Trends in College Wage Premiums in India, 1987 - 2012

Aakansha Jain

July 2025

Abstract

This paper traces how the college-versus-high-school earnings gap evolved in India during the decades when tertiary enrollment was rising rapidly. Drawing on five waves of the *National Sample Survey*—43 (1987-88), 50 (1993-94), 55 (1999-2000), 61 (2004-05), and 68 (2011-12)—we estimate conventional Mincer wage equations and complement them with the two-part Cragg models that separate the extensive margin (labor-force participation) from the intensive margin (weekly pay, conditional on work). The repeated cross-sections are organized into synthetic birth cohorts to track wage premiums over time. Although university graduates continue to command sizeable labor-market advantages, the premium for the youngest cohorts appears to have flattened during the period when seat capacity expanded.

JEL codes: I26, I23, J31, J24, O15

1 Introduction

Following the 1991 market reforms, India recorded its fastest-ever expansion of higher education: the number of universities rose from fewer than 200 in 2000 to roughly 800 by 2015, while student enrollment jumped from 8.6 million to 33.3 million (Joshi and Ahir, 2016). Yet even after this remarkable growth, the country's Gross Enrollment Ratio (GER) stood at only about 24 percent in 2015-well below China's 40 percent in 2016 (World Bank, 2018) and the United States' 88 percent in the late 2010s (World Bank, 2024). Against this backdrop of rapid but still-incomplete massification, our study asks: *how has the wage premium for university graduates relative to higher-secondary graduates evolved across cohorts and age groups?* Tracking that premium during India's enrollment boom offers a window onto whether expanding access has narrowed—or merely redistributed—the returns to a degree in the labor market.

Several other developing economies—including China, Taiwan, and Thailand—experienced comparable surges in tertiary enrollment. In each case, the proportion of 18- to 23-year-olds in higher education more than doubled within a generation after admission quotas were relaxed and private institutions proliferated (Li et al., 2025). While policymakers often celebrate this broader access, economists caution that an over-supply of degree holders can

exceed labor-market demand—particularly when technological change and trade exposure accelerate the pace at which skill requirements evolve (Goldin and Katz, 2008; Autor, 2014).

From an individual perspective, tertiary education continues to offer both private and social returns. A first university degree is uniquely valued for the occupation-specific skills that shape a graduate's labor-market standing. Students therefore enroll with two central expectations: securing suitable employment and earning a higher wage premium relative to higher-secondary completers. Reflecting these objectives, our analysis concentrates on (i) the probability of being employed after graduation and (ii) the wage differential conditional on employment.

Recent Indian studies document a rapid rise in private colleges—many charging high fees yet offering limited assurances of instructional quality (Emont, 2024; Agarwal, 2006). Economic theory predicts that when the supply of degree-holders grows faster than high-skill employment opportunities, the graduate wage premium will narrow (Hanushek and Woessmann, 2008). Consistent with this mechanism, analyses of Periodic Labor Force Survey (PLFS) data reveal that the mean college-versus-high-school wage premium remained broadly unchanged between 2011-12 and 2018-19; quantile regressions further show a modest decline in the lower half of the wage distribution (Chen et al., 2022). This marks a distinct shift from the steady upward trend documented for 1983-2004 by (Kijima, 2006) and (Azam, 2010), suggesting that India's labor market began adjusting to the surge in graduates during the 2010s.

Existing studies highlight three ways the college wage premium can rise or fall. First, supply if the number of graduates grows faster than skilled jobs, relative wages shrink (Card and Lemieux, 2001; Goldin and Katz, 2008). Second, demand new technologies and the spread of knowledge-based services make firms pay more for tertiary skills (Autor, 2014; Azam, 2010). Third, composition returns vary by where and what one studies—well-resourced, selective colleges, and degrees in engineering or management, tend to pay more than general arts or science(Kingdon and Unni, 2001; Chakraborty, 2019; Bhandari and Thakur, 2013). Cross-country evidence also shows that China's heavy focus on technical fields has kept its premiums higher than India's broader mix(Bharti and Yang, 2021).

Higher-education growth has not been uniform across India. In some states, private technical colleges have multiplied, while many public universities remain resource-constrained. As a result, wage returns may vary by institutional quality and field of study. Labor-market segmentation—by formality, location, caste, or gender—can further condition these returns, although our empirical focus is on cohort differences in the college premium.

