of education in the occasional wage market (vs -0.07% for women) and a 5.8% return in the salary market (vs. -0.63% for women). Women may be exhibiting negative returns, due to selection into working for a wage. But the problem is that only 41% of men and 14% of women work for wages in India and the economy has not yet developed enough to enable widespread formal employment. Madhesaran and Attewal (2007) use NSS data for 1983, 1993-94 and 1999-2000 to examine the wage gap between higher castes and Scheduled Castes/Tribes (SCs/STs) in regular salaried urban labour market. They conclude that, in spite of the efforts of the Government of India to remove inequalities in higher education and government and public sector employment[6], SCs still continue to lag behind the general population in terms of

[5] These return rates are simple average over the three years (1983, 1993, 1999), manually calculated by the authors.

[6] Policy makers in India believed that having access to places reserved for SC/ST candidates would help reduce some of the educational disparities. Three broad categories of reservations, or quotas for SCs and STs are available, viz.: employment, educational, and political. With rapid population growth and an increase in mass education, college admissions have become increasingly competitive in India. Hence, 15% and 7.5% of the places in higher education are reserved for SCs and STs, respectively, although states can observe different quotas based on their respective SC/ST populations. 15% of government jobs at all levels are reserved for SCs, and 7.5% for STs. This includes jobs in central government, in state government and in public sector units. Since government is the largest formal sector employer, accounting for over 66% of all jobs in India, this is a substantial benefit.

educational attainment level and an overwhelming majority of individuals belonging to the SCs/STs are found in low-skill, low-paying jobs.

Agrawal (2011) uses data from the nationally representative household survey- India Human Development Survey (IHDS), conducted in 2004-05, using standard Mincerian wage equations, (with log of hourly wages as the dependent variable), separately for rural and urban sectors to estimate the private returns to education in India. The paper shows that an additional year of education corresponds to 8.5% higher returns in hourly wages. It also finds positive correlation between returns to education and levels of education, contradicting the hypothesis of diminishing returns to education. The findings of the study indicate that returns to education are lower for rural than for urban residents (e.g., 4.64% vs. 6.59% for primary education). This finding is consistent with those in Duraisamy (2002). In general, the disadvantaged social groups of the society tend to earn lower wages and family background is an important determinant affecting the earnings of individuals.,

Geetha Rani's (2014) paper studies the impact of different levels of education, religion, caste as well as the impact of living in urban and rural communities on earnings in India. The study uses data from a large cross-section sample of India Human Development Survey, conducted in 2005 to estimate Mincer and augmented Mincer equations, considering log of hourly wage as the dependent variable and education measured in years of schooling as a continuous variable, experience and experience square as explanatory variables. The basic Mincer equation estimates the average rates of return to education as 14% and when controls (such as ability of the individuals) are introduced then the estimates of returns to education reduce to 10%.

Mendiratta and Gupt (2013), using the pseudo panel approach, estimate returns to education in India for different levels of education, on the basis of location and gender. They make use of the standard Mincer equation using employment-unemployment data from the 61$^{st}$ (2004-05) and 66$^{th}$ (2009-10) round of the NSSO. Their study finds that the average return to education is around 15% per year of education, but OLS underestimates the returns at 10.8%. So, they conclude that education is more rewarding at higher levels and returns to education do not decline after secondary level. A gender wise comparison of returns shows that returns at initial levels of education (primary and middle) are lower for females but for higher levels of education the situation reverses. Returns to female education for technical diploma / certificate are as high as 37.13% whereas they are 24.55% for male education.

Mitra (2019), using the 68th Round Employment-Unemployment (EUS) (2011-12) data has estimated the marginal returns to education and calculated rates of return for different levels of education across various disaggregation like male/female, social group wise and public/private sector, using quantile regression method. The study has found that the rates of return for elementary education at the lowest wage quantile is 1.25% whereas for the highest wage quantile it is 3.63%. For secondary education the values are respectively 7.5% and 5%. For higher secondary it is 5% and 9%. Finally, for Graduate and above level of education it is 9.33% and 15.33%. Rate of returns to education show a linear increase across educational levels for the highest wage quantile but not so for the lowest. For females, at the lowest wage quantile returns to education for elementary education is 1.5%, which increases to 15.33% for graduate and above level of education. However, for males the same increase is only to 6.67%, from 1.38%. At the highest wage quantile, the increase for females is from 3.5%-20%, whereas for males it is 3.38%-14.33%. The study thus concludes (similar to Azam's study, 2012) that rates of return are higher at the higher end of wage distribution across India.

This is the sole paper which has studied sectors and it has found that for the public sector, the rates of return are the highest for the secondary level (14.8%) and gradually diminish for the graduate and above level (4.64%). In the private sector, the rates of return are the lowest for elementary education (0.93%) and highest for the graduate and above level (15.87%). The study also found that at the highest wage quantile in the private sector, the rate of return for the graduate and above level of STs is 4.43% and that for other castes it is close to 21%. We make a similar attempt to study sectors with PLFS data. Meanwhile, in addition to private and public sectors, we also consider the third sector.

From the literature summary in Table 1, we see that all but one study uses the National Sample Survey (NSS) data in their research. The only exception is Agrawal (2011), who uses the India Human Development Survey (IHDS) data. However, none of these studies use the PLFS dataset. To the best of our knowledge, our paper is the first that uses the PLFS dataset to estimate the returns to education in India.

Table 1: A brief summary of literature on returns to education in India

| Paper | Returns to education (%) | Methodology | Data | Location |
|---|---|---|---|---|
| Mitra (2019) | No average estimates; 1.25-3.63 for elementary; 2.00-9.00 for secondary; 9.33-15.33 for graduate and above | Mincerian equation; Quantile regression | 68th(2011–2012) round of the National Sample Survey (NSS) | National (456,999 individuals ) |
| Mendiratta and Gupt (2013) | 11.7 | Mincerian equation | 61st(2004-05) round of the National Sample Survey | National(-) |
| | 10.76 | Mincerian equation | 66th(2009-10) round of the National Sample Survey | National(-) |
| | around 15 | Pseudo panel approach using WLS | 61st and 66th round of the NSS | National(-) |
| | 7.35 for elementary; 7.68-15.13 for secondary; | Extended earnings function | 61st round of the NSS | National(-) |
| | 4.75 for elementary; 6.89-13.5 for secondary ; | Extended earnings function | 66th round of the NSS | National(-) |
| Duraisamy (2002) | 7.9 for elementary; 7.4 for middle; 17.3 for secondary; 9.3 for higher secondary; 11.7 for college; | Extended earnings equation | NSS (1993-1994) | National (83,900) |
| Vasudeva Dutta (2006) | 3.29 for elementary; 2.35 for middle; 5.31 for secondary; 9.02 for graduate school; | Mincerian earnings equation | NSS, January–December 1983 | National, male ( 27,356) |
| | 2.13 for elementary; 1.69 for middle; 4.27 for secondary; 9.15 for graduate school; | Mincerian earnings equation | NSS, July 1993– June 1994 | National, male ( 26,387) |
| | 2.43 for elementary; 2.02 for middle; 4.64 for secondary; 10.26 for graduate school; | Mincerian earnings equation | NSS, July 1999–June 2000 | National, male ( 27,295) |
| Agrawal (2011) | 8.5 | Mincerian earnings equation | India Human Development Survey (IHDS) 2005 | National (46,965) |

## 4. Methodology

We start by estimating a standard type of Mincerian education production function as follows.

$$Y = \alpha_0 + \beta edu_i + \gamma X + \varepsilon$$

where $Y$ is the natural logarithm of monthly wages, which remains so in all specifications This means that we estimate a log-linear model and, therefore, the coefficients can be interpreted as semi-elasticities: they measure the percentage change in wages for each unit increase in the regressor[7]. $edu_i$ is a categorical variable representing the different education qualifications in India, as is described in Table A1. $X$ is a vector of control variables and ε is the random disturbance term. The control variables include age, age-squared, gender, whether the individual is from the urban or rural sample, household size and household type, marital status, religion, social groups (unique to India[8]), firm's ownership, and the sector of industry where the worker is employed. A table with the variables' definition is provided in the Annex (Table A1).

For casual workers we also control for the monthly hours of work. Unfortunately, the working hours variable is neither available for regular workers nor for the self-employed. We also include interactions between marital status and gender, as existing literature has found strong evidence that females and males are affected differently by marital status in terms of wages (see, among others, Chen and Pastore, 2021). Chen and Pastore (2021) find that males experience a marriage premium of around 23% whereas females suffer a marriage penalty of around 15%. Indeed, typically, marriage reduces the reservation wage of husbands, pushing them to work more, while it increases the reservation wage of wives, pushing them to work less.

One of the major challenges in estimating returns to education is endogeneity (Card, 1999). Because education decisions are often an outcome of family decisions which is affected by parental socio-economic background and children's innate ability, it is difficult to untangle these factors and estimate the pure returns to education. One

---

[7] In the case of independent dummy variables, like overeducation, the semi-elasticity interpretation is flawed and, following Halvorsen and Palmquist (1980), it should be computed as: $\left(e^{\beta}-1\right)*100$. This formula measures the percentage change in the median wage, which is less affected by outliers. Nonetheless, many authors interpret also the estimated coefficients of dummy variables directly as semi-elasticity. This is acceptable when the estimated coefficient is sufficiently close to zero. In what follows, we will leave to the reader the calculation of exact semi-elasticities.

[8] Refer to Appendix for a discussion on the Indian Caste System.

of the mainstream methodologies to deal with the endogeneity issue is the Instrumental Variable (IV) method. In this paper, we have tried two IVs: the proportion of individuals who completed primary school by state (36 states in total) and the proportion of individuals who completed compulsory education by state. These two instruments are relevant in that the proportion of individuals who complete primary/compulsory education is a good predictor of the average education attainment level in each state. Meanwhile, these instruments satisfy the exclusion restriction in that no individual decision can alter the state-level proportions. Although our first stage F statistics are larger than 300, far above the empirical threshold of 10 or 11, we find that the IV estimates greatly inflate the OLS estimates (sometimes more than double the OLS estimates). Detailed results are discussed in Section 6.3.

Because of the lack of improvement from using IV, we apply a more recent heteroskedasticity-based method – the Lewbel method (Lewbel, 2012; Baum and Lewbel, 2019). The method is an improvement of conventional IV in that it can identify endogenous regressors through exploiting the heteroskedasticity in the error term and not completely relying on an external instrument. Meanwhile, the Lewbel method has been demonstrated efficient in studying returns to education in China (Asadullah and Xiao, 2019). Additionally, if external instruments are available, the Lewbel method can be used as a robustness check of the conventional IV method (Baum and Lewbel, 2019). The Lewbel results will be discussed further in Section 6.3 as well.

## 5. Data and Variables

### 5.1 Periodic Labour Force Survey

We have used a large sample unit/individual level dataset on employment and unemployment in India from the PLFS for 2018–2019[9] (PLFS, 2019). Although the third annual round of the PLFS data conducted during July 2019-June 2020 has been released recently, yet we do not consider it for our paper due to the fact that there seems to be some discrepancies in estimates provided by the said dataset. Labour market distress was caused in India as economic activities came to a standstill due to the stringent nationwide lockdown called to halt the spread of Covid-19; further, this resulted in falling GDP growth, which should have been depicted by the PLFS 2019-20 data; rather, very surprisingly, the data reveal that the unemployment rate, as measured by the usual status, fell from 6.1 percent in 2017-18 to 4.8 percent in 2019-20 [10], which seems

[9] Data downloaded from http://microdata.gov.in/nada43/index.php/catalog/146
[10] Anand. I (2021), 'A disconcerting picture behind the headline numbers: There is evidence to suggest that the PLFS data may underestimate the loss of earnings and fall in consumption', The HINDU, August 3, 2021. https://edumo.in/wp-content/uploads/2021/08/The-Hindu-Newspaper-pdf-3-August-2021.pdf

puzzling during such severe economic distress. Also, rise in the average income of salaried workers, as depicted by PLFS 2019-20, does not match with other data generated by small scale surveys conducted during the lockdown period. These surveys reported massive earnings loss during the lockdown[11]. Hence, in our opinion, the PLFS 2019-20 data may underestimate the loss of earnings during the lockdown and in such an eventuality, the returns to education would not be accurate.

