

Impact of Owner's education on Productivity of Unincorporated Enterprises in India

Manjari Agrawal, Amit Kr. Eshore, and Anshika Chaudhary

Msc. Economics, Indira Gandhi Institute of Development and Research, Mumbai

Abstract

This paper examines the role of education of the establishment's owner in affecting labour productivity in India's vast unincorporated sector, a segment that employs millions yet consistently lags behind in productivity and efficiency. We find that owners' education is strongly and consistently associated with higher labour productivity: establishments run by more educated owners outperform those run by illiterate owners, with productivity premiums rising steadily from lower to higher education levels. Moreover, under-educated owners face a substantial productivity deficit, whereas over-education does not have a negative impact on productivity.

Introduction

A substantial part of India's workforce is engaged in self-employment activities in the informal sector, with 58% of the workforce being self-employed in 2023-24. This number has been on a steady rise due to sluggish growth and limited availability of wage employment in the formal sector. India's labour statistics define the self-employed as people who operate their own farm or non-farm enterprise, or are engaged independently in a profession or trade on their own, or with one or a few partners in household enterprises (PLFS 2023-24). Self-employment has been categorised into three types: own-account workers who run their own enterprise without employees, employers who run the enterprise by hiring paid employees, and unpaid helpers who work in a household enterprise for no wages. Around 18% of the self-employed workforce are unpaid family workers, and the remaining mostly captured by the non-agricultural unincorporated sector. This sector employs 12 crore workers and contributes about 44% to India's GVA. Despite this, the sector is characterized by significantly lower labour productivity compared to the formal sector, which impedes national economic progress. This makes it important to ask, what are the determinants of productivity and profitability of the establishments in this sector?

From the 7.34 crore estimated establishments belonging to unincorporated 'manufacturing', 'trade' and 'other services' sector, around 6.34 crore are estimated to be Own Account Establishments (OAE), i.e., an establishment which runs without any hired worker employed on a regular basis. Thus, it is safe to say that the owner is the manager and the primary human capital input, making owner-specific characteristics critical determinants of the establishment's productivity.

This study tries to analyze the role of education of the establishment's owner in affecting the labour productivity. We find that owners' education is strongly and consistently associated with higher labour productivity: establishments run by more educated owners outperform those run by illiterate owners, with productivity premiums rising steadily from lower education levels up to more than 100% for graduate-level owners. These effects remain robust even after extensive owner, firm, state, and industry controls. This link between owner's education and productivity in context of unincorporated enterprises in India can originate through several channels. Becker's Human Capital theory argues that education provides owners valuable cognitive, technical and analytical skills that enhance their productivity (Becker, 1975). Education also induces better managerial practices employed, leading to higher productivity (McKenzie and Woodruff, 2017). Moreover, education makes owners adaptable to existing and changing environments, for example, educated individuals have a comparative advantage in adopting new technologies, thus making them more productive (Welch, 1970; Bartel & Lichtenberg, 1987).

This study also brings in the concept of over-education and under-education of owner and its impact on labour productivity. Generally, a worker is said to be overeducated if they have acquired more education than is required to perform their jobs. Educational mismatch thus refers to the difference between the worker's attained level of education and the education required in the job. Now the vast majority of educational mismatch literature focuses on individual employees and their wages, starting with Duncan & Hoffman, 1981, who show that the return to an additional year of surplus education is about half the size of the return to an additional year of education, thus a "wage penalty". Shifting from employee personal returns to firm level productivity, the impact of overeducation is a bit ambiguous, with some studies showing positive impact on productivity, some negative and some insignificant. (S McGuinness, 2006; P Grunau, 2016; Kampelmann and Rycx, 2014) However, the effect of owners' educational mismatch on firm productivity remains largely unexplored in the literature. This may be because, in the formal sector, productivity is driven primarily by employees' skills and efforts, making the owner's education mismatch less relevant for study.

The concept of Education-Occupation Mismatch has also been well-documented in the Indian labor market (e.g., Datta & Mishra, 2019). These studies implicitly include the self-employed, yet they fail to capture the unique economic function of the self-employed owner within their own firm. In the unincorporated sector, the owner is simultaneously the worker, manager, and risk-taker. Therefore, the owner's human capital mismatch is not merely a wage determinant, but a critical driver of the firm's productivity. Our paper uniquely isolates the self-employed owner-manager in the Indian unincorporated sector to assess how educational mismatch affects the productivity of the establishment. Owners with surplus education exhibit a clear productivity advantage, showing 14% to nearly 30% higher labour productivity than adequately educated owners, while under-educated owners face a substantial productivity deficit of 14% to 27%. These results indicate that surplus education enhances owners' managerial and technological capabilities, whereas lacking the required education significantly impairs enterprise productivity, even after extensive controls.

The rest of the paper is structured in the following way: Section 1 provides the literature review and the conceptual framework. Section 2 studies the impact of education level of owner on labour productivity, controlling for other firm and owner characteristics. Section 3 discusses the methodology

for calculating overeducation and undereducation , and then studies their impact on productivity. Section 4 concludes the paper.

Literature Review

The economic motivation for this study stems from the persistent and significant productivity gap that characterizes the Indian economy. The unincorporated, or informal, sector, while providing employment to over 120 million people and contributing a massive share of the nation's Gross Value Added , operates at substantially lower efficiency levels than its formal counterparts. Formal sector firms are reported to be approximately 3.5 times more productive than informal enterprises (Rabbani & Raj SN, 2024). This divergence is partly attributed to external factors such as inadequate infrastructure and limited access to finance (Rabbani & Raj SN, 2024). However, the human capital within these firms is also strongly implicated. Studies examining Total Factor Productivity Growth in India often find that the positive effects of literacy and human capital are concentrated in the formal sector, suggesting that a lack of human capital impact contributes to the lower productivity growth in the informal sector (Kathuria et al., 2013). This vast, underperforming economic segment urgently requires an investigation into its internal, modifiable drivers of efficiency, with the owner's human capital being paramount.