Earlier analyses of National Sample Survey (NSS) data paint a more nuanced picture. (Duraisamy, 2002) and (Kijima, 2006) show that the college-versus-high-school wage premium rose for all experience groups between the 1980s and early 2000s, but the gains were larger for mid-career and senior workers than for recent entrants. In contrast, (Azam, 2010) finds that the increase over 1983-2004 was driven mainly by workers aged 25-34, with little change among older cohorts. Together, these studies suggest that different experience groups are not perfect substitutes in India's labor market. However, those studies stop at or before NSS round 61 (2004-05) and thus do not capture more recent expansions in higher education. To observe how the college-versus-high-school wage differential evolves over time, the present analysis pools five repeated cross-sections of the National Sample Survey (1987-88, 1993-94, 1999-00, 2004-05, and 2011-12). This pooled approach yields a harmonized dataset large enough to follow successive birth cohorts through different stages of working life and to

compare age groups within each survey year. Covering a period in which the Gross Enrollment Ratio nearly tripled, the data permit an exploration of how the wage gap may evolve for both younger graduates and older workers.

By documenting how the college-versus-high-school wage gap varies across cohorts and age groups, this study offers descriptive evidence for ongoing discussions about higher-education expansion and labor-market segmentation. The remainder of the article proceeds as follows: Section 2 surveys prior research on returns to tertiary education; Section 3 lays out the empirical model; Section 4 describes the NSS data and variable construction; Section 5 presents the core findings; Section 6 presents a spatial decomposition of the young–old premium gap and a parallel measure of zone-level college supply; Section 7 discusses limitations and future research directions; and Section 8 concludes.

2 Literature Review

Early debates on how education raises wages turned on the classic tension between human-capital theories (e.g., Becker, 1964) and signaling models (e.g., Spence, 1973). While the former view posits that schooling imparts skills that boost productivity, the latter suggests education may simply reveal high-ability workers. Empirically, the Mincerian earnings function (Mincer, 1974) became a standard tool, regressing log wages on schooling and experience to estimate returns to education.

Subsequent work introduced selection considerations, notably through the Roy model (Willis and Rosen, 1979), emphasizing that individuals self-select into education based on expected benefits. This framework highlights the possibility of heterogeneous returns: those with higher ability or better alignment with specific fields may gain more from college than others. Modern empirical methods instrument for schooling or use selection-correction estimators to separate true returns from unobserved ability bias.

The implications of large-scale expansions in higher education have also been studied in a general-equilibrium setting (Heckman et al., 1998). Rapid tertiary growth can temporarily compress wage premiums if the supply of graduates outpaces demand, though skill-biased technological change may later restore or augment these premiums (Acemoglu, 2002). Evidence from countries such as China suggests that wage premiums can rebound once industries upgrade to absorb newly abundant skilled labor (Ou and Zhao, 2016).

However, not all degrees yield equal labor-market outcomes. Lax accreditation and under-regulated private institutions can produce nominal credentials with limited skill acquisition (Agarwal, 2006). Indeed, field of study is often crucial: degrees in engineering, finance, or computer science tend to command higher returns than those in less specialized areas (Balestra and Backes-Gellner, 2017). Cross-country comparisons further underscore the importance of foundational quality. (Bharti and Yang (2021)), for instance, contrast China's emphasis on primary and secondary schooling before a science-oriented tertiary expansion with India's more "top-down" approach, where rapid expansion in college seats sometimes outstripped improvements in earlier schooling levels.

Inequities in access and outcomes also persist. Studies drawing on the experience of historically marginalized groups in India (Hnatkovska et al., 2012) show some reduction in caste-based wage gaps from improved educational attainment, but access to quality insti-

tutions remains uneven (Tilak and Choudhury, 2021). High tuition at private colleges, in particular, may exacerbate social divides (Mitra, 2021).