The PLFS covers a large sample size of households across the states and the Union Territories. It follows a multistage stratified sampling to cover the households. The dataset provides details of employment and unemployment and information on other socio-economic and demographic features of individuals. The National Sample Survey (NSS) format of interview schedule is followed by PLFS while collecting data on employment and unemployment, relative to other socio-economic and demographic information. The erstwhile NSSO-EUS (Employment Unemployment Survey) and current PLFS contain rich data on educational attainment and socio-economic characteristics (age, religion, caste and monthly expenditures) at the individual level.

We have restricted the sample for our analysis to only the working-age groups (i.e., people belonging to the age group 15–59 years) who, at the time of survey worked either as paid employees (salaried/casual labourer) or self-employed[12] (employer or own account worker for the family enterprise) for the majority of the time in the last year (at least six months) and use the primary activity (Usual Principal Status[13]) or occupation of work. The monthly earnings schedule records the earnings in the last referenced 30 days for self-employed and paid workers (casual/salaried) and is taken as our main earnings variable[14].

## 5.2 Variables

The existing literature (see, among others, Klasen and Pieters 2012; World Development Report 2012; Raveendran 2016) suggests that important determinants of wage work participation in India are human capital endowment (education and work experience), socio-economic and cultural factors, access to resources (skills and capital through technical education).

[11] A telephonic survey of 4000 workers across 12 states by Azim Premji Foundation (2020) in collaboration with the Centre of Civil Society showed that 80% of the workers in urban areas reported employment loss, while 50% of the remaining reported income losses or even no salary disbursement.

[12] Data on the earnings of self-employed, besides the income of wage-employees in India was collected for the first time during PLFS 2017-18 and this was repeated during PLFS 2018-19 as well.

[13] In PLFS data 'Usual Principal Status' (UPS) of a person is identified by using a reference period of 365 days preceding the date of survey. A person is considered as being in the workforce if he/she is gainfully employed for a major part of the preceding 365 days

[14] INR74 (Approx)=1USD and INR87(Approx)=1 Euro.

Accordingly, we have split explanatory variables into the following categories: individual, household and social characteristics. For the measurement of the returns to education, two variables are of importance viz. wages and the years of education attained. In the PLFS questionnaires, the recall period for waged earnings is one month. Years of education is considered as the measure of human capital accumulated and we have calculated the same based on a mapping of the Indian education system onto the International Standard Classification of Education[15].

PLFS data provides educational attainment levels and clubs together post-graduate and above level education in a single code. Hence, we have followed Kingdon and Theopold (2008), to convert levels of education into years of education (See Table A4 in Annex). We only consider years of education completed, without any repeats and any years of uncompleted education: due to the sheepskin effect, completed education, with the accompanying degree captures more accurately the level of human capital accumulated than the years spent in schooling.

## 5.3Descriptive statistics

PLFS 2018–2019 was conducted between July 2018 and June 2019. It covered 55,812 rural households (239817 individuals, 57%) and 45,767 urban households (180940 individuals, 43%) in India (420,757 individuals in total). 13.44% respondents belonged to Scheduled Tribe, 16.92% were Scheduled Castes, 39.77% belonged to Other Backward Class and 29.88% belonged to Others (Upper Castes). Hinduism being the major religion of the country, 73.92% respondents were Hindus, 14.845 belonged to Islam, 6.93% were Christians, 2.12% were Sikhs, 1.09% practiced Buddhism and around 1% belonged to the other religions (Jainism, Zoroastrianism and others).

Detailed descriptive statistics are presented in Table A2. Regular workers on average earn 4496 Rupees per month and self-employed workers earn a similar amount of 4512 Rupees, whereas casual workers earn less than 1000 Rupees (most likely because casual workers work much fewer hours per month). Around 36.2% of the working-age sample is regular workers and more than 45.7% is self-employed, whereas only 18.1% is casual. 23.5% of the sample is female, indicating that females are potentially underrepresented in the working-age population in the current sample.

[15] The mapping is prepared by the OECD, accessible via https://gpseducation.oecd.org/Content/MapOfEducationSystem/IND/IND_2011_EN.pdf. The exact mapping in our case is as follows: compulsory education is equivalent to 8 years of education; secondary education is equivalent to 10 years of education; academic upper secondary education is equivalent to 12 years of education (whereas vocational upper secondary education is equivalent to 13 years of education); academic college and tertiary technical education is equivalent to 15 years of education (tertiary professional education is equivalent to 16 years of education); master's education is equivalent to 17 years of education; and doctoral education is equivalent to 22 years of education.

Closer investigation reveals several other gender differences: first, among male (female) workers, 89.7% (84.7%) work in the private sector, 9.4% (13.3%) work in the public sector, and 0.9% (1.9%) work in the third sector; second, an average male has 8.88 years of education whereas an average female only has 6.83 years of education; third, the employment rate for male is 92.4%, much higher than the employment rate of 73.9% for female. We have investigated gender distribution across agriculture, manufacture, and other sectors. Unsurprisingly, all sectors are male-dominant (male-to-female ratio greater than 3:1), indicating female's under-participation in the labour market overall. Meanwhile, we have looked into sector distribution within each gender. Of males, 29.45% work in agriculture and 13.14% work in manufacture sector. Of females, 37.62% work in agriculture and 13.53% work in manufacture sector.

The average age of working-age people is 37.7, substantially younger than the average age of 43.3 in the Chinese labour market (Chen and Pastore, 2021). This is interesting as it indicates India has younger labour supply. With a young and massive population, India could potentially enjoy larger population dividend if its labour force was better educated. India could take cue from China, which leads in the field of literacy with a much more extensive system of education, whereby it invested first in mass literacy (basic education) before expanding higher education while India did the opposite, focusing on higher education (Dreze and Sen, 1991).

For the survey, rural households are visited once, but the urban households are revisited. Since the estimates of most of our indicators, including the distribution of workers across gender, caste, religion, location and sectors (public, private and non-profit sector) can be generated from the first visit only, so we have used the first visit data for both the rural and urban households in this analysis.

# 6. Results

## 6.1OLS results

Table 2 presents the Mincerian regression results for regular workers, self-employed workers, and casual workers, respectively. All three regressions have relatively large adjusted $R^2$ (0.421 for regular workers, 0.355 for self-employed workers, and 0.496 for casual workers) for a cross-section estimate, indicating strong explanatory power of our specifications. Column (1) shows that, compared to those with no formal education, workers who are literate below primary earn 7% higher monthly wages. Correspondingly, primary education raises average earnings by 13.8%, middle school 24.9%, secondary school 34.0%, higher secondary school 45.6%, graduate and diploma 72.5%, and postgraduate and above 88.9%. By and large, these estimates are comparable to findings in the literature (see, for instance, Duraisamy, 2002; Vasudeva Dutta, 2006).

To make the results comparable to the broader literature, we also calculate the yearly returns[16] and report them in Panel B of Table 2. Compared to regular workers with no formal education, an additional year of literacy education raises an average regular worker's earnings by 2.3% (primary education: 3.4%, middle school: 3.7%, secondary school: 4.5%, higher secondary: 5.8%, graduate and diploma: 9.8%, post graduate and above: 8.2%). The story is similar for self-employed workers, except that an additional year of middle school education pays better than an additional year of secondary or higher secondary education. This is possibly because, compared to knowledge acquired through secondary education, knowledge acquired through early entry into the job market is more beneficial to self-employed workers. This is echoed by the overall lower return rate for self-employed than for regular workers. However, graduate and above levels of education still pay significantly better (5.7% for graduate and 6.9% for post graduate) than all the other levels of education. Interestingly, levels of education do not seem to affect casual workers' earnings significantly.

Age, a proxy of potential work experience, has a quadratic relationship with wages (e.g., for self-employed workers, annual return rate reaches the maximum at the age of 26.5; for regular and casual workers, the coefficient for the square term is extremely small, thus the return rate is virtually always increasing during one's work life). Rural workers earn 19.0% less than urban workers. Wages negatively correlate with household size: probably the effort in economic activity is reduced because of effort in non-market work. In a developing and surplus labour economy like India, open unemployment is low as people are desperate to be employed in any kind of work, because remaining unemployed is not an option for many poverty-stricken households, due to the lack of passive income support schemes for the unemployed. This is one of the main reasons for depression of wages and is also a pointer to the fact that people seek any kind of jobs (even if they are low salaried), as long as they are unemployed and do not earn some income. With an increase in the household size, if an increased number of household members seek casual wage employment, then each worker simply works for less time than earlier and a major section of the workforce is underemployed and engaged in low productivity work. Thus, the earnings from these activities are on average quite low, highlighting that most workers are trapped in low paying activities.

Women are paid 28.5% less than men, *ceteris paribus*. After interacting with marital status, we see from Table 2 that married males enjoy a wage premium of around 14% whereas married females suffer a wage penalty of around -25%, which is additional to the gender gap. In other words, married women on average are paid 53.7% less than single men. In other words, married females are the most disadvantageous group in this case. Kingdon (1998) showed that although the rates of return for women

[16] When education qualifications are controlled as dummies, scholars typically utilise the following formula in calculating the yearly returns corresponding to a specific education qualification (Duraisamy, 2002). $r_i=(\beta_i-\beta_i)/(t_i-t_{i-1})$ where $r_i$ represents the yearly return rate to education qualification $edu_i$. $\beta_i$ and $\beta_{i-1}$ are the returns to education qualification at the *i*th and *i−1*th level. *t* and *t−1* are the years of schooling corresponding to their respective qualifications.

increased with levels of education, yet males had a higher rate of return and this is corroborated by our results.

The difference between religious groups is less dramatic. Turning to social groups, ceteris paribus schedule tribes have 7.6% higher earnings than other backward classes whereas scheduled castes have 6.3% lower earnings. These differences may be also due to unobserved differences of the workforce belonging to each tribe, which would be the case if the individuals belonging to the scheduled tribes were more motivated or skilled than the others.

Importantly, compared to those in the private sector, regular workers in the public sector earn 41.0% more. Regular workers in the public sector are also paid much higher than those in the third sector (41.0% vs. -5.4%, $p<0.01$). The public sector is on average a better employer than the private sector (or the third sector), but this depends also on the fact that all public sector jobs are similar due to the strong regulation of wages, whereas in the private sector there are very different conditions. However, employment opportunities in the public sector in India are very restrictive as it is extremely competitive, and availability is low. Entry to public sector jobs is possible by passing tough and highly competitive exams, which are either all India based or state specific. Entry conditions are quite tough thus making it impossible for general caste individuals to easily get a job. Preferential treatment (such as age relaxation, waiver of application fees etc.) is endowed on SC/STs. Caste based reservations for regular salaried work in the public sector, which is the most highly coveted category of jobs among the entire population, has been an important positive discriminatory labour market policy in India. Accordingly, Scheduled Castes have a 15% and Scheduled Tribes a 7.5% reservation in all public educational institutions and government or quasi-government jobs (which form the major part of all regular salaried jobs) (Das & Vasudeva-Dutta, 2007).