The theoretical link between owner education and firm productivity is well-established through Human Capital Theory, arguing that investment in education provides the cognitive and managerial skills necessary for effective production (Becker, 1964). For the unincorporated sector, this link is particularly important because the owner functions as the sole manager, worker, and strategist. Numerous studies confirm a direct, positive association between owner/manager education and firm performance. Analysis across developing economies in Asia and Africa finds that informal firms managed by individuals with secondary or higher education exhibit significantly higher labor productivity (28% to 33%) compared to those with only primary schooling (Amin & Islam, 2022). The literature also suggests that education enhances firm productivity largely through improved management quality. Owners with higher human capital are more likely to employ better business practices (e.g., record-keeping, marketing, inventory control) which is a key driver of higher firm productivity in micro- and small enterprises (McKenzie & Woodruff, 2017). This relationship is vital in the informal sector, where education (or literacy) is argued to influence Total Factor Productivity by enhancing the workforce's ability to absorb and utilize new technology (Natarajan & Duraisamy, 2008).

Other than education, the literature also suggests several other determinants of labour productivity, like establishment characteristics, other owner characteristics and the business environment. The literature talks about making the owner's gender a crucial control in the informal sector of developing countries. Studies across various regions, including Ghana, Argentina, and India, have shown that female-owned informal enterprises tend to be significantly less productive and smaller than their male-owned counterparts (Avenyo & Kraemer-Mbula, 2021; Ghani et al., 2013; Amin, 2011). In India's informal sector, female-owned firms are, on average, 45 percent less productive than male-owned firms (Gang et al., 2021). Amin, 2011 shows that owner education level, firm size, sector, location, and hours of operation collectively explain only about 25 % to 30 % of the gender-based gap in labor productivity and firm-size. The unexplained component is often linked

to gendered social norms that restrict mobility and impose household burdens, leading female owners to operate primarily from home, which is estimated to explain 19% of the gender gap in productivity of unorganized manufacturing firms in India (Bose, 2023). Controlling for gender is thus essential to avoid attributing productivity differences due to social norms to educational level and mismatch effects.

Social group disparities, especially the caste system in India act as fundamental constraints on establishment productivity and thus must be controlled for. Research using Indian business data consistently finds that businesses owned by historically marginalized groups, such as Scheduled Castes and Tribes (SC and STs), perform significantly worse than non-SC & ST businesses in terms of earnings and owner human capital (Deshpande & Sharma, 2016). This performance disparity is exacerbated by evidence suggesting that SC/ST-owned firms register significantly lower growth compared to upper-caste firms (Deshpande & Sharma, 2013).

Firm age serves as an important proxy for a firm's accumulated experience and survival efficacy. Generally, older firms are found to be more productive than younger informal firms, aligning with survival selection effects. This positive relationship is partly attributed to the firm's tendency to formalize over time; as firm age increases, the probability of the firm adopting more formal features and operating at a higher scale of informality increases (Posti et al., 2025). Since formalization is associated with higher productivity, age acts as an indirect positive factor. Controlling for age accounts for the learning-by-doing and survival of the fittest mechanisms that can bias productivity of the establishments.

The adoption and usage of Information and Communication Technologies (ICT), especially digital payments, is emerging as a critical modern determinant of informal firm productivity. Empirical evidence from developing economies shows that informal enterprises that use digital technology (like mobile phones, computers, or the internet) exhibit significantly higher labor productivity, with productivity being 65% to 77% higher for users than for non-users (Danquah & Owusu, 2021). In the Indian informal sector, the adoption of digital transaction tools, such as digital payments, is correlated with higher sales and boosted sales by 12–15% even during slow economic periods like the COVID-19 lockdown (Dutta et al., 2023). Given that owner education often determines the ability to adopt and utilize such technologies, controlling for ICT usage is essential to prevent confounding the direct effect of education with its indirect effect through technological adoption.

Over-education is usually defined as a situation where a worker's attained education exceeds the level required for her job, and is part of the broader education mismatch literature that also includes under-education and adequate matches (Duncan & Hoffman, 1981; McGuinness, 2006). Surveying mainly OECD economies, McGuinness (2006) shows that overeducation is widespread and potentially costly both to workers and firms: surplus schooling is typically rewarded at a lower rate than required schooling and several firm-level studies associate over-education with lower productivity and wasted training investments. She also highlights that under-education tends to be more detrimental to productivity than over-education, though evidence on the latter is mixed.

The study of educational mismatch has evolved from analyzing individual wage penalties to examining its consequences for firm and regional productivity, highlighting allocative efficiency con-

cerns. Firm-level productivity studies make this asymmetry evident. Using linked employer–employee data for Germany, Grunau (2016) estimates the impact of overeducated and undereducated workers in establishments’ workforces. She finds that under-education significantly impairs establishment-level productivity, suggesting that employers have strong incentives to avoid recruiting under-qualified workers. Over-education, in contrast, has a small negative but statistically insignificant effect on productivity, suggesting that surplus schooling may be absorbed within firms without substantial losses. However, other European studies present a more nuanced view of over-education’s effect on the firm. Kampelmann and Rycx (2014), using linked panel data, demonstrated that the over-education of workers can have a positive, significant impact on firm productivity, suggesting that the benefits derived from workers’ latent skills (the spillover effect) can outweigh potential costs. Majumder and Mukherjee (2025) analyze skill mismatch in the Indian labour market using a panel-data framework, treating over- and under-education as components of a broader allocative inefficiency. They show that mismatch has a sizeable, statistically significant negative impact on labor productivity at the aggregate level, with estimates suggesting around a 25 percent productivity loss attributable to misallocation of skills. This solidifies the argument that educational mismatch is a macro-relevant constraint on productivity growth.

A parallel strand links over-education explicitly to labour-market segmentation between formal and informal employment. Herrera-Idárraga et al. (2012, 2015) study Colombia and suggests that the productivity and welfare consequences of over-education are likely to be more pronounced in segmented, dual labour markets typical of developing economies. For developing countries, however, direct evidence on over-education has long been constrained by data limitations, especially the absence of information on required schooling by occupation. Mehta et al. (2011) proposed an alternative approach that is particularly suited to such contexts. They proposed defining “unskilled” jobs by their low returns (below 7%) to secondary/tertiary education and applied this to India and three other developing economies to found rising over-education in unskilled service jobs, strongest in the Philippines but limited in India and Thailand. They also show that job-level mean and modal schooling are poor proxies for required education and can systematically misclassify workers as over- or under-educated. However we employ the realized matches statistical approach to capture mismatch by comparing attained education with the mean as the statistical threshold (Verdugo and Verdugo 1989).