Empirical approaches to measuring education-related wage gaps typically begin with the Mincer (1974) framework but differ in the choice of omitted schooling group. U.S. studies often take high-school graduates as a base (e.g., Card and Lemieux, 2001); policy-oriented work sometimes pools all workers with ≤HS schooling together (e.g., Azam, 2010); and full-ladder specifications distinguish multiple tiers of education (e.g., Goldin and Katz, 2008 in the U.S.; Duraisamy, 2002 in India). Cohort-specific extensions (e.g., Kijima, 2006) further explore how premiums vary across birth cohorts under changing supply-demand conditions.

Finally, global evidence highlights that long-run wage growth aligns more closely with the accumulation of cognitive skills than simple enrollment rates (Hanushek and Woessmann, 2008). Where systems expand tertiary education without strengthening foundational schooling, the labor-market payoff may weaken. Illustrative cases from East Asia and Latin America suggest varied outcomes depending on baseline school quality, regulation, and demand for skilled labor (Ou and Zhao, 2016).

In sum, human-capital and signaling theories form the cornerstone of research on education and wages. Contemporary work stresses selection, institutional quality, and general-equilibrium effects, demonstrating that while college degrees can boost earnings, these gains hinge on the interplay of student ability, institutional standards, and labor-market conditions. For India, where private provision continues to expand rapidly, ongoing research aims to disentangle whether increasing graduate supply will sustain robust returns or yield more modest wage differentials as credentials proliferate.

3 Empirical Specification

This section develops five complementary specifications that together portray the association between higher education and weekly wages in the National Sample Survey (NSS).

3.1 Baseline Mincer Regression

We begin by asking how weekly wages differ between college graduates and individuals whose highest qualification is Grade 12, after accounting for age, gender, caste, religion, and state of residence. To answer that question, we estimate a parsimonious Mincer equation:

$$\ln w_{ist} \;=\; \beta_0 \;+\; \beta_1\,\mathrm{College}_{ist} \;+\; \mathbf{X}'_{ist}\boldsymbol{\delta} \;+\; \alpha_s \;+\; \varepsilon_{ist}, \tag{1}$$

where w_{ist} is the real weekly wage of individual i in state s and survey round t. The indicator $\mathrm{College}_{ist}$ equals 1 for a bachelor's degree or higher and 0 for exactly Grade 12. The control vector $\mathbf{X}_{ist}$ contains age, age^2, and dummies for sex, caste, and religion—individual-specific characteristics that are largely time invariant and unrelated to innate academic ability yet capture well-documented wage differentials (e.g., between men and women or across caste groups). By conditioning on these covariates, we aim to isolate the portion of the wage gap plausibly attributable to schooling rather than to these background factors. State fixed effects α_s absorb time-invariant labor-market conditions and policies within each state.

The ordinary least squares (OLS) coefficient β_1 represents a conditional mean difference:

$$\beta_1 \;=\; \mathbb{E}\big[\ln w_{ist} \mid \text{College} = 1, \mathbf{X}_{ist}, \alpha_s\big] \;-\mathbb{E}\big[\ln w_{ist} \mid \text{College} = 0, \mathbf{X}_{ist}, \alpha_s\big], \tag{2}$$

averaged over the observed distribution of $\mathbf{X}_{ist}$ and α_s. In plain terms, β_1 indicates by what percentage weekly wages differ, on average, between otherwise similar individuals who do and do not hold a college degree.

A caveat remains: unobserved traits—such as ability, family wealth, or school quality—may influence both college attendance and wages. Unless the zero-conditional-mean assumption $\mathbb{E}[\varepsilon_{ist} \mid \text{College}_{ist}, \mathbf{X}_{ist}, \alpha_s] = 0$ holds, β_1 should be interpreted as a *partial correlation*, not a causal return to college.

3.2 Two-Part (Cragg) Model

A substantial fraction of high-school and college graduates report zero weekly wages, particularly in the earlier NSS rounds (Tables A.7 and A.8). Among high-school graduates the share earning zero wages falls from about 88 % in 1987-1988 to 72 % in 2011-2012; for college graduates the decline is from roughly 85 % to 55 %. Although these incidences decrease over time (Figure B.2), they remain large enough to require a model that distinguishes non-earners from positive earners.