The story is similar for self-employed workers and for casual workers. We do note that the difference in returns to different levels of education become smaller. In other words, education contributes less to the wage increase for self-employed and casual workers. This is in line with the notion that the public sector is rewarding education more than any other sector where law enforcement is less binding. One natural conjecture is that the lower return rate in the private sector is a result of larger share of informal jobs in those sectors, because informal jobs typically pays lower than regular jobs. We have investigated the share of informal jobs in each sector. We find evidence in support of this conjecture. Specifically, 22.9% workers in the private sector work casually, whereas only 5.5% workers in the public sector are casual.

Notably, self-employed females suffer a wage penalty of 91.6% compared to self-employed men, due to the fact that a major share of women work in agriculture and allied sectors (37.62%), compared to men (29.45%). PLFS 2018-19 data states that a major share of the increase in female self-employment has taken place in rural agricultural own account workers (OAWs); a large section of the rural female workforce was thus affected by falling real incomes. PLFS 2018-19 Report states that the situation for female self-employed OAWs, is also worrying. Their earnings decreased over the period 2017-18 to 2018-19, for both rural and urban areas[17]. Interestingly enough, married males no longer enjoy a wage premium when working casually whereas widowed or divorced females enjoy around 50% wage premiums when they are self-employed. This is probably because they are the only breadwinner of the family. Casual workers in the public sector are paid much worse than those in the private sector. It seems that, if one chooses to work casually, then one should go for the private sector instead of the public sector. The monthly working hours for casual workers are statistically significant but economically minuscule - one additional hour relates to 0.5% wage increase.

[17] Gains were seen for female employers only. However, a miniscule fraction of the rural female workforce are self-employed employers.

Table 2: OLS regressions of ln monthly wage - general education

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Regular | se | Self-employed | se | Casual | se |
| Panel A: Original estimates | | | | | | |
| **A1: Education qualifications (baseline: no formal education)** | | | | | | |
| Literate below primary | 0.070*** | (0.021) | 0.037*** | (0.014) | 0.024** | (0.011) |
| Primary | 0.138*** | (0.015) | 0.077*** | (0.010) | 0.048*** | (0.008) |
| Middle school | 0.249*** | (0.014) | 0.171*** | (0.009) | 0.072*** | (0.008) |
| Secondary school | 0.340*** | (0.014) | 0.200*** | (0.010) | 0.065*** | (0.009) |
| Higher secondary school | 0.456*** | (0.014) | 0.254*** | (0.012) | 0.043*** | (0.012) |
| Graduate and diploma | 0.725*** | (0.014) | 0.427*** | (0.013) | 0.068*** | (0.022) |
| Postgraduate and above | 0.889*** | (0.017) | 0.565*** | (0.025) | 0.163** | (0.077) |
| A2: years of education | 0.055*** | (0.001) | 0.027*** | (0.001) | 0.006*** | (0.001) |
| Age | 0.023*** | (0.002) | 0.053*** | (0.002) | 0.012*** | (0.002) |
| Age square | -0.000*** | (0.000) | -0.001*** | (0.000) | -0.000*** | (0.000) |
| Rural (yes=1) | -0.190*** | (0.012) | -0.318*** | (0.007) | -0.142*** | (0.011) |
| Household size | -0.014*** | (0.001) | 0.017*** | (0.002) | -0.004*** | (0.001) |
| Gender (female=1) | -0.285*** | (0.016) | -0.916*** | (0.033) | -0.244*** | (0.025) |
| **Marital status (baseline: never married)** | | | | | | |
| Currently married | 0.139*** | (0.009) | 0.065*** | (0.012) | -0.010 | (0.009) |
| Widowed | -0.028 | (0.038) | -0.142*** | (0.029) | -0.084*** | (0.026) |
| Divorced/separated | 0.006 | (0.048) | -0.053 | (0.048) | -0.018 | (0.034) |
| Female X Currently married | -0.252*** | (0.019) | -0.042 | (0.035) | -0.085*** | (0.026) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Female X Widowed | -0.095** | (0.046) | 0.615*** | (0.048) | 0.056 | (0.037) |
| Female X Divorced/separated | -0.061 | (0.065) | 0.460*** | (0.092) | 0.002 | (0.060) |
| **Religion (baseline: Hinduism)** | | | | | | |
| Islam | -0.015* | (0.009) | -0.031*** | (0.008) | 0.045*** | (0.009) |
| Christianity | 0.083*** | (0.012) | 0.157*** | (0.014) | 0.130*** | (0.016) |
| Sikhism | 0.018 | (0.019) | 0.386*** | (0.021) | 0.035*** | (0.013) |
| Buddhism | 0.093*** | (0.023) | 0.144*** | (0.026) | -0.146*** | (0.017) |
| Others | 0.079*** | (0.027) | 0.127*** | (0.024) | 0.079*** | (0.024) |
| **Social groups (baseline: other backward class)** | | | | | | |
| Scheduled tribe | 0.076*** | (0.012) | -0.036*** | (0.010) | -0.102*** | (0.008) |
| Scheduled caste | -0.063*** | (0.009) | -0.083*** | (0.009) | -0.005 | (0.006) |
| Others | 0.090*** | (0.007) | 0.122*** | (0.007) | -0.039*** | (0.009) |
| **Household type (baseline: self-employed agriculture)** | | | | | | |
| Self-employed in:non-agriculture | 0.212*** | (0.011) | 0.089*** | (0.007) | 0.104*** | (0.014) |
| Regular wage/salary earning | 0.235*** | (0.014) | -0.242*** | (0.016) | 0.139*** | (0.013) |
| Casual labour in: agriculture | -0.039 | (0.041) | -0.323*** | (0.040) | 0.092*** | (0.011) |
| Casual labour in: non-agriculture | -0.042 | (0.032) | -0.295*** | (0.025) | 0.232*** | (0.011) |
| Others | 0.115*** | (0.027) | -0.079** | (0.033) | 0.144*** | (0.027) |
| **Sector (baseline: private)** | | | | | | |
| Public | 0.410*** | (0.007) | -0.128 | (0.149) | -0.178*** | (0.015) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Third | -0.054*** | (0.017) | 0.198*** | (0.063) | 0.110*** | (0.043) |
| Monthly working hour - causal worker | | | | | 0.005*** | (0.000) |
| R square | 0.422 | | 0.355 | | 0.496 | |
| Adjusted R square | 0.421 | | 0.355 | | 0.496 | |
| Robust SE | Yes | | Yes | | Yes | |
| N | 38996 | | 49329 | | 19502 | |

**Panel B: yearly returns to different education qualifications**

| | | | | | | |
|---|---|---|---|---|---|---|
| Literate below primary | 0.023*** | (0.007) | 0.012*** | (0.005) | 0.003 | (0.004) |
| Primary | 0.034*** | (0.010) | 0.020*** | (0.007) | 0.002 | (0.007) |
| Middle | 0.037*** | (0.004) | 0.031*** | (0.003) | 0.003 | (0.003) |
| Secondary | 0.045*** | (0.004) | 0.015*** | (0.004) | -0.000 | (0.005) |
| Higher Secondary | 0.058*** | (0.005) | 0.027*** | (0.005) | -0.014* | (0.008) |
| Graduate and Diploma | 0.090*** | (0.003) | 0.057*** | (0.004) | 0.006 | (0.009) |
| Post Graduate and above | 0.082*** | (0.006) | 0.069*** | (0.013) | 0.003 | (0.046) |

Notes: monthly working hour is only available for casual workers. Years of education are controlled in two alternative ways: in A1, education is controlled as categorical variable with different education qualification categories; in A2, education is controlled as continuous variable with years of education. Odd-numbered columns are coefficients and even-numbered columns are robust standard errors. *** p<0.01, ** p<0.05, * p<0.1

Table 3: OLS regressions of ln monthly wage - technical education

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Regular | se | Self-employed | se | Casual | se |
| Technical degree in: engineering/ technology | 0.426*** | (0.020) | 0.541*** | (0.056) | 0.135*** | (0.036) |
| Technical degree in: other subjects | 0.222*** | (0.021) | 0.306*** | (0.049) | 0.192** | (0.089) |
| Diploma or certificate (below graduate) in: | | | | | | |
| engineering/technology | 0.123*** | (0.017) | 0.180*** | (0.044) | 0.048 | (0.065) |
| other subjects | 0.109*** | (0.020) | 0.104*** | (0.038) | -0.011 | (0.065) |
| Diploma or certificate (graduate and above) | 0.155*** | (0.026) | 0.327*** | (0.053) | -0.154 | (0.112) |
| No. of years in Formal Education | 0.051*** | (0.001) | 0.025*** | (0.001) | 0.006*** | (0.001) |
| Age | 0.027*** | (0.002) | 0.053*** | (0.002) | 0.012*** | (0.002) |
| Age square | -0.000*** | (0.000) | -0.001*** | (0.000) | -0.000*** | (0.000) |
| Rural (yes=1) | -0.199*** | (0.013) | -0.320*** | (0.007) | -0.142*** | (0.011) |
| Household size | -0.013*** | (0.001) | 0.018*** | (0.002) | -0.004*** | (0.001) |
| Gender (female=1) | -0.264*** | (0.016) | -0.912*** | (0.033) | -0.246*** | (0.025) |
| Currently married | 0.135*** | (0.009) | 0.063*** | (0.012) | -0.008 | (0.009) |
| Widowed | -0.026 | (0.039) | -0.138*** | (0.029) | -0.083*** | (0.026) |
| Divorced/separated | 0.001 | (0.049) | -0.051 | (0.048) | -0.016 | (0.034) |
| Female X Currently married | -0.254*** | (0.019) | -0.041 | (0.035) | -0.084*** | (0.026) |
| Female X Widowed | -0.091** | (0.046) | 0.614*** | (0.047) | 0.056 | (0.037) |
| Female X Divorced/separated | -0.058 | (0.065) | 0.455*** | (0.093) | 0.002 | (0.060) |
| Islam | -0.014 | (0.009) | -0.033*** | (0.008) | 0.045*** | (0.009) |
| Christianity | 0.083*** | (0.012) | 0.151*** | (0.013) | 0.131*** | (0.016) |
| Sikhism | 0.017 | (0.019) | 0.382*** | (0.021) | 0.035*** | (0.013) |
| Buddhism | 0.081*** | (0.024) | 0.134*** | (0.026) | -0.148*** | (0.017) |
| Others | 0.085*** | (0.027) | 0.135*** | (0.024) | 0.077*** | (0.025) |
| Scheduled tribe | 0.079*** | (0.012) | -0.034*** | (0.010) | -0.102*** | (0.008) |
| Scheduled caste | -0.057*** | (0.009) | -0.081*** | (0.009) | -0.005 | (0.006) |
| Others | 0.094*** | (0.007) | 0.123*** | (0.007) | -0.038*** | (0.009) |
| Self-employed in:non-agriculture | 0.208*** | (0.011) | 0.085*** | (0.007) | 0.104*** | (0.014) |
| Regular wage/salary earning | 0.236*** | (0.014) | -0.248*** | (0.016) | 0.140*** | (0.013) |
| Casual labour in: agriculture | -0.033 | (0.041) | -0.324*** | (0.040) | 0.091*** | (0.011) |
| Casual labour in: non-agriculture | -0.043 | (0.032) | -0.298*** | (0.025) | 0.233*** | (0.011) |
| Others | 0.131*** | (0.027) | -0.071** | (0.033) | 0.142*** | (0.026) |
| Public | 0.414*** | (0.007) | -0.127 | (0.155) | -0.179*** | (0.015) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Third | -0.047*** | (0.017) | 0.214*** | (0.064) | 0.112*** | (0.042) |
| Monthly working hour - causal worker | | | | | 0.005*** | (0.000) |
| R square | 0.421 | | 0.355 | | 0.496 | |
| Robust SE | Yes | | Yes | | Yes | |
| N | 38984 | | 49323 | | 19504 | |

Notes: monthly working hour is only available for casual workers. *** p<0.01, ** p<0.05, * p<0.1

## 6.2 Subsample analysis

This section presents OLS Mincerian type of earnings estimates for different subsamples. The dependent variable in all the graphs is the natural logarithm of monthly wages for regular workers. Each dot represents the estimated coefficient. We start by comparing women vs. men, followed by rural vs. urban workers. Then we look at subsamples by different sector. Lastly, we dive into the different social groups (i.e., scheduled tribes, scheduled castes, other backward classes, and others), which are unique to India.