Data and Methodology

Description of Data and Variables

We have used the Annual Survey Of Unincorporated Sector Enterprises (ASUSE) as our data source. We mainly focus on the latest 2023-24 survey, but similar analysis is done for both 2021-22 and 2022-23 datasets as well. This survey is conducted by The National Sample Survey Office (NSSO), MoSPI, Government of India on an annual full fledged basis since 2021. Earlier rounds of this survey exist, but they were not conducted annually and did not collect information on the owner’s educational attainment. This data collects the economic and operational characteristics of unincorporated non-agricultural establishments in manufacturing, trade and other services sectors of India. Broadly, this unincorporated sector also covers the informal sector enterprises. This survey is carried out following an ‘establishment approach’, where the respondents were selected

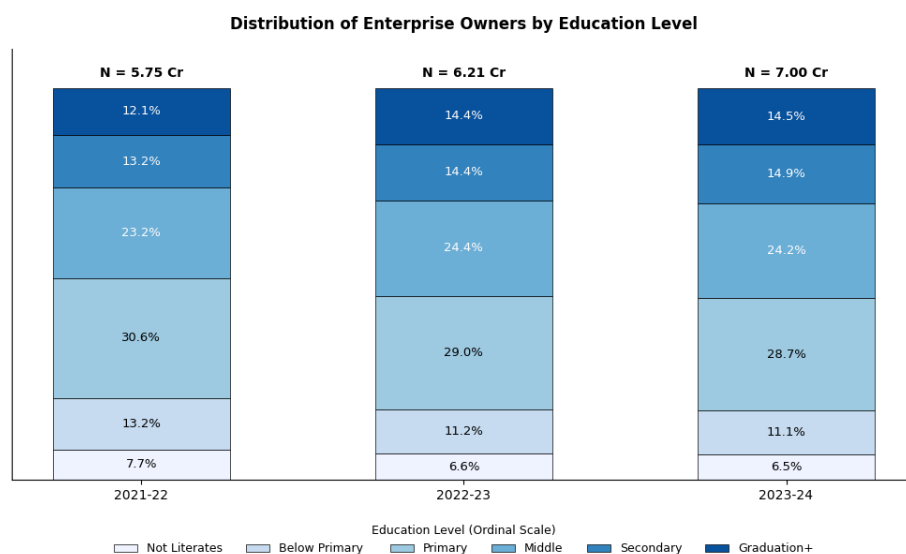


Figure 1: Distribution of Enterprises by the owner's education level

establishments. An establishment is an enterprise, or part of an enterprise that is situated in a single location and in which either only a single productive activity is carried out or in which the principal productive activity accounts for most of the value added.

Now, ASUSE collects data on the highest level of general education obtained by the proprietor or the major partner in a partnership. Thus, Self-Help Groups, Societies, Trusts, Clubs, Co-operatives, etc. have been dropped from our analysis and only proprietorship and partnership establishments have been considered.

Figure 1 shows the distribution of establishments according to the highest education level of the owner from 2021-22 to 2023-24. We see that majority of the establishments have owners who have completed primary education but have not completed higher secondary education. Within the time frame of three years, we find some insightful results. The percentage share of establishments owned by individuals with education levels below secondary has consistently dropped over the three years. On the other hand, the share of establishments owned by individuals with education levels equal and above higher secondary has increased. This pattern implies that that while establishments with owners having middle education levels remain the most common, the growth of new establishments is being driven by individuals with higher formal education.

We measure labour productivity of the establishments by calculating the Gross Value Added per Worker. Our main dependent variable is the log of GVA per worker. We use the annual GVA of the establishments for our analysis. For non-perennial establishments with a 30 days reference period, we calculate annual GVA by taking into account the number of months the establishment was operating. For calculating total workers, we include both full time and part time workers (formal, informal & unpaid family member) as well as the working owner. Figure 2 shows us the mean labour productivity by education level of owner from 2021 to 2023.

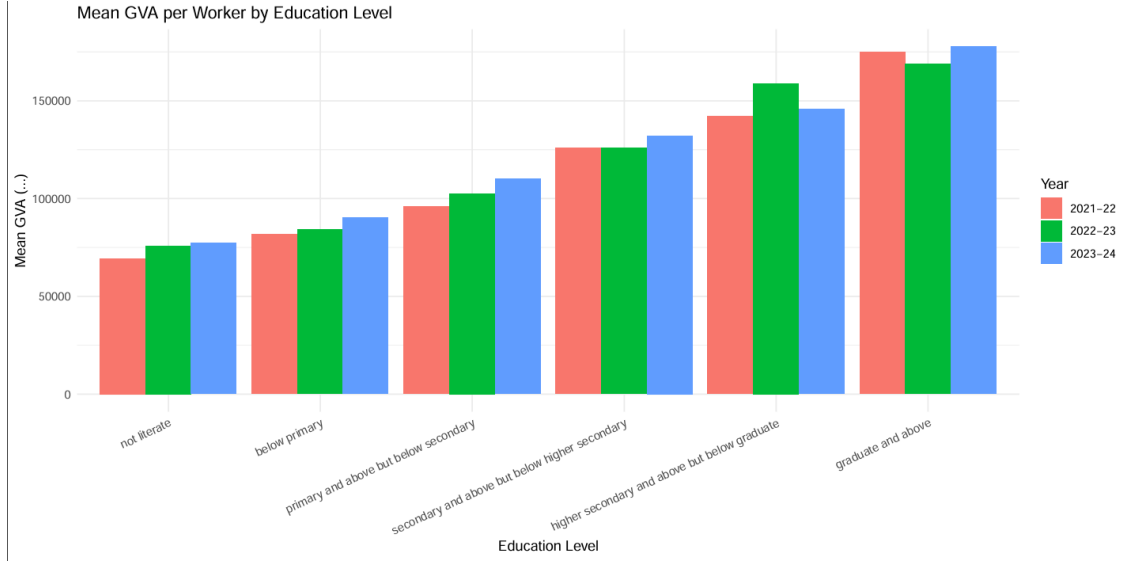


Figure 2: Mean GVA per worker by the education level of owner from 2021-23

The most significant observation from Figure 2 is the monotonic positive relationship between the owner's education level and the firm's labor productivity. As education levels rise through primary, secondary, and higher secondary, the Mean GVA per Worker increases in a stepped fashion, confirming the core tenet of Human Capital Theory that higher owner education leads to higher firm efficiency.