Following Cragg (1971), let d^*_{ist} be an individual's latent propensity to earn any wage and let y^*_{ist} be the latent *log* wage conditional on working. The two-part system is

$$\text{Participation:} \qquad d^*_{ist} = \mathbf{Z}'_{ist}\boldsymbol{\alpha} + u_{ist}, \qquad d_{ist} = \begin{cases} 1 & \text{if } d^*_{ist} > 0, \\ 0 & \text{otherwise}, \end{cases} \tag{3}$$

$$\text{Wage level:} \qquad y^*_{ist} = \mathbf{X}'_{ist}\boldsymbol{\beta} + v_{ist}, \qquad y_{ist} = \begin{cases} y^*_{ist} & \text{if } d_{ist} = 1, \\ 0 & \text{if } d_{ist} = 0, \end{cases} \tag{4}$$

where $\mathbf{Z}_{ist}$ and $\mathbf{X}_{ist}$ contain college, age, age square, and other controls same as the baseline specification; the errors u_{ist} and v_{ist} are assumed independent. Equation (3) is estimated by Probit, and Equation (4) by truncated regression restricted to positive earners. This separation allows different factors—and different coefficients—to govern (i) the probability of receiving any wage and (ii) the level(in logs) of wages once employed, a flexibility not available in a single-equation Tobit. Related wage applications include Elek et al. (2011) on under-reported earnings and Alam et al. (2015) on child-labor wages.

The Cragg model clarifies whether a college-high-school differential reflects *participation* (entry into paid work) or *premiums* (wage levels among workers). Pooling non-earners with earners in a single regression conflates these channels. A remaining caveat is that zero-wage individuals are treated as labor-force participants; misclassification of true non-participants could bias the participation equation. Refinements such as those in Hnatkovska et al. (2012) could sharpen this distinction, but such extensions are deferred to future work. Appendix Tables A.11 and A.12 provide complementary descriptive evidence on the incidence of zero wages and the occupational composition of those reporting them.

3.3 Cohort and Age-Bin Regressions

A breakdown of the sample by only "young" versus "old" age groups can mask key differences among specific birth cohorts. For instance, a cutoff such as "under 40 vs. over 40" in a single survey year combines individuals born a decade or more apart, even though they may have experienced substantially different educational opportunities and labor-market entry conditions. The 2000s college boom—marked by sharp growth in institutions, seats, and the Gross-Enrollment Ratio (Joshi and Ahir, 2016, Figure B.3)—may have narrowed the college wage premium, though not uniformly across workers. To capture such cohort-specific shifts, we follow each five-year birth cohort through successive survey rounds instead of relying on broad age splits.

To study life-cycle dynamics, individuals born in the same five-year interval form birth cohort c. For each cohort-survey cell (c, t) we estimate

$$\ln w_{icst} = \beta_{0,ct} + \beta_{1,ct}\,\text{college}_{icst} + \mathbf{X}'_{icst}\boldsymbol{\delta}_{ct} + \alpha_s + \varepsilon_{icst}, \tag{5}$$

where i indexes individuals in cohort c, s states, and t survey year. All other variables are defined the same as the baseline specification. Comparing $\beta_{1,ct}$ across survey years traces the evolution of the college premium *within* cohort c as it ages.

Complementing this pseudo-panel, each survey wave is partitioned into five-year age bins $j \in \{20\text{–}24, 25\text{–}29, \dots\}$. For every age bin j in year t we run

$$\ln w_{ijst} = \beta_{0,jt} + \beta_{1,jt}\,\text{college}_{ijst} + \mathbf{X}'_{ijst}\boldsymbol{\delta}_{jt} + \alpha_s + \varepsilon_{ijst}, \tag{6}$$

where i indexes individuals in age bin j. The same control set $\mathbf{X}_{ijst}$ and state fixed effects α_s are employed. The coefficients $\beta_{1,jt}$ give a cross-sectional snapshot of how the college premium varies by age within survey year t.

Taken together, Equations (5) and (6) offer a unified picture: the cohort regressions document how a given cohort's premium changes over successive rounds, while the age-bin regressions reveal contemporaneous differences across age groups in any single wave.

3.4 Alternative Reference Categories

Since the reference category for college can importantly shape the measured gap, we further experiment with alternate comparison groups. Rather than comparing college graduates exclusively with high-school completers, we progressively expand the base category to include individuals who stopped schooling at primary, secondary, or any level below the tertiary threshold. Table 3.1 illustrates an example of how these alternate reference-group definitions are organized in practice. Analyzing how the premium changes when the comparison group shifts helps clarify whether observed differentials are driven primarily by contrasts with those who complete high school or extend to lower schooling levels.