From Figure 1, we see that, although differences are not statistically significant, returns to men's education are slightly higher than those to women's education at the literate (5.9% vs. 1.4%, p=0.380) or primary level (11.1% vs. 9.8%, p=0.713). Yet, returns to men's education are significantly (p<0.01 in all cases) lower than that to women at the middle-school level (19.3% vs. 28.4%) and all the levels above (28.6% vs. 38.0% for secondary, 38.2% vs. 58.7% for higher secondary, 59.9% vs. 101.5% for graduate and diploma, 71.7% vs. 119.0% for postgraduate and above). In other words, the gender gap in returns to education in favour of women increases with the level of educational qualifications. These results are in line with previous findings (Psacharopoulos 1994; for the case of India, see Duraisamy and Malathy 1993; Duraisamy 2002) which suggests that higher levels of education benefit women more than men.

This finding should not be overlooked. It suggests, in particular, that, especially for women, higher levels of education would lead to higher wages, beyond the threshold of reservation wages. This suggests, in turn, that increasing the level of education of women would help drawing them more frequently into the labour force. According to Blau and Kahn (2017), from the human-capital perspective, women's rising labor-force participation is expected to raise the returns to their investment in higher education and thus to narrow the educational gender gap.

This finding is of great policy interest: girls should be encouraged to continue education, especially given the wage penalty faced by women in general. The higher returns to education of women are generally interpreted as a consequence of self-selection of the most talented and motivated women into employment: if the least motivated do not work, this is increasing the average returns to education by excluding from the estimates the least motivated and talented and, hence, those women who would earn a lower income. Only the best paid are included in the estimates. This is not the case for men, since almost all men work, also the least skilled and paid. Pastore and

Verashchagina (2006), in their study of Belarus, reported that the returns to education of men become higher than those of women, when total income is considered, suggesting that men fare better than women in secondary jobs and entrepreneurial activities. Their study has also reported that women receive higher wages from their main job than men across all educational groups, though the gap tends to reduce for individuals holding a high secondary school diploma.

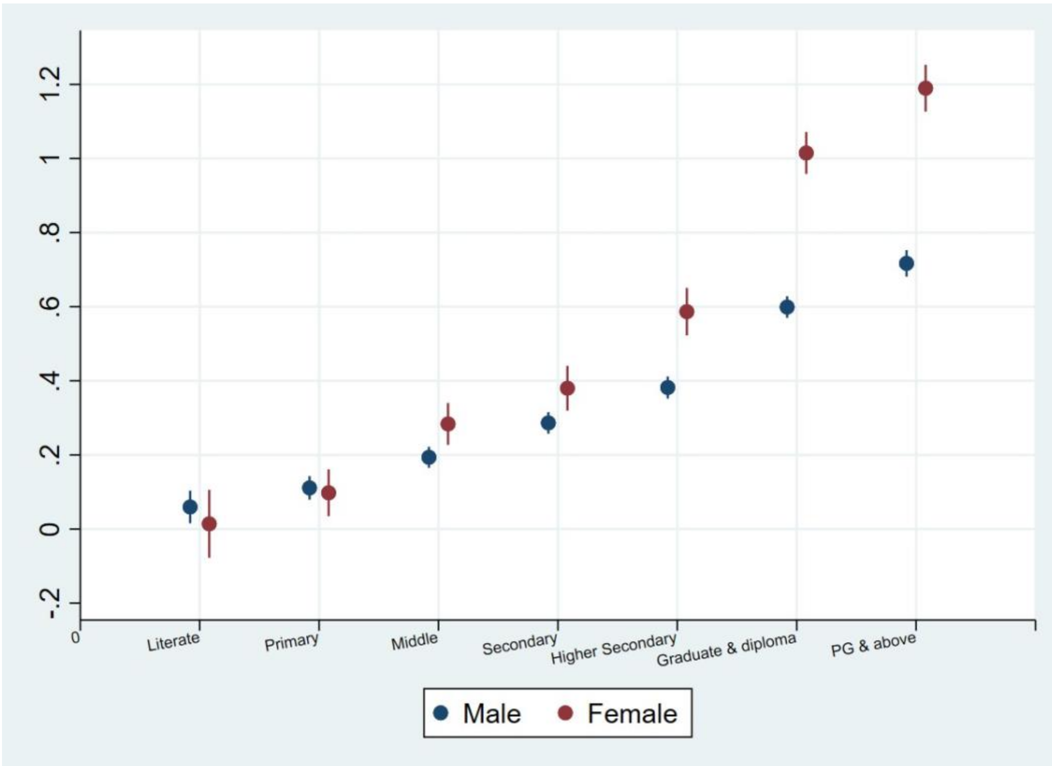



Figure 1: Subsample analysis - gender

Figure 2 reports the coefficients for the rural/urban subsample, respectively. It shows that urban workers are paid similarly (p>0.50 in all cases) to rural workers for higher secondary education (44.5% vs. 44.3%) and below (6.4% vs. 6.9% for literate, 13.9% vs. 12.0% for primary, 24.3% vs. 23.5% for middle, 34.6% vs. 31.6% for secondary). Yet, not surprisingly, urban workers are paid significantly higher than rural workers for graduate and diploma education (74.3% vs. 62.3%, p<0.01) and for postgraduate and above (89.4% vs. 75.1%, p<0.01). Duraisamy (2002) also reports that returns per year of schooling are, in general, significantly lower in rural areas than in the urban areas, but returns to lower level of education (primary, secondary) are higher in rural than urban areas (69% more for men and 32% more for women). Agarwal's study (2011) using IHDS data indicates that returns to primary education are lower for rural than for urban residents (e.g., 4.64% vs. 6.59%).

Discrepancy in the locational distribution of workers by activity status across rural and urban areas in India may be a possible explanation. Rural areas have a higher percentage (45.7%) of self-employed individuals, followed by regular salaried/wage employment (36.2%) and casual wage-employment (18.1%)[18]. In India casual wage employment and self-employment are generally considered unreliable employment categories as compared to regular salaried/wage employment, thus the rural areas seem to comprise a more vulnerable share of the workforce, who are entitled to lower returns to education. Role of vocational training may be considered in such an eventuality as it helps self-employed individuals earn a higher income (Bairagya 2021). We do not delve into the aspect of vocational education any further as this is beyond the scope of our paper.

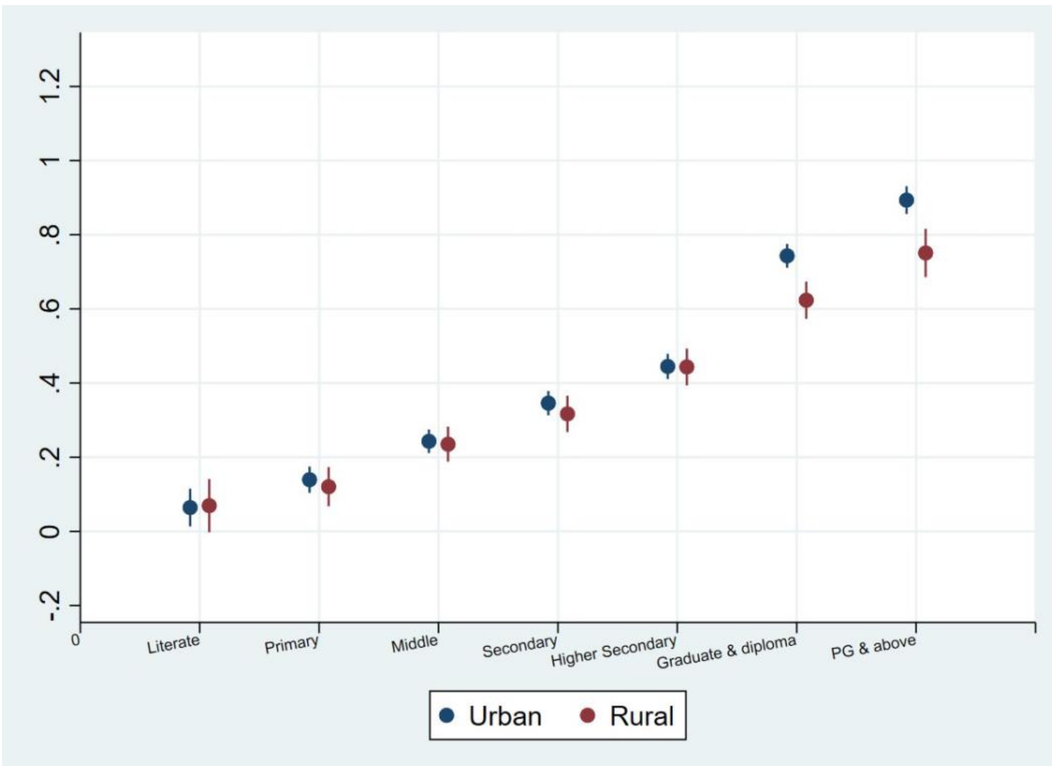



Figure 2: Subsample analysis - rural vs. urban

Figure 3 shows that the public sector pays the highest wages among the three sectors considered, in line with our observation in the OLS regression in Table 5. The confidence interval for the third sector is large mainly due to the relatively small sample number of observations (1361) compared to the private sector (108192) and the public sector (12665) (as can be seen from Table A3). When comparing the public sector with the private sector, the differences are not statistically significant (p>0.15) for lower levels of education (11.1% vs. 4.7% for literate, 5.8% vs. 13.2% for primary, 26.9% vs. 21.7% for middle). The differences become statistically significant for higher levels of education (38.3% vs. 30.0% for secondary (p=0.08), 58.3% vs. 36.6% for higher secondary (p=0.000), 78.5% vs. 68.1% for graduate and diploma (p=0.03), 94.6% vs. 84.5% for postgraduate and above (p=0.04)). In both the private and the public sectors, the rates of return to education increase with

[18] As mentioned previously in the Descriptive Statistics section.

consecutive higher levels of education, that is, they are the highest from graduate level onwards. When comparing the public sector with the third sector, the differences are never statistically significant despite the levels of education (11.1% vs. 21.0% for literate, 5.8% vs. 4.6% for primary, 26.9% vs. 21.1% for middle, 38.3% vs. 33.3% for secondary, 58.3% vs. 48.1% for higher secondary, 78.5% vs. 65.7% for graduate and diploma, 94.6% vs. 78.8% for postgraduate and above). Overall, higher levels of education seem to pay better to regular workers in the public sector, compared to those in the private sector. Although we do not observe statistically significant different returns between the public and the third sector, we conjecture this is a result of small sample size issue. Had we had more observations from the third sector, we could have seen that the public sector pays better than the third sector on average, as this is what we observed from Table 2 in the OLS regression. Validation of this conjecture is left to future research.