So, our main regression equation is :

$$\log(\text{GVA per worker})_i = \alpha + \beta_1 \text{Education Level}_i + \beta_2 \text{Owner Char.}_i + \beta_3 \text{Est. Char.}_i + \sum_{s=2}^{37} \gamma_s + \sum_{j=10}^{96} \delta_j + \epsilon_i$$

Here, γ_s represents state fixed effects and δ_j represents industry fixed effects using NIC 2008 code. We first run the regression without any owner and establishment characteristics. Then we add other controls one by one.

We control for several determinants of labour productivity to avoid the problem of omitted variable bias. This also helps in narrowing down the potential channels through which owner's education affects labour productivity. The motivation for various controls was previously discussed in the literature review. We control for owner's characteristics using their social group, gender (using ownership type) and whether they have technical education. Under establishment characteristics, we control for the age using the years of establishment; location ; nature of operation (seasonal, casual, perennial); whether they have a bank account; and whether they use computer and internet.

Baseline Regression Results

Our regression results are based on the OLS methodology with robust standard errors. All the specifications show control for state and industry effects. The results in Table 1 reveal a positive significant difference between the impact of education and no education on labour productivity, across all specifications. This means that owners with higher education run establishments with significantly higher labour productivity compared to those who are illiterate. It is also observed that the magnitude of the coefficient increases with level of education. We find that establishments run by graduate and above owners are around 119 % more productive than establishments run by illiterate owners, while establishments run by Secondary level owners are around 82 % more productive. The productivity premium is seen to be rising monotonically from below-primary to graduate-and-above education levels. These effects remain robust even after controlling for ownership type, social group, location, technology use, experience, as well as state and industry fixed effects. It should be noted that the impact of education levels declines as we keep controlling for other variables.

Controlling for owner characteristics, we observe that female proprietorships are 47% less productive than male proprietorships. Also non-backward social groups run establishments that are 35% more productive than ST owned establishments. This may reflect structural disadvantages faced by ST groups. Not having a bank account leads to 20% less productivity. Also taking into account establishment characteristics, we find that outside household establishments are significantly more productive than within household establishments. Moreover, establishments with more years of operation have a significant positive impact. This shows experience and learning-by-doing does improve productivity. Also not using internet and computer has significant negative impact on labour productivity.

We also interacted the general and technical education categorical variables and used it as a regressor in our analysis. Our findings indicate a strong, diminishing complementary relationship between general and technical education: acquiring technical training yields a substantial 12.6% boost to productivity for low-educated owners, but this marginal return significantly decreases as general education levels increase, suggesting partial substitutability between the two human capital investments.

The interaction of general education levels with computer as well as internet usage has also been used as a regressor. Digital proficiency, encompassing both computer and internet use, is a dominant factor driving firm productivity and exhibits strong complementarity with general education. While both digital tools provide substantial gains across all levels, the largest proportional increases (over 32%) are observed among the lowest-educated owners, suggesting that investing in digital access and literacy is the most effective way to empower small businesses and compensate for deficits in formal schooling.

We did a similar analysis for establishments surveyed in ASUSE 2021–22 as well as ASUSE 2022–23. and the results obtained were consistent with the main findings.