Table 3.1: Alternative Reference Groups for the *College* Indicator

Specification	**Reference group (base category)**
College vs. Higher Secondary	Workers who completed *exactly* Grade 12
College vs. Higher Secondary + Primary to Middle	Workers who completed any class from Grade 1 through Grade 12 but never entered tertiary education
College vs. Higher Secondary + Primary to Middle + Literate Below Primary	All workers with some schooling below the bachelor's level, including literate adults without formal credentials
College vs. All Non-College (incl. Illiterates)	Everyone whose highest attainment is below a bachelor's degree, illiterate individuals included

3.5 Sub-sample Analysis: Gender Heterogeneity

India's labor market has long shown a pronounced gender divide. Across the five NSS Employment-Unemployment rounds analyzed, the female labor-force participation rate (FLFP) among 21- to 40-year-olds declined from roughly 32 percent to 23 percent, while the male rate hovered near 75 percent, dipping only slightly in the most recent round.(Bhat et al., 2021) One factor behind the overall decline is that more young people now remain in education longer than in the past.

Against this backdrop we ask whether the college wage premium differs by gender. We start by mapping the share of people with at least a bachelor's degree by sex, survey year, and age group (21–25, 26–30, 31–35, 36–40), as shown in Figure B.6. Among women aged 21–25, college attainment rises from about 4 percent in 1987–1988 to roughly 12 percent by 2011–2012; for their male counterparts the increase is from 7 to 19 percent. Since 2000, the share of college graduates has climbed much more steeply for both women and men, hinting that the surge in university places ("expansion effect") reflects a growing demand for college degrees among younger people.

B.7 repeats the exercise after restricting the sample to individuals no longer enrolled in any educational institution at the time of the survey—the population used in our wage regressions. Predictably, the youngest male cohort shows the lowest share of college completers in the labor force, many having pursued further study. The pattern is less pronounced for women, suggesting earlier labor,market entry or higher rates of permanent withdrawal.

To quantify gender-specific returns we re-estimate Equation (1) separately for women. Denote the coefficient on *College* in the female sample by β_1^{F}. An analogous regression for men yields β_1^{M}. Comparing these two coefficients shows whether college closes or widens the wage gap between men and women with the same level of education.

Together, the baseline Mincer regressions, the two-part Cragg model, cohort-survey pseudo-panels, alternative reference-group comparisons and sub-sample analysis provide a comprehensive view of the wage-education relationship. Although these specifications are descriptive rather than causal, taken as a whole they consistently highlight enduring wage differentials by schooling across cohorts and survey waves.

4 Data

We use data from five large quinquennial rounds of the National Sample Survey(NSS) Unemployment and Employment: Rounds 43 (1987-1988), 50 (1993-1994), 55 (1999-2000), 61 (2004-2005), and 68 (2011-2012) [1] These surveys are conducted throughout the year in multiple sub-rounds to minimize seasonality and cover the entire country, except for a few remote regions. The NSS follows a multistage stratified sampling design, with villages or urban blocks as first-stage units and households as ultimate stage units. Within each sampled unit, households are then selected through a second stage of stratification based on livelihood classes or hamlet-group subdivisions. Although specific stratification procedures have changed somewhat over time, the NSS provides household-level multipliers/weights to account for these differences and ensure representative estimates Hnatkovska et al. (2012).

Each of these five rounds includes roughly 120,000 households and about 600,000 individuals, capturing household composition, demographic characteristics, labor-force participation, and wages. To construct a consistent working sample, we retain only individuals aged 20-60 who were not enrolled in any educational institution at the time of the survey, reported positive wages, and indicated their highest educational attainment. Focusing on this age range helps capture prime working-age individuals who have completed formal schooling. The population age distribution as well as education attainment levels are shown in Figure B.1

Tables A.1 and A.2 provide the corresponding unweighted and weighted values of our working sample. We apply NSSO sampling weights with minor adjustments to align the pooled data across rounds. Following Hnatkovska et al. (2012), **all wages are converted to real weekly values** using official state poverty lines with rural Maharashtra in 1983 as the base. Observations missing critical variables such as education or wage data are excluded.