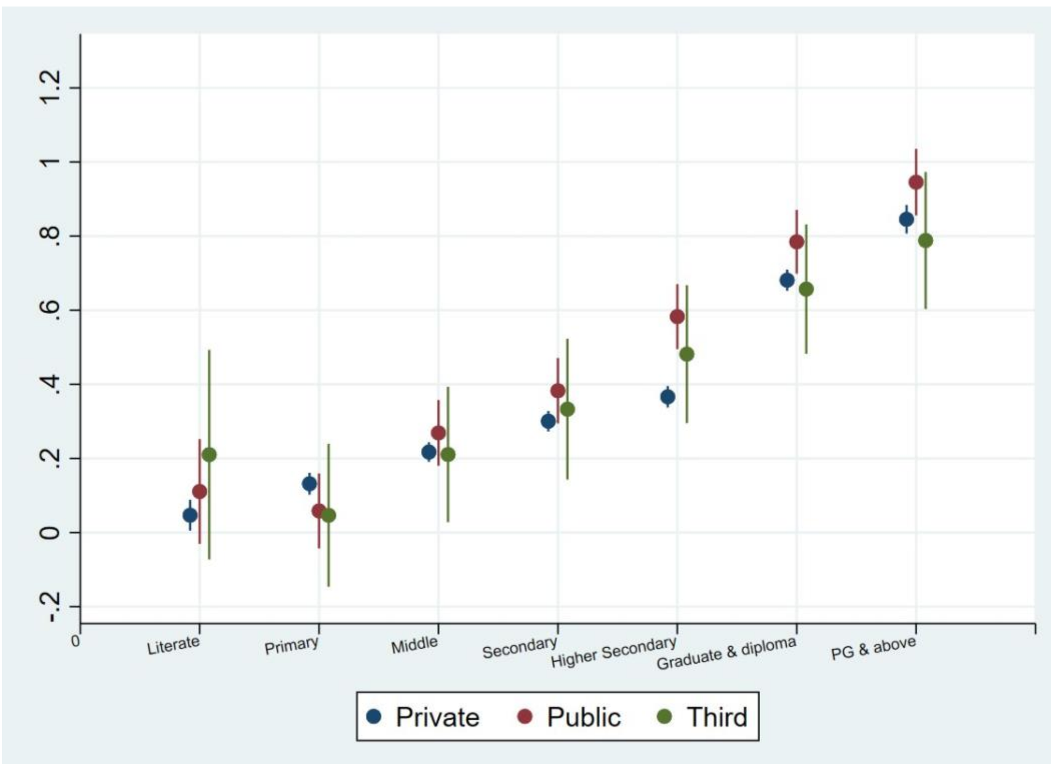



Figure 3: Subsample analysis- sector

Figure 4 reports the returns to workers relative to different castes. It shows that the returns to education relative to the scheduled tribe are the highest across all levels of education. To be more precise, we report the estimates by the levels of education across different social groups in Table 4. We also do pairwise coefficient difference tests among different social groups. Compared to OBC, regular workers from the ST are paid significantly higher in four out of the seven education levels (in the remaining three scenarios, the differences are not statistically significant. These are primary education, graduate and diploma, and postgraduate and above level of education).

The story between the SC and the ST is quite similar: workers from the ST are paid significantly higher in five out of the seven education levels (the two exceptions are primary education and postgraduate and above level of education). Between the ST and the fourth category (i.e., others), again, five out of the seven education levels are significantly different and the non-significant results are from the two highest levels of education.

By and large, returns to the scheduled tribe are the highest among all social groups. Our findings are corroborated by literature, as Sikdar (2019), found that although few numbers of STs are able to achieve higher education, yet when they do so they are successful in getting regular salaried jobs. Intuitively, since scheduled tribes are the more disadvantaged social group in India, individuals from the scheduled tribe would face limited job opportunities had they not been educated. In other words, education opens up more opportunities for the disadvantaged group, be it workers from the scheduled tribe or female workers. Borooah and Iyer (2005) found in their study that in a favourable situation, individuals from a backward caste will have high rates of return. This points towards self-selection of the most motivated and talented individuals into employment of people from the scheduled tribes. Although our results from the Heckman selection model do not provide statistically significant (Chi-square=2.15, p=0.143) results in support of this conjecture, the lack of statistical significance could be because of model specification issues. It is noted that the results from Heckman selection model are susceptible to correct model specification (Heckman, 1979). For shortness' sake, we do not report the Heckman estimates, which are however available on request.

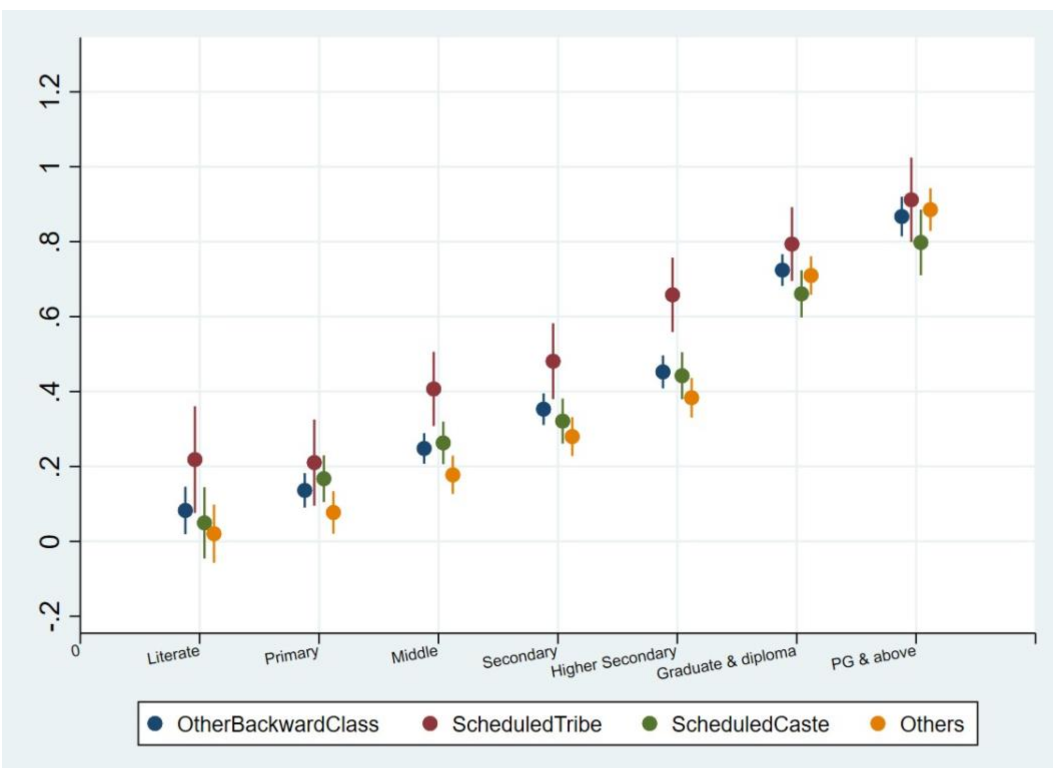



Figure 4: Subsample analysis -social group

Table 4 Returns to education by social groups

| | Literate | Primary | Middle | Secondary | Higher Secondary | Graduate & diploma | PG & above |
|---|---|---|---|---|---|---|---|
| OBC | 0.082** | 0.136*** | 0.248*** | 0.353*** | 0.452*** | 0.724*** | 0.867*** |
| | (0.032) | (0.023) | (0.021) | (0.022) | (0.023) | (0.022) | (0.027) |
| ST | 0.218*** | 0.210*** | 0.407*** | 0.481*** | 0.658*** | 0.794*** | 0.912*** |
| | (0.073) | (0.059) | (0.051) | (0.052) | (0.051) | (0.050) | (0.057) |
| SC | 0.049 | 0.167*** | 0.263*** | 0.321*** | 0.442*** | 0.661*** | 0.798*** |
| | (0.049) | (0.032) | (0.029) | (0.031) | (0.032) | (0.032) | (0.045) |
| OT | 0.020 | 0.077*** | 0.177*** | 0.280*** | 0.384*** | 0.710*** | 0.886*** |
| | (0.040) | (0.029) | (0.026) | (0.027) | (0.027) | (0.026) | (0.029) |

Notes: OBC=Other Backward Class, ST=Scheduled Tribe, SC=Scheduled Caste, OT=others

## 6.3Causal inference attempts and robustness checks

As explained in the Methodology section, to identify causal relationships, two methods are attempted for regular workers. We choose regular workers because it is the most standard group, compared to self-employed workers or casual workers. We first try the conventional IV approach. We experiment with several instrumental variables, but none fully satisfied either the relevance or the exclusion restriction. Following a previous study on China (Chen and Pastore, 2021), we construct the IV by calculating the proportion of people who completed elementary (or compulsory, alternatively) education in each of the 36 regions in India. With one instrumental variable, we can have maximum one instrumented variable. This means the education dummies are inappropriate in this case. Thus, we use years of education instead. The first stage F statistics in all specifications are greater than 300, indicating we have rather strong IVs. Arguably, the IVs are exogenous given that each individual's education choice unlikely affects the regional education rate. Looking at the IV results using either IVs in Table 5, we can have two conclusions. First, the positive returns to education is robust. Second, the IV results did not improve the OLS results much, as coefficients are even inflated to a large extent, which is contrary to the performance of a good IV that coefficients get lower (Card, 1999). Clearly, the IV results are not very helpful.

We then try the more recent heteroskedasticity-based Lewbel method (Lewbel, 2012; Baum and Lewbel, 2019). The intuition of the Lewbel method is as follows. Instruments are constructed based on the heteroskedasticity in the error term. Heteroskedasticity can be tested via the Breusch-Pagan test. In our case, the BP tests always return Chi squares greater than 100, giving us strong confidence that the heteroskedastity assumption is satisfied, which ensures the correlation between the endogenous regressor and the constructed instrument. From Table 5, we see that the Lewbel estimates are rather close to the OLS estimates. Again, there is not much improvement in the precision of the estimates. Nonetheless, both the IV estimates and the Lewbel estimates serve as robustness checks to the OLS estimates. Additionally, our attempts may inform other researchers' future exploration on similar topics.

Table 5: OLS, IV, and Lewbel estimates for regular workers

| | OLS (1) | IV (2) | Lewbel (3) | IV (4) | Lewbel (5) |
|---|---|---|---|---|---|
| Years of education | 0.055*** (0.001) | 1.571* (0.860) | 0.057*** (0.002) | 0.116*** (0.014) | 0.058*** (0.002) |

Notes: the IV in columns (2) and (3) is the proportion of people who completed elementary education in each region. The IV in columns (4) and (5) is the proportion of people who completed compulsory education. *** p<0.01, ** p<0.05, * p<0.1

## 7. Summary remarks

In this paper, we utilise the Periodic Labour Force Survey (PLFS) for 2018–2019 to estimate the returns to education in India, updating the existing literature which covered the period up to 2011, about a decade earlier. To the best of our knowledge, our paper is the first to estimate the returns to general and technical education using this wave of the PLFS data. We find that, compared to regular workers with no formal education, an additional year of literacy education raises average regular workers' earnings by 2.3%, primary education about 3.4%, middle school education about 3.7%, secondary school education about 4.5%, higher secondary education about 5.8%, graduate and diploma about 9.8%, post graduate and above level of education about 8.2%. Compared to Duraisamy (2002), our calculated return rates overall are smaller in magnitude, indicating a common trend as larger share of a population are better educated. In other words, an increase in the supply of educated worker likely leads to a decrease in the returns to education, *ceteris paribus*.