Table 1: OLS Regression Results: Determinants of log(GVA per Worker)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Education (Base: Not Lit.)											
Below Primary	0.197*** (0.007)	0.197*** (0.007)	0.132*** (0.006)	0.126*** (0.006)	0.114*** (0.006)	0.118*** (0.006)	0.118*** (0.006)	0.116*** (0.006)	0.112*** (0.006)	0.107*** (0.006)	0.102*** (0.006)
Primary	0.351*** (0.006)	0.351*** (0.006)	0.246*** (0.006)	0.229*** (0.005)	0.202*** (0.005)	0.199*** (0.005)	0.194*** (0.005)	0.195*** (0.005)	0.184*** (0.005)	0.179*** (0.005)	0.173*** (0.005)
Middle	0.498*** (0.006)	0.496*** (0.006)	0.378*** (0.006)	0.351*** (0.006)	0.306*** (0.006)	0.295*** (0.006)	0.282*** (0.006)	0.283*** (0.006)	0.264*** (0.005)	0.260*** (0.005)	0.253*** (0.005)
Secondary	0.599*** (0.006)	0.593*** (0.006)	0.481*** (0.006)	0.447*** (0.006)	0.384*** (0.006)	0.366*** (0.006)	0.343*** (0.006)	0.349*** (0.006)	0.320*** (0.006)	0.316*** (0.006)	0.309*** (0.006)
Graduate+	0.788*** (0.007)	0.775*** (0.007)	0.656*** (0.006)	0.610*** (0.006)	0.512*** (0.006)	0.485*** (0.006)	0.446*** (0.006)	0.456*** (0.006)	0.411*** (0.006)	0.410*** (0.006)	0.402*** (0.006)
Tech Edu (Yes)		0.056*** (0.004)	0.075*** (0.004)	0.076*** (0.004)	0.057*** (0.004)	0.052*** (0.004)	0.048*** (0.004)	0.053*** (0.004)	0.029*** (0.004)	0.023*** (0.004)	0.025*** (0.004)
Ownership (Prop. Male)											
Proprietary (Fem)			-0.648*** (0.004)	-0.639*** (0.004)	-0.601*** (0.003)	-0.472*** (0.003)	-0.453*** (0.003)	-0.437*** (0.003)	-0.429*** (0.003)	-0.456*** (0.003)	-0.456*** (0.003)
Proprietary (Trans)			0.062 (0.171)	0.049 (0.170)	0.131 (0.168)	0.221 (0.165)	0.242 (0.164)	0.241 (0.163)	0.247* (0.151)	0.267* (0.140)	0.248* (0.140)
Partnership (Same HH)			0.167*** (0.019)	0.160*** (0.019)	0.034** (0.018)	0.011 (0.018)	-0.011 (0.017)	-0.006 (0.017)	-0.041** (0.017)	0.129*** (0.018)	0.127*** (0.018)
Partnership (Diff HH)			-0.144*** (0.018)	-0.147*** (0.018)	-0.269*** (0.018)	-0.304*** (0.017)	-0.319*** (0.017)	-0.298*** (0.017)	-0.312*** (0.017)	-0.117*** (0.017)	-0.118*** (0.017)
Social Group (Base: ST)											
SC				0.106*** (0.007)	0.110*** (0.007)	0.090*** (0.007)	0.090*** (0.007)	0.089*** (0.007)	0.082*** (0.007)	0.074*** (0.006)	0.061*** (0.006)
OBC				0.204*** (0.006)	0.200*** (0.006)	0.170*** (0.006)	0.164*** (0.006)	0.160*** (0.006)	0.152*** (0.006)	0.145*** (0.006)	0.131*** (0.006)
Others				0.305*** (0.007)	0.282*** (0.006)	0.248*** (0.006)	0.234*** (0.006)	0.228*** (0.006)	0.217*** (0.006)	0.212*** (0.006)	0.199*** (0.006)
Not Known				0.308*** (0.017)	0.307*** (0.017)	0.271*** (0.017)	0.262*** (0.017)	0.264*** (0.017)	0.272*** (0.017)	0.253*** (0.017)	0.238*** (0.016)
Bank Account (Owner)											
Enterprise					0.342*** (0.004)	0.277*** (0.004)	0.207*** (0.004)	0.201*** (0.004)	0.146*** (0.004)	0.229*** (0.004)	0.225*** (0.004)
No Account					-0.224*** (0.003)	-0.195*** (0.003)	-0.183*** (0.003)	-0.183*** (0.003)	-0.152*** (0.003)	-0.159*** (0.003)	-0.157*** (0.003)
Location (within HH)											
Perm. Structure						0.411*** (0.003)	0.367*** (0.003)	0.374*** (0.003)	0.356*** (0.003)	0.358*** (0.003)	0.353*** (0.003)
Temp. Structure						0.282*** (0.007)	0.276*** (0.007)	0.282*** (0.007)	0.278*** (0.007)	0.265*** (0.007)	0.270*** (0.007)
No Structure						0.351*** (0.008)	0.346*** (0.008)	0.341*** (0.008)	0.343*** (0.007)	0.328*** (0.007)	0.345*** (0.007)
Mobile Market						0.299*** (0.008)	0.301*** (0.008)	0.294*** (0.008)	0.298*** (0.008)	0.272*** (0.008)	0.285*** (0.007)
Street Vendors						0.345*** (0.005)	0.349*** (0.005)	0.343*** (0.005)	0.345*** (0.005)	0.312*** (0.005)	0.326*** (0.005)
Est Type (Hired)							-0.205*** (0.003)	-0.201*** (0.003)	-0.167*** (0.003)	-0.439*** (0.004)	-0.435*** (0.004)
Years of Operation (<1 Year)											
1-3 Years								0.136*** (0.006)	0.137*** (0.006)	0.136*** (0.006)	0.131*** (0.006)
>3 Years								0.239*** (0.005)	0.243*** (0.005)	0.250*** (0.005)	0.244*** (0.005)
Computer Used (Base: Yes)											
									-0.111*** (0.005)	-0.207*** (0.005)	-0.200*** (0.005)
Internet Used (Base: Yes)											
									-0.186*** (0.003)	-0.195*** (0.003)	-0.192*** (0.003)
Log of Total Workers											
										-0.300*** (0.003)	-0.289*** (0.003)
Nature of Operation (Perennial)											
Seasonal											-1.003*** (0.016)
Casual											-2.691*** (0.112)
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	466847	466847	466847	466847	466847	466847	466847	466847	466847	466847	466847
Adj. R^2	0.299	0.299	0.361	0.367	0.386	0.413	0.420	0.424	0.432	0.448	0.460

Note: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Education Mismatch:

Now we bring in over-education and under-education in our regression equation. To do this we need the attained level of education for a particular owner and the required level of education for a particular industry. While the former can be obtained from establishment-level data from ASUSE, to estimate the latter, the literature provides various methods. We use realized matches (RM) approach. This is a statistical approach that captures mismatch from the market's perspective and thus involves interaction from both demand and supply side factors. The Realized Matches approach compares the owner's attained education with the mean education level of owners in the particular industry. The issue with this approach is that it is not capturing the actual level of education required in the industry (Verhaest & Omeij, 2006). However it also has an advantage as it allows us to see the intrinsic requirements in the industry at present. This approach could also be done by comparing with the mode but was avoided due to multi-modal problems that could be faced.

Now to obtain the mean of general education attained by owner in each industry according to the NIC 2008 codes under survey coverage of ASUSE, we convert the levels of education in the following way using the conversion adapted from Duraisamy (2002). Not literate (1) corresponds to zero years of general education obtained; Below primary (2) corresponds to 3 years of education; Primary and above but below secondary (3) corresponds to 5 years of education; Secondary and above (4) corresponds to 10 years of education; Higher secondary and above (5) corresponds to 12 years of education; Graduate and above (6) corresponds to 15 years of education. To get information on the required education by the industry, we use the 2 digit National Industrial Classification (NIC) codes for 2008, a numerical system used in India to categorize all economic and business activities. ASUSE covers industries under manufacturing, trading and other services. We calculate the average years of general education for every NIC 2008 code using the information on education, industry and sampling weights. Then, following Verdugo and Verdugo (1989), we create a threshold of minus and plus one standard deviation from mean. This range establishes the adequate education required for each industry classification. So in this scenario, the owner would be overeducated if their education is more than mean plus standard deviation and would be under-educated if their education is less than mean minus standard deviation. More precisely, suppose e_i represents attained years of education of an individual 'i', and e_o and s_o are the mean and standard deviation of years of education respectively for their industry, o. Then as per RM, a worker will be considered Overeducated if: $e_i > e_o + s_o$; Undereducated if: $e_i < e_o - s_o$; and Adequately educated if: $e_i - s_o \leq e_i \leq e_o + s_o$. Using this, we find the share of over-educated, under-educated and adequately educated owners across industries and also across Broad Activity Categories from 2021-22 to 2023-24

From Figure 3(a) and 3(b), we observe that the share of over-education has increased across years for all BACs. However, the share of under-education has increased in the services sector.

Now with the available information, we create a dummy variable for over-education and under-education at the establishment-level. If the owner is over-educated, the over-education dummy takes value 1, 0 otherwise and if the owner is under-educated, the under-education dummy takes value 1, 0 otherwise. The new regression equation now is as follows:

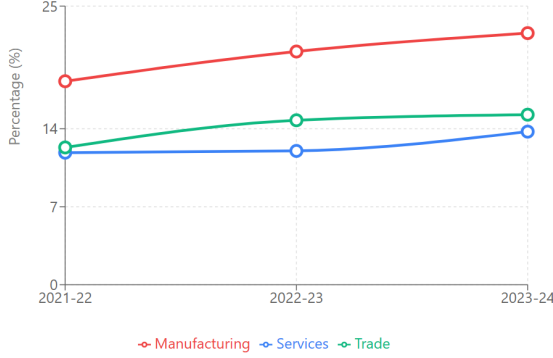


Figure 3(a): Over-education Trend across BAC from 2021-23

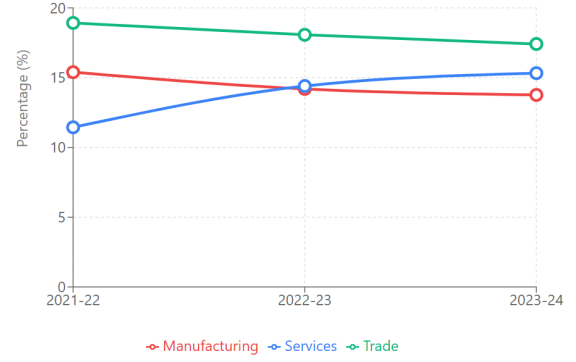


Figure 3(b): Under-education Trend across BAC from 2021-23

$$\log(\text{GVA per worker})_i = \alpha + \beta_1 \text{OverEducated}_i + \beta_2 \text{UnderEducated}_i + \beta_3 \text{Owner Char.}_i + \beta_4 \text{Est. Char.}_i + \sum_{s=2}^{37} \gamma_s + \sum_{j=10}^{96} \delta_j + \epsilon_i$$

Regression Results:

Over-education signals potential inefficiencies if the returns to surplus education are found to be lower than those of the required education. It generally implies that the resources are being wasted because the marginal levels of education received are costing more than the productivity advantage. However, we find the following results:

Across all model specifications, the dummy variable for over-education, which is defined relative to the base category of adequately educated owners, remains positive, stable in magnitude, and highly statistically significant. In the baseline model, the coefficient of 0.253 implies that establishments owned by over-educated individuals have about 28.8% higher labour productivity than those whose owners possess education exactly adequate for their industry. Even in the fully controlled specification, where the coefficient falls to 0.133, the productivity premium remains sizeable at around 14.2%. This consistent positive association suggests that owners with more education than required for the activity bring additional managerial skills, better decision-making capacity, and greater ability to adopt productive technologies and practices, all of which contribute to substantially higher enterprise productivity.

On the other hand, the dummy variable capturing under-education, also defined relative to adequately educated owners as the base category, remains negative and highly statistically significant across all model specifications. In the simplest model, establishments owned by under-educated individuals have about 27% lower labour productivity compared to those whose owners possess the level of education appropriate for their industry (based on the coefficient of -0.317 in the log-linear specification). Even after introducing extensive controls, the productivity penalty remains large

at about 14%. This robustness indicates that under-education relative to industry requirements has an independent and economically meaningful negative impact on enterprise productivity, beyond what can be explained by other observable characteristics. The effect of all the control variables are similar to the previous analysis.