A notable pattern is that regular workers' yearly return to education increases with the level of education qualifications, which is against the hypothesis of diminishing returns to education. Similar patterns are observed for self-employed workers, but casual workers do not seem to benefit as much from education. These estimates are largely in line with the previous literature relative to other years, though we are the first to provide estimates for regular workers, self-employed workers, and casual workers, separately. Previous studies either did not mention this caveat (Mendiratta and Gupt-Arthaniti, 2013) or could provide estimates for regular workers only (Duraisamy, 2001) or left self-employed workers out of the conversation (Vasudeva Dutta, 2006). Meanwhile, we are the first to provide estimates for literacy-level education and to separately estimate the returns for college education and postgraduate education. These additional layers enable us to estimate the return rate more accurately. For example, unsurprisingly, after separating college education and postgraduate education, our estimates of the returns to college education is 9%, lower than the 11.7% in Duraisamy (2001).

Additionally, we are among the first to estimate the returns to technical education, controlling for general education. We find that, for regular workers, technical degree in engineering / technology pays the best (42.6%) among all the technical degrees. Interestingly, among those with technical degree in engineering / technology, self-employed workers are paid even better than regular workers (54.1% vs. 42.6%). This is possibly because the more capable engineers / technicians self-select to be self-employed. Compared to general education, obtaining technical degree in engineering / technology is roughly equivalent to obtaining higher secondary education degree.

By and large, regular workers are better paid compared to those who are self-employed or work casually. Education levels higher than compulsory secondary schooling cause an increase in propensity to take part in paid work. This is because the returns to education are insignificant and low for lower levels of education. The returns increase significantly along with the increase in educational levels. Such increasing pattern of private rates of return imply that for an individual, it is profitable to invest in higher education.

Additionally, we estimate the returns to different fields of study and, in particular, for technical education. Understandably, technical education in engineering or technology pays the best among all the technical education qualifications. Interestingly though, for people with technical education qualifications, those who are self-employed are paid better than those who are regular workers. One possibility is the self-selection into different career paths. Especially given the unprecedented technological advancement, people with higher technical or engineering skills are more likely to start up a company and seize greater benefit, compared to those with lower skills.

We have also identified some interesting patterns in analyzing the heterogeneity across different subsamples. Substantial heterogeneity in returns is observed; for example, estimates are higher for females compared with their male counterparts. Starting from middle-school up through postgraduate and above levels of education, females enjoy significantly, both economically and statistically, higher returns to education than males: 28.4% vs. 19.3% at middle-school level, 38.0% vs. 28.6% for secondary, 58.7% vs. 38.2% for higher secondary, 101.5% vs. 59.9% for graduate and diploma, 119.0% vs. 71.7% for postgraduate and above.

Apparently, the return gap between women and men increases with the level of educational qualifications. Most prominently, the more educated women are, the greater the wage premium they can enjoy, proving that the gains from bringing more women into the workforce may be considerable since only a small fraction of women work for wages but their education has increased substantially. Although it is possible that more talented women self-select to continue education, women in general can still benefit from policy intervention that encourages them to continue education.

Education does not seem to contribute much to narrow the rural-urban differences. For higher secondary education and levels below, the differences in returns are not statistically significantly different. Notably, urban workers are paid better than rural workers if both groups complete graduate and diploma education (74.3% vs. 62.3%, $p<0.01$) or postgraduate and above level of education (89.4% vs. 75.1%, $p<0.01$). This is saying that education is not helping rural workers catch up with their urban peers. If anything, the earning gap is enlarged for those who complete higher levels of education. This calls for attention towards enhancing the role of vocational education in rural areas. There is a need for an appropriate policy attention towards enhancing the participation of rural individuals (majority of whom are self-employed) in the formal vocational training programs.

People from scheduled tribes enjoy the highest returns to education among all the social groups, at every education level, proving that education can greatly improve their situation. But it is notable that wage data captures only that section of STs who are actually in wage employment. Presumably that is a miniscule section, which are the 'creamy layer' and once employed they get wages at par with the other castes. They might have availed of the benefit of reservations over successive generations and presently may have achieved success in their respective areas. This is the classic case where the elites within the group have monopolized the gains from employment.

Education is a powerful tool for upward social mobility, which can help to build an inclusive society by reducing socioeconomic disparities. Our finding that education level below primary has the lowest return implies that universalization of elementary education alone will not suffice in the economy because modern industry needs higher and specialised education, thus making it imperative to improve secondary and above education levels. On the one hand, in urban areas and for some special groups (women and some castes) it is important to increase the share with tertiary education to reach equality within the group; on the other hand in most of the country, especially in rural areas and peripheral regions, it is important to increase basic levels of education. In a developing country like India, returns to education in the labour market may be of limited use because most of the workforce does not participate in the formal labour market. Hence, a policy of solely increasing education among masses is insufficient if not complimented with other changes in the Indian economy to promote widespread wage employment. Also it would be important to promote regular employment, as a means to guarantee better paid jobs to the largest number of individuals and better wages for women.

## References


Agrawal, T. (2011). Returns to education in india: Some recent evidence, indira gandhi institute of development research. Technical report, Working Paper 2011–017, Mumbai.

Asadullah, M Niaz (2006) "Returns to Education in Bangladesh," Education Economics, 14(4), 453-468.

Asadullah, M. N. and Xiao, S. (2020). The changing pattern of wage returns to education in post-reform china. Structural Change and Economic Dynamics, 53:137-148.

Azam, M. (2012). Changes in Wage Structure in Urban India, 1983–2004: A Quantile Regression Decomposition, *World Development*, 4(6): 1135-1150, https://doi.org/10.1016/j.worlddev.2012.02.002

Bahl, S., Bhatt, V. and Sharma, A. (2021), "Returns to formal and informal vocational education and training in India", International Journal of Manpower, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/IJM-04-2021-0211

Bairagya, I. (2021), Impact of formal vocational training on the earnings of self-employed individuals in rural India, *Journal of Business Venturing Insights*, Volume 16, e00269, ISSN 2352-6734, https://doi.org/10.1016/j.jbvi.2021.e00269

Baum, C. F., & Lewbel, A. (2019). Advice on using heteroskedasticity-based identification. The Stata Journal, 19(4), 757-767.

Becker, G. (1964). Human capital (new york: National bu-reau of economic research, 1964). *BeckerHuman Capital1964*.

Bennell, P. (1996). Rates of return to education: Does the conventional pattern prevail in subsaharan africa? *World development*, 24(1):183–199.

Bhandari, L. and Bordoloi, M. (2006). Income differentials and returns to education. *Economic and Political Weekly*, pages 3893–3900.

Blau, F. D and Kahn, L. M. (2017), The Gender Wage Gap: Extent, Trends, and Explanations, Journal of Economic Literature, 55(3), 789–865

Borooah, V & Iyer, S. ( 2005). "Vidya, Veda, and Varna: The influence of religion and caste on education in rural India," Journal of Development Studies, Taylor & Francis Journals, vol. 41(8), pages 1369-1404.

Card, D. (1999). The causal effect of education on earnings. Handbook of labor economics, 3, 1801-1863.

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

Chauhan, C. P. S.(2008) Education and caste in India, Asia Pacific Journal of Education, 28:3, 217-234, https://doi.org/10.1080/02188790802267332

Chen, J. and Pastore, F. (2021). "study hard and make progress every day": Updates on returns to education in China. Technical report, Global Labor Organization (GLO).

Das, M. B., and Dutta, P. V. (2007). Does caste matter for wages in the Indian labor market. *Washington, DC, USA: The World Bank*.

Drèze, J and Sen. A (1991), Hunger and Public Action, DOI: 10.1093/0198283652.001.0001 https://oxford.universitypressscholarship.com/view/10.1093/0198283652.001.0001/acprof-9780198283652-chapter-11

Duraisamy, M., and P. Duraisamy. 1993. Returns to scientific and technical education in India. Margin*,* 7(1): 396–406

Duraisamy, P. (2002). Changes in returns to education in india, 1983–94: by gender, age-cohort and location. *Economics of Education Review*, 21(6):609–622.

Fasih, T., Kingdon, G., and Patrinos, H. (2012). Heterogeneous returns to education in the labor market.

Foster, A. D., & Rosenzweig, M. R. (1996). Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution. *The American Economic Review*, *86*(4), 931–953. http://www.jstor.org/stable/2118312

Fulford, S. (2014). Returns to Education in India, *World Development*, 59 :434-450, https://doi.org/10.1016/j.worlddev.2014.02.005

Geetha Rani, P. (2014). Disparities in earnings and education in india. *Cogent Economics & Finance*, 2(1):941510.

Ghurye, G. S. (1969). *Caste and race in India*. Popular Prakashan.

Government of India (2020), Annual Report: Periodic Labour Force Survey 2018-19, National Sample Survey Office, New Delhi.

Halvorsen, R., R. Palmquist, et al. (1980). The interpretation of dummy variables in semilogarithmic equations.American economic review 70(3), 474–475.

Heckman, J. 1979. Sample selection bias as a specification error. Econometrica 47: 153--161.

Kanjilal-Bhaduri S. and Pastore, F. (2018): 'Returns To Education And Female Participation Nexus: Evidence From India', Indian Journal of Labour Economics; September 2018, Volume 61, Issue 3, pp 515–536.

Kijima, Y. (2006), Why did wage inequality increase? Evidence from urban India 1983–99, *Journal of Development Economics*, 81 (1), Pages 97-117, ISSN 0304-3878, https://doi.org/10.1016/j.jdeveco.2005.04.008.

Kingdon, G. G (1998) Does the labour market explain lower female schooling in India?, The Journal of Development Studies, 35:1, 39-65, DOI: 10.1080/00220389808422554

Kingdon, G., and N. Theopold. 2008. Do returns to education matter to schooling participation? Evidence from India. Education Economics 16(4): 329–350

Klasen, S., and J. Pieters. 2012. Push or pull? Drivers of female labor force participation during India's economic boom. IZA Discussion Paper No. 6395.

Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. Journal of Business & Economic Statistics, 30(1), 67-80.

Lutz, W., Cuaresma, J. C., Kebede, E., Prskawetz, A., Sanderson, W. C., and Striessnig, E. (2019). Education rather than age structure brings demographic dividend. *Proceedings of the National Academy of Sciences*, 116(26):12798–12803.

Madheswaran, S., & Attewell, P. (2007). Caste discrimination in the Indian urban labour market: Evidence from the National Sample Survey. *Economic and Political Weekly*, *42*(41), 4146–4153.

Mankiw, N. G., Romer, D., and Weil, D. N. (1992). A contribution to the empirics of economic growth. *The quarterly journal of economics*, 107(2):407–437.

Mendiratta, P. and Gupt, Y. (2013). Private returns to education in india by gender and location: A pseudo-panel approach. *Arthaniti-Journal of Economic Theory and Practice*, 12(1-2):48–69.

Mitra, A. (2019). Returns to education in india: Capturing the heterogeneity. *Asia & the Pacific Policy Studies*, 6(2):151–169.

Mok, K. H. and Marginson, S. (2021). Massification, diversification and internationalisation of higher education in china: Critical reflections of developments in the last two decades. *International Journal of Educational Development*, 84:102405.