Table 2: Regression Results: Determinants of GVA per worker (Over and Under Education)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Education (Base: Matched)											
Under-educated	-0.317*** (0.004)	-0.313*** (0.004)	-0.245*** (0.003)	-0.225*** (0.003)	-0.195*** (0.003)	-0.184*** (0.003)	-0.176*** (0.003)	-0.175*** (0.003)	-0.159*** (0.003)	-0.156*** (0.003)	-0.152*** (0.003)
Over-educated	0.253*** (0.003)	0.242*** (0.003)	0.244*** (0.003)	0.224*** (0.003)	0.177*** (0.003)	0.162*** (0.003)	0.138*** (0.003)	0.147*** (0.003)	0.128*** (0.003)	0.134*** (0.003)	0.133*** (0.003)
Tech Edu (Yes)		0.086*** (0.004)	0.104*** (0.004)	0.103*** (0.004)	0.077*** (0.004)	0.070*** (0.004)	0.064*** (0.004)	0.068*** (0.004)	0.041*** (0.004)	0.035*** (0.004)	0.037*** (0.004)
Ownership Type (Base: Prop. M)											
Proprietary (Fem)			-0.660*** (0.004)	-0.650*** (0.004)	-0.607*** (0.003)	-0.478*** (0.003)	-0.457*** (0.003)	-0.442*** (0.003)	-0.433*** (0.003)	-0.460*** (0.003)	-0.460*** (0.003)
Proprietary (Trans)			0.097 (0.162)	0.081 (0.162)	0.162 (0.164)	0.253 (0.158)	0.273 (0.154)	0.272 (0.153)	0.274 (0.146)	0.293** (0.134)	0.273** (0.135)
Partnership (Same HH)			0.187*** (0.019)	0.177*** (0.019)	0.041** (0.018)	0.017 (0.018)	-0.007 (0.018)	-0.003 (0.017)	-0.039*** (0.017)	0.131*** (0.018)	0.129*** (0.018)
Partnership (Diff HH)			-0.124*** (0.018)	-0.129*** (0.018)	-0.261*** (0.018)	-0.298*** (0.017)	-0.314*** (0.017)	-0.294*** (0.017)	-0.309*** (0.017)	-0.114*** (0.017)	-0.115*** (0.017)
Social Group (Base: ST)											
SC				0.105*** (0.007)	0.109*** (0.007)	0.090*** (0.007)	0.090*** (0.007)	0.089*** (0.007)	0.082*** (0.007)	0.073*** (0.007)	0.061*** (0.006)
OBC				0.210*** (0.007)	0.205*** (0.006)	0.174*** (0.006)	0.167*** (0.006)	0.163*** (0.006)	0.155*** (0.006)	0.148*** (0.006)	0.134*** (0.006)
Others				0.323*** (0.007)	0.297*** (0.006)	0.260*** (0.006)	0.245*** (0.006)	0.238*** (0.006)	0.227*** (0.006)	0.221*** (0.006)	0.207*** (0.006)
Not Known				0.308*** (0.017)	0.307*** (0.017)	0.270*** (0.017)	0.261*** (0.017)	0.263*** (0.017)	0.271*** (0.017)	0.252*** (0.017)	0.237*** (0.016)
Bank Account Holder (Owner)											
Enterprise					0.362*** (0.004)	0.292*** (0.004)	0.216*** (0.004)	0.212*** (0.004)	0.153*** (0.004)	0.236*** (0.004)	0.231*** (0.004)
No Account					-0.234*** (0.003)	-0.203*** (0.003)	-0.190*** (0.003)	-0.190*** (0.003)	-0.157*** (0.003)	-0.164*** (0.003)	-0.162*** (0.003)
Location (Within HH)											
Permanent Structure						0.418*** (0.003)	0.370*** (0.003)	0.377*** (0.003)	0.358*** (0.003)	0.360*** (0.003)	0.355*** (0.003)
Temporary Structure						0.270*** (0.007)	0.264*** (0.007)	0.270*** (0.007)	0.267*** (0.007)	0.254*** (0.007)	0.259*** (0.006)
No Structure						0.334*** (0.008)	0.330*** (0.008)	0.325*** (0.008)	0.329*** (0.007)	0.314*** (0.007)	0.331*** (0.007)
Mobile Market						0.284*** (0.008)	0.287*** (0.008)	0.281*** (0.008)	0.286*** (0.008)	0.259*** (0.008)	0.273*** (0.007)
Street Vendors						0.329*** (0.005)	0.334*** (0.005)	0.328*** (0.005)	0.332*** (0.005)	0.299*** (0.005)	0.313*** (0.004)
Est Type (Hired)											
							-0.216*** (0.003)	-0.213*** (0.003)	-0.176*** (0.003)	-0.448*** (0.003)	-0.445*** (0.003)
Years of Operation (<1 Year)											
1-3 Years								0.136*** (0.006)	0.138*** (0.006)	0.137*** (0.006)	0.131*** (0.006)
>3 Years								0.236*** (0.005)	0.241*** (0.005)	0.248*** (0.005)	0.241*** (0.005)
Computer Used (Base: Yes)											
									-0.112*** (0.005)	-0.209*** (0.005)	-0.201*** (0.005)
Internet Used (Base: Yes)											
									-0.194*** (0.002)	-0.203*** (0.002)	-0.200*** (0.002)
Log of Total Worker											
										-0.301*** (0.003)	-0.290*** (0.003)
Nature of Operation (Perennial)											
Seasonal											-1.006*** (0.016)
Casual											-2.689*** (0.113)
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	466847	466847	466847	466847	466847	466847	466847	466847	466847	466847	466847
Adj. R^2	0.288	0.288	0.353	0.360	0.381	0.409	0.416	0.421	0.430	0.446	0.458

Note: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Conclusion

This paper set out to examine whether owner education and educational mismatch shape labour productivity in India's vast unincorporated sector, a segment that employs millions yet consistently lags behind in productivity and efficiency. Using three consecutive rounds of ASUSE (2021–22 to 2023–24), our analysis provides robust evidence that owner human capital is a significant determinant of informal enterprise performance. Labour productivity rises monotonically with the owner's educational attainment, and even after conditioning on various owner and establishment characteristics, higher levels of schooling continue to deliver sizeable productivity premium. The importance of complementary capabilities, such as technical/vocational training, digital adoption, access to banking, and years of operational experience, further underscores that education interacts with other forms of human and technological capital to shape firm productivity.

A key contribution of this study is the incorporation of educational mismatch using the Realized Matches approach. Contrary to much of the firm-level evidence from other countries, we find that over-education is not a drag on productivity. Instead, owners whose attained schooling exceeds the statistical norm for their industry exhibit a large and persistent productivity premium, even after extensive controls. This suggests that in environments characterised by managerial constraints, informational frictions, and limited adoption of efficient practices, which are typical features of India's informal sector, surplus schooling enhances decision-making capacity rather than generating inefficiencies. In stark contrast, under-education carries a substantial productivity penalty, reinforcing the idea that insufficient schooling constrains the ability of owners to manage, compete, and adapt to evolving market conditions.

From a policy perspective, the results highlight three priorities: (i) expanding access to quality schooling and vocational training for current and future entrepreneurs; (ii) promoting digital literacy and affordable access to technology; and (iii) designing targeted support to under-educated owners who face structural barriers to productivity growth. Strengthening human capital can generate broad-based improvements in managerial capability, adoption of technology, and ultimately, enterprise productivity.

In sum, this paper provides new evidence that owner education, both its level and its alignment with industry norms—plays a decisive role in shaping productivity in India's unincorporated sector. By bringing mismatch considerations into the analysis, it also broadens the understanding of how human capital operates within establishments in a developing country. With availability of more data in future, further analysis can be done using the trend of over-education and under-education of owners.

References

- [1] Amin, M. (2011). *Labor productivity, firm-size and gender: the case of informal firms in Argentina and Peru*. [World Bank]
- [2] Amin, M., & Islam, A. (2022). *Does manager education play a role in the productivity of informal firms in developing economies? Evidence from firm-level surveys*. [Review of Development]

Economics, 26, 962–984.].

- [3] Avenyo, E. K., & Kraemer-Mbula, E. (2021). [*Innovation and the performance of informal enterprises in developing countries: A gender perspective*]. [International Journal of Gender and Entrepreneurship, Vol. 13, No. 4, pp. 277-301].
- [4] Bartel, A., & Lichtenberg, F. (1987). [*Technical Change, Learning, and Wages*]. [NBER Working Paper No. w2732].
- [5] Becker, G. S. (1964). [*Human capital: A theoretical and empirical analysis*]. [Chicago: University of Chicago Press].
- [6] Becker, G. S. (1975). [*Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education (2nd ed.)*]. [New York: National Bureau of Economic Research.].
- [7] Bose, S. (2023). [*The penalty of work from home: gender gap in productivity of unorganised manufacturing firms in India*]. [Small business economics, 2023 - Springer].
- [8] Danquah, M., & Owusu, S. (2021). [*1) Digital technology and productivity of informal enterprises: Empirical evidence from Nigeria*]. [econstor.eu].
- [9] Datta, A., & Mishra, A. (2019). [*Effect of Education–Occupation Mismatch on Wages in India*]. [The Indian Journal of Labour Economics, Springer;The Indian Society of Labour Economics (ISLE), vol. 62(3), pages 389-411, September.].
- [10] Deshpande, A., & Sharma, S. (2013). [*Entrepreneurship or Survival? Caste and Gender of Small Business in India*]. [No 228, Working papers from Centre for Development Economics, Delhi School of Economics].
- [11] Deshpande, A., & Sharma, S. (2016). [*Disadvantage and discrimination in self-employment: caste gaps in earnings in Indian small businesses*]. [Small Business Economics, Springer, vol. 46(2), pages 325-346, February.].
- [12] Dutta, S., et al. (2023). [*Informal sector in India and adoption of digital technologies*]. [Indian Growth and Development Review, 2023 - emerald.com].
- [13] Duncan, G., & Hoffman, S. (1981). [*The Incidence and Wage Effects of Overeducation*]. [Economics of Education Review, 1, 75-86.].
- [14] Gang, I. N., et al. (2021). [*The gender productivity gap: Evidence from the Indian informal sector*]. [WIDER Working Paper Series wp-2021-183, World Institute for Development Economic Research (UNU-WIDER).].
- [15] Ghani, E., Kerr, W., & O’Connell, S. (2013). [*Spatial Determinants of Entrepreneurship in India..* [Regional Studies, 48(6), 1071–1089.].
- [16] Grunau, P. (2016). [*The impact of overeducated and undereducated workers on*

- establishment-level productivity: First evidence for Germany*]. [International Journal of Manpower, 2016 - emerald.com].
- [17] Herrera-Idárraga, P., et al. (2012). [*Informality and Overeducation in the Labor Market of a Developing Country*]. [XREAP2012-20, Available at SSRN: <https://ssrn.com/abstract=2172058> or <http://dx.doi.org/10.2139/ssrn.2172058>].
- [18] Herrera-Idárraga, P., et al. (2016). [*Regional Wage Gaps, Education and Informality in an Emerging Country: The Case of Colombia*]. [Spatial Economic Analysis, Taylor Francis Journals, vol. 11(4), pages 432-456, October.]].
- [19] Kampelmann, S., & Rycx, F. (2012). [*The impact of educational mismatch on firm productivity: Evidence from linked panel data*][Economics of Education Review, 2012, vol. 31, issue 6, 918-931]
- [20] Kathuria, V., Raj, R. S. N., & Sen, K. (2013). [*Impact of Human Capital on Manufacturing Productivity Growth in India*]. [Springer Books, Human Capital and Development, edition 2, chapter 0, pages 23-37, Springer.].
- [21] Majumder, R., & Mukherjee, D. (2025). [*Skill Mismatch in Indian Labour Market and its Impact on Productivity: A Panel Data Analysis*]. [Available at SSRN: <https://ssrn.com/abstract=5423594> or <http://dx.doi.org/10.2139/ssrn.5423594>].
- [22] McGuinness, S. (2006). [*Overeducation in the labour market*]. [Journal of economic surveys, 2006 - Wiley Online Library].
- [23] McKenzie, D., & Woodruff, C. (2017). [*Business Practices in Small Firms in Developing Countries*]. [Management Science 63(9):2967-2981.].
- [24] Mehta, A., et al. (2011). [*Overeducation in developing economies: How can we test for it, and what does it mean?*]. [Economics of Education Review, Volume 30, Issue 6, December 2011, Pages 1334-1347].
- [25] Natarajan, R., & Duraisamy, P. (2008). [*Efficiency and productivity in the Indian unorganized manufacturing sector: did reforms matter?*]. [International Review of Economics, Springer; Happiness Economics and Interpersonal Relations (HEIRS), vol. 55(4), pages 373-399, December].
- [26] Posti, L., et al. (2025). [*An empirical study of firm characteristics, degree of informality, and productivity heterogeneity in the Indian informal sector*]. [Applied Economics, 2025 - Taylor Francis].
- [27] Rabbani, M., & Raj, S. N. (2024). [*The role of financial constraints in firm transition—Evidence from Indian manufacturing*]. [Journal of International Development, John Wiley Sons, Ltd., vol. 36(4), pages 2193-2228, May.].
- [28] Verdugo, R., & Verdugo, N. (1989). [*The Impact of Surplus Schooling on Earnings: Some*

Additional Findings]. [The Journal of Human Resources, Vol. 24, No. 4 (Autumn, 1989), pp. 629-643].

[29] Welch, F. (1970). [*Education in Production*]. [Journal of Political Economy, 1970, vol. 78, issue 1, 35-59].