Montenegro, C. E and Patrinos, H. A. (2021). A data set of comparable estimates of the private rate of return to schooling in the world, 1970–2014, International Journal of Manpower, Vol. ahead-of-print No. ahead-of-print. DOI 10.1108/IJM-03-2021-0184

OECD (2019) Education at a Glance 2019, Country Note, India, https://www.oecd.org/education/education-at-a-glance/EAG2019_CN_IND.pdf#:~:text=India%20Education%20at%20a%20Glance%3A%20OECD%20Indicators%20%28OECD%2C,of%20education%20systems%20in%20OECD%20and%20partner%20countries.

OECD (2021), Education at a Glance: OECD Indicators, India, https://www.oecd-ilibrary.org/sites/b304737b-en/index.html?itemId=/content/component/b304737b-en

Pastore, F. and Verashchagina, A. (2006). Private returns to human capital over transition: A case study of Belarus. Economics of Education Review, 25(1):91–107.

Psacharopoulos, G. (1985). Returns to education: A further international update and implications. *Journal of Human resources*, pages 583–604.

Psacharopoulos, G. 1994. Returns to investment in education: A global update. World Development, 22(9): 1325–1343

Psacharopoulos, G. and Patrinos, H. A. (2002). Returns to investment in education: A further update. *Available at SSRN 367780*.

Psacharopoulos, G. and Patrinos, H. A. (2004). Returns to investment in education: a further update. *Education economics*, 12(2):111–134.

Raveendran, G. 2016. The Indian labour market: A gender perspective. Discussion paper-8, Prepared for Progress of the World's Women 2015–16. UN Women.

Schultz, T. W. (1960). Capital formation by education. *Journal of political economy*, 68(6):571–583.

Sianesi, B. and Van Reenen, J. (2000). The returns to education: a review of the macro-economic literature.

Sikdar, S. (2019). Rate of return to education in India: Some insights. Working Paper No. 270. National Institute of Public Finance and Policy

Thorat. S & Newman. K.S. (2007), Caste and Economic Discrimination: Causes, Consequenses and Remedies, Economic and Political Weekly, Vol. 42, No. 41 (Oct 13-19,2007), Pages 4121-4124

Vasudeva Dutta, P. (2006). Returns to education: New evidence for india, 1983–1999. *Education Economics*, 14(4):431–451.

Waughray, A. (2010). Caste discrimination and minority rights: The case of india's dalits. *International Journal on Minority and Group Rights*, 17(2):327–353.

World Development Report. 2012. Gender equality and development.

www.statista.com › statistics › 1001039India - unemployment rate by education level 2019 | Statista

## Annex

**Table A1: Definitions of variables used in OLS Regression Equations**

| Variable | Definition |
|---|---|
| **Dependent Variables** | |
| Log Monthly wage - regular worker | Log of total earnings during the preceding calendar month for regular/salaried wage activity in INR |
| Log Monthly wage - self-employed worker | Log of gross earnings during the last 30 days for self-employment activity in INR |
| Log Monthly wage - casual worker | Log of Monthly total wage earnings of casual workers in INR |
| **Independent Variables** | |
| Monthly working hour - casual worker | Total hours actually worked during the month |
| **Personal Variables** | |
| Age | Age in years |
| Age Square | Square of age |
| Female | Gender dummy; male=0, female=1 |
| Male | Gender dummy: male=1, female=0 |
| Never married | Marital Status dummy; never married=1,<br>married, divorced/separated, widowed=0 |
| Currently married | Marital Status dummy; married=1,<br>never married, divorced/separated, widowed=0 |
| Widowed | Marital Status dummy; widowed=1, |

| | |
|---|---|
| | never married, married, divorced/separated=0 |
| Divorced/Separated | Marital Status dummy; married=1,<br><br>never married, divorced, widowed, separated=0 |
| Female Currently Married | Interaction of Female*Currently Married |
| Female Widowed | Interaction of Female*Widowed |
| Female Divorced/ Separated | Interaction of Female*Divorced/Separated |
| General Education | Number of years of education (as defined in table A2) |
| No formal schooling | Years of education gained=1; yes=1, no=0 |
| Literate below primary | Years of education gained=3; yes=1, no=0 |
| Primary | Years of education gained=5; yes=1, no=0 |
| Middle school | Years of education gained=8; yes=1, no=0 |
| Secondary school | Years of education gained=10; yes=1, no=0 |
| Higher secondary school | Years of education gained=12; yes=1, no=0 |
| Graduate and diploma | Years of education gained=15; yes=1, no=0 |
| Postgraduate and above | Years of education gained=17; yes=1, no=0 |
| Technical education | Years of technical education (as defined in table A4) |
| No technical education | Years of technical education gained=0; yes=1, no=0 |
| Technical degree in: engineering/ technology | Years of education gained=18; yes=1, no=0 |
| Technical degree in: other subjects | Years of education gained=18; yes=1, no=0 |

| | |
|---|---|
| Diploma or certificate (below graduate level) in: engineering/technology | Years of education gained=15; yes=1, no=0 |
| Diploma or certificate (below graduate level) in: other subjects | Years of education gained=15; yes=1, no=0 |
| Diploma or certificate (graduate and above level) | Years of education gained=17; yes=1, no=0 |
| **Demographic variables** | |
| Household size | Number of members of the household |
| ***Household type*** | |
| Self-employed in:agriculture | Single major source of income of household for the last 365 days is from self-employment in agricultural activities; yes=1, no=0 |
| Self-employed in:non-agriculture | Single major source of income of household for the last 365 days is from self-employment in non-agricultural activities; yes=1, no=0 |
| Regular wage/salary earning | Single major source of income of household for the last 365 days is from regular/salaried employment; yes=1, no=0 |
| Casual labour in: agriculture | Single major source of income of household for the last 365 days is from casual labour in agriculture; yes=1, no=0 |
| Casual labour in: non-agriculture | Single major source of income of household for the last 365 days is from casual labour in non-agriculture; yes=1, no=0 |
| Others | Household which does not have any income from economic activities; yes=1, no=0 |
| ***Religion*** | |

| | |
|---|---|
| Hinduism | Religion dummy; yes=1, no=0 |
| Islam | Religion dummy; yes=1, no=0 |
| Christianity | Religion dummy; yes=1, no=0 |
| Sikhism | Religion dummy; yes=1, no=0 |
| Buddhism | Religion dummy; yes=1, no=0 |
| Others | Religion dummy; yes=1, no=0 |
| ***Social Groups*** | |
| Scheduled tribe | Social group dummy; yes=1, no=0 |
| Scheduled caste | Social group dummy; yes=1, no=0 |
| Other backward class | Social group dummy; yes=1, no=0 |
| Others | Social group dummy; yes=1, no=0 |
| ***Location*** | |
| Rural | Location dummy; yes=1, no=0 |
| Urban | Location dummy; yes=1, no=0 |
| ***Sector*** | |
| Private | Enterprise type dummy; yes=1, no=0 |
| Public | Enterprise type dummy; yes=1, no=0 |
| Non-Profit Organisation (Third sector) | Enterprise type dummy; yes=1, no=0 |

Table A2: Descriptive Statistics

| Variable | Mean | SD | Max. | Min. | N |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| Monthly wage - regular worker | 4496.084 | 8520.799 | 41300 | 0 | 122290 |
| Monthly wage - self-employed worker | 4511.730 | 7581.425 | 61000 | 0 | 122290 |
| Monthly wage - casual worker | 888.511 | 2231.746 | 10208.57 | 0 | 122290 |
| Ln (Monthly wage - regular worker) | 9.319 | (0.712) | 10.629 | 4.605 | 39024 |
| Ln (Monthly wage - self-employed worker) | 9.064 | (0.771) | 11.019 | 4.605 | 49370 |
| Ln (Monthly wage - casual worker) | 8.518 | (0.504) | 9.231 | 5.144 | 19518 |
| Gender (female=1) | 0.235 | 0.424 | 1 | 0 | 122268 |
| Age | 37.669 | 10.646 | 59 | 15 | 122290 |
| Years of education | 8.396 | 4.958 | 28 | 0 | 122290 |
| Rural (yes=1) | 0.582 | 0.493 | 1 | 0 | 122290 |
| Household size | 4.778 | 2.044 | 21 | 1 | 122290 |
| Marital status | | | | | |
| Never married | 0.182 | 0.386 | 1 | 0 | 122290 |
| Currently married | 0.775 | 0.418 | 1 | 0 | 122290 |
| Widowed | 0.036 | 0.186 | 1 | 0 | 122290 |
| Divorced/Separated | 0.007 | 0.084 | 1 | 0 | 122290 |
| Religion | | | | | |
| Hinduism | 0.753 | 0.431 | 1 | 0 | 122290 |
| Islam | 0.133 | 0.339 | 1 | 0 | 122290 |
| Christianity | 0.070 | 0.256 | 1 | 0 | 122290 |
| Sikhism | 0.020 | 0.141 | 1 | 0 | 122290 |
| Buddhism | 0.013 | 0.112 | 1 | 0 | 122290 |
| Others | 0.011 | 0.105 | 1 | 0 | 122290 |
| Social groups | | | | | |
| Scheduled tribe | 0.142 | 0.349 | 1 | 0 | 122290 |
| Scheduled caste | 0.169 | 0.374 | 1 | 0 | 122290 |
| Other backward class | 0.396 | 0.489 | 1 | 0 | 122290 |
| Others | 0.293 | 0.455 | 1 | 0 | 122290 |
| Household type | | | | | |
| Self-employed in:agriculture | 0.418 | 0.493 | 1 | 0 | 122290 |
| Self-employed in:non-agriculture | 0.309 | 0.462 | 1 | 0 | 122290 |
| Regular wage/salary earning | 0.149 | 0.356 | 1 | 0 | 122290 |
| Casual labour in: agriculture | 0.052 | 0.222 | 1 | 0 | 122290 |
| Casual labour in: non-agriculture | 0.060 | 0.238 | 1 | 0 | 122290 |
| Others | 0.011 | 0.104 | 1 | 0 | 122290 |
| Technical education No technical education | 0.960 | 0.197 | 1 | 0 | 122194 |
| Technical degree in: engineering/ technology | 0.007 | 0.081 | 1 | 0 | 122194 |
| Technical degree in: other subjects | 0.008 | 0.089 | 1 | 0 | 122194 |
| Diploma or certificate (below graduate level) in: engineering/technology | 0.010 | 0.098 | 1 | 0 | 122194 |

| | | | | | |
|---|---|---|---|---|---|
| Diploma or certificate (below graduate level) in: other subjects | 0.010 | 0.099 | 1 | 0 | 122194 |
| Diploma or certificate (graduate and above level) | 0.006 | 0.078 | 1 | 0 | 122194 |
| General education No formal schooling | 0.158 | 0.365 | 1 | 0 | 122218 |
| Literate below primary (3) | 0.046 | 0.209 | 1 | 0 | 122218 |
| Primary (5) | 0.124 | 0.330 | 1 | 0 | 122218 |
| Middle school (8) | 0.233 | 0.423 | 1 | 0 | 122218 |
| Secondary school (10) | 0.159 | 0.366 | 1 | 0 | 122218 |
| Higher secondary school (12) | 0.116 | 0.320 | 1 | 0 | 122218 |
| Graduate and diploma (15) | 0.128 | 0.334 | 1 | 0 | 122218 |
| Postgraduate and above (17) | 0.035 | 0.184 | 1 | 0 | 122218 |
| Sector | | | | | |
| Private | 0.885 | 0.319 | 1 | 0 | 122290 |
| Public | 0.104 | 0.305 | 1 | 0 | 122290 |
| Non-Profit (Third sector) | 0.011 | 0.105 | 1 | 0 | 122290 |

Notes: the data include waged workers (age range 15-59) from the 2018-2019 Periodic Labour Force Survey in India. Variable categories with a proportion below 0.005 are collapsed into one.

Table A3: Educational representation among Public Sector, Private Sector and Non-Profit Organisation (third sector)

| | Sector | | | | | | |
|---|---|---|---|---|---|---|---|
| | Private | | Public | | Non-Profit Organisations | | |
| Education Levels | Freq. | % | Freq. | % | Freq. | % | Total |
| No formal schooling | 18708 | 0.173 | 554 | 0.044 | 48 | 0.035 | 0.158 |
| Literate, but below primary | 5357 | 0.050 | 214 | 0.017 | 25 | 0.018 | 0.046 |
| Primary | 14479 | 0.134 | 638 | 0.050 | 84 | 0.062 | 0.124 |
| Middle | 26413 | 0.244 | 1887 | 0.149 | 216 | 0.159 | 0.233 |
| Secondary | 17292 | 0.160 | 1991 | 0.157 | 143 | 0.105 | 0.159 |
| Higher Secondary | 11777 | 0.109 | 2270 | 0.179 | 156 | 0.115 | 0.116 |
| Graduate and Diploma | 11447 | 0.106 | 3787 | 0.299 | 442 | 0.325 | 0.128 |
| Post Graduate and above | 2719 | 0.025 | 1324 | 0.105 | 247 | 0.181 | 0.035 |
| Total | 108192 | 1 | 12665 | 1 | 1361 | 1 | 1 |

**Table A4: Transformation of education coding to years of education**

| **General Educational Attainment Code** | NSS Code | Imputed Years of education |
|---|---|---|
| Not Literate | 1 | 0 |
| Literate through attending NFEC/AEC, TLC or others | 2,3,4 | 1 |
| Literate, but below primary | 5 | 3 |

| | | |
|---|---|---|
| Primary | 6 | 5 |
| Middle | 7 | 8 |
| Secondary | 8 | 10 |
| Higher Secondary | 10 | 12 |
| Graduate and Diploma | 11,12 | 15 |
| Post Graduate and above | 13 | 17 |
| **Technical Educational Attainment Code** | | |
| No technical Education | 1 | 0 |
| Degree | 2 | 18 |
| Diploma/certificate (below graduate) | 3,4,5,6,7 | 15 |

Source: Kingdon and Leopold, 2008; Table 1. Note: NFEC = Non-Formal Education Centre, TLC = Total Literacy Campaign, AEC = Alternative Education Centre

## Appendix

### System of Caste in India

India is traditionally deeply stratified on the basis of caste and religion. As a system of hereditary social stratification, caste (Social Groups in our dataset) is associated primarily with India, where it has existed for over three thousand years. The first mention of the concept of caste is found in the literary records of the Indo–Aryan culture of the Gangetic plain (Ghurye, 1969). The Indian caste factor has strong historical roots as it was built up with a recognition of variations in human nature and as a plan to fit these variations into a hierarchical structure according to the needs of society. Even during present times, caste status plays an important role in determining opportunities and outcomes for the people of the country.

Castes in India are divided between five main religions, namely, Hinduism, Islam, Christianity, Sikhism and Buddhism. Efforts to remove caste based inequality led some monarchs and British colonials to recognize most of the lower caste and religious populations as historically marginalized. Consequently, in independent India, the constitution went to great length in formulating affirmative action for marginalized populations, namely, Dalits, Adivasis, and to a lesser extent Shudras. In official lexicon Dalits are identified as Scheduled Castes (SCs), Adivasis as Schedule Tribes(STs) and Shudras as Other Backward Class (OBCs). Together, these groups are identified to be placed lower in the caste and religious hierarchy while the Hindu Upper Castes are understood to be placed higher.

According to Ghurye (1969) the term "Scheduled Castes", now standardised in the Constitution, was first used by the Simon Commission and embodied in Section 309 of the Government of India Act, 1935, and later included in the Government of India (Scheduled Caste) Order of 1936 (Chauhan, 2008). The term was used by the British government to designate all castes and classes previously included under the term "depressed classes". Because of the nature of their work and rank in the varna[19] hierarchy, they were labelled as "untouchables" (synonymous

[19] Varna means colour, which does not refer to skin colour or racial characteristics but to a system of colour symbolism reflecting the social hierarchy, see G. Flood (1998), *An Introduction to Hinduism* (Cambridge: CUP, 1998)

with the current dalits, meaning “downtrodden”), and thus were denied equality of opportunities in all socioeconomic fields, including education.

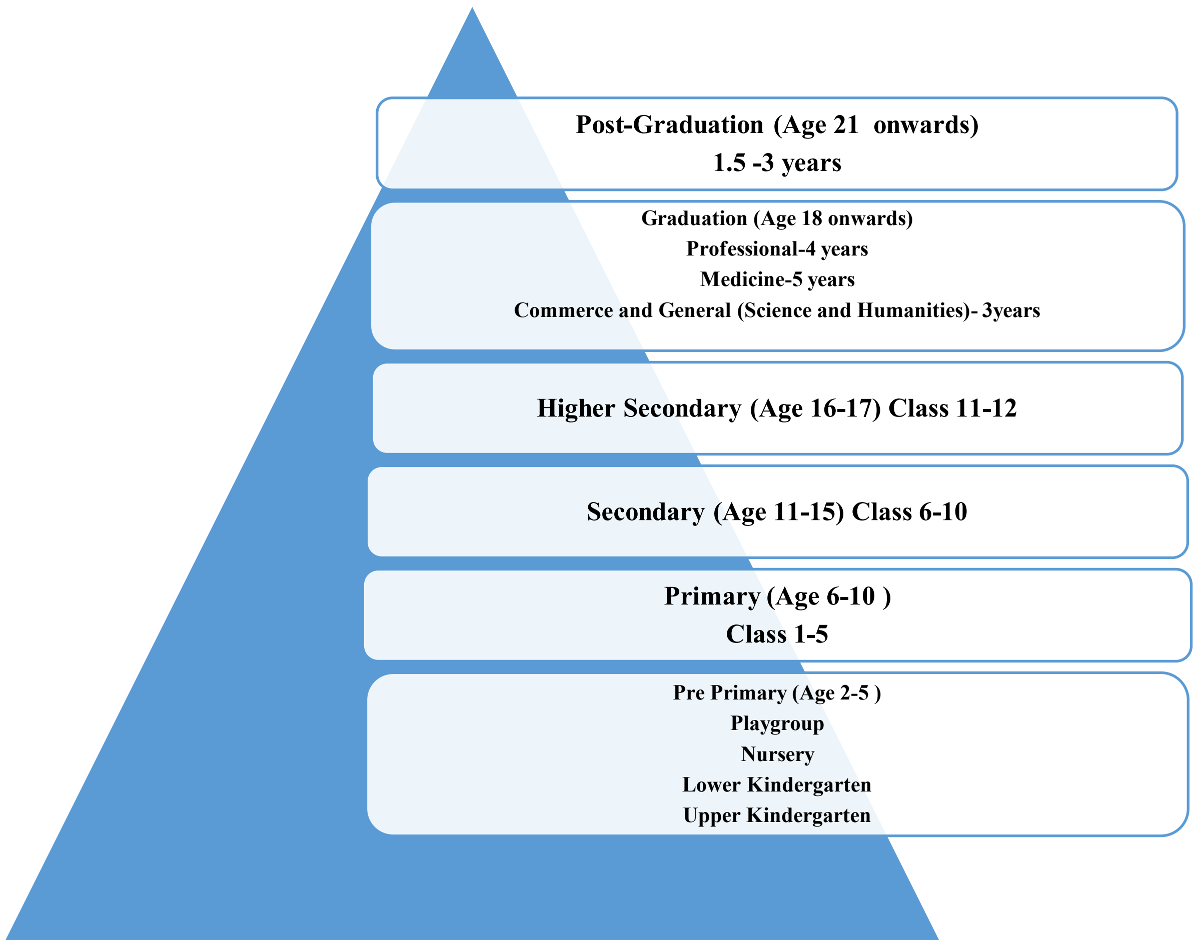
Literature on economic discrimination in India has extensively documented caste based discrimination. Scheduled Castes and Scheduled Tribes continue to be one of the most deprived social groups in the country and lag behind upper caste Hindus and other religions on indicators of social and economic development. Caste–based social division has recently transformed into a political agenda in the country (Chauhan, 2008). Three of the major socioeconomic categories are defined by caste combinations: Forward Castes (FCs) or Upper Castes (UCs), Scheduled Castes and Scheduled Tribes (SC/ST), and Other Backward Castes (OBCs). Though the Indian government has been making efforts since independence to bridge the socioeconomic gap between the advantaged and disadvantaged groups, yet, SCs and STs have remained socially, economically and culturally deprived because of their specific occupational and geographical conditions, hence remaining separated from mainstream Hindu society by virtue of their extreme poverty. The idea of untouchability is the main reason for backwardness of the SCs, although it has wide variations across the North and South of India.

**Appendix**

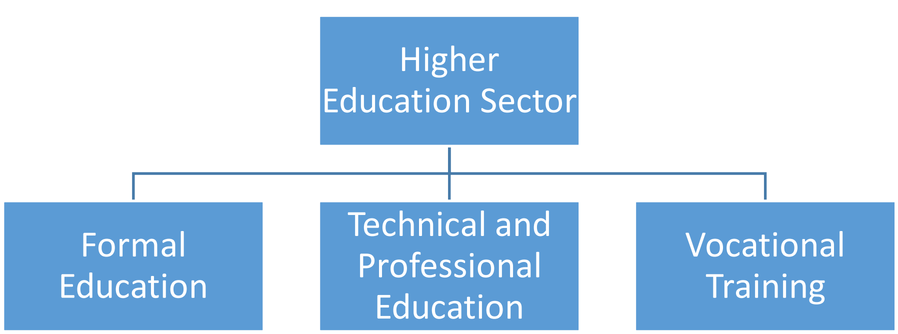
## The Indian Education System

The Indian education system can broadly be considered as a pyramidal structure:

11-12, 48-49, 58-61, as mentioned in Waughray, A. (2010). Caste Discrimination and Minority Rights: The Case of India's Dalits, *International Journal on Minority and Group Rights*, *17*(2), 327-353.




The University Grants Commission (UGC) is the statutory body responsible for determining and maintaining standards of higher education (NEP[20]2020). The composition of the Higher Education is shown below:




[20] https://www.education.gov.in/sites/upload files/mhrd/files/NEP Final English 0.pdf

The recent All India Survey on Higher Education (AISHE) (2019–2020) Report, released by the department of higher education mentions that the number of institutions of national importance has increased from 75 to 135 since 2015–2016. The gender parity index for all the social groups improved from 0.92 to 1.01 in the last five years (Mok and Marginson, 2021). Improvement in the gender parity index for all social groups is a good indication of sustainable education and infrastructures development. However, one worrisome aspect highlighted by the AISHE (2019–2020) report is, that, only 4% colleges have enrolment more than 3000 students. This is a matter of grave concern which should be taken into immediate consideration by the higher education institutions of India (NEP). This issue needs to be addressed strategically for long–term sustainability as although Indian higher education system is among the largest in the world, yet the share of unemployment corresponding to educational levels is significant.