We recode education into six categories: (1) illiterate, (2) literate but below primary, (3) primary to secondary, (4) higher secondary, (5) graduate/diploma, and (6) postgraduate. Because some early rounds do not distinguish between secondary and higher secondary, those categories are combined in (4). Categories (5) and (6) are merged as "college" in the entire analysis although removing postgraduates (6) separately does not alter the main findings (results not included in the paper). We also control for sex, caste, religion, sector, and landholding quantiles, which proxy for household economic status.

Our primary outcome is weekly wages. The NSS measures daily earnings (excluding in-kind payments) for the week preceding the survey, whether the individual worked full-time or half-time, and we sum these daily amounts into a weekly wage value. Since our core focus is on the wage gap between college and high-school graduates, short-term fluctuations in weekly earnings are assumed to be relatively small in these occupations. Descriptive statistics in Tables A.3 and A.4 show a growing share of college-educated individuals and a widening raw wage gap over time. Figure B.5 further shows that the percentage of younger college graduates is higher than older college graduates, especially after the year 2000.

[1] NSSO (1988),NSSO (1994), NSSO (2000), NSSO (2005) and NSSO (2012)

5 Results

This section reports five sets of results: Baseline Mincer Regressions and Extensions (Section 5.1); Two-Part (Cragg) Model (Section 5.2); Cohort and Age-Bin Analyses (Section 5.3); Alternative Reference Groups (Section 5.4); and a Sub-sample Analysis of Gender Heterogeneity (Section 5.5).

5.1 Baseline Mincer Regressions and Extensions

We begin by estimating equation (1) with the logarithm of weekly wages as the dependent variable and high-school graduates as the omitted category for the *college* dummy.Table 9.1 reports the baseline results.PanelA restricts the sample to younger workers (ages 20–40) and PanelB to older workers (ages 41–60). Across all five National Sample Survey (NSS) rounds the coefficient on *college* is positive, highly significant, and rises over time. For younger workers it climbs from roughly 0.10 in 1987–1988 to about 0.63 in 2011–2012, implying an increase from a 10 percent to an 80 percent wage gap after exponentiation. The older cohort shows a similar, though slightly attenuated, trajectory. These estimates are consistent with Azam (2010), who analyze hourly wages in NSS data through 2004 and likewise find that the premium is larger for younger than for older workers—a gap that persists through 2012.

The other covariates behave as expected. In 1987–1988 the female coefficient is slightly positive (0.09), consistent with a small, select group of wage-earning women, but it turns negative (≈ -0.16 to -0.24) once female labor-force participation broadens in the late 1990s. Scheduled-Caste/Tribe and Muslim indicators remain negative throughout, reflecting persistent social wage gaps.

To test whether the raw premium merely captures location or wealth advantages, we re-estimate (1) after adding an urban dummy and household land-holding quartiles. Table 9.2 lists the adjusted *college* coefficients; full regression output appears in Tables A.9 and A.10. The premium declines modestly, especially in later waves, suggesting that part of the gap is correlated with urban residence or family assets.

Urban residence and land ownership are closely linked to both educational attainment and earnings, but the direction of this relationship is ambiguous. Families with greater initial resources are more likely to live in cities or to hold land and are also more likely to finance college, whereas college graduates often move to urban labor markets and may purchase land later in life. These variables therefore mix background advantages with life-cycle outcomes and may also proxy for unobserved traits—such as ability—that we cannot measure directly. When we include them, the college coefficient falls modestly, indicating that part of the raw association between college and wages is shared with urban and land status. We treat this decline as a descriptive adjustment rather than evidence of a "net" effect, because the underlying mechanisms remain intertwined.

We also estimate regressions analogous to Table 9.1 but replace the continuous age terms (age_{ist} and age^2_{ist}) with discrete five-year age-group dummies. The resulting estimates, summarized in Tables A.5 and A.6, confirm that the main patterns are robust to this alternative functional form for age.

In a further specification, we include district fixed effects instead of state effects for the last three rounds (1999–2000, 2004–2005, and 2011–2012), as district identifiers are not available

in the earlier data.Table 9.3 presents these results.Relative to state fixed effects, district-level dummies absorb finer-grained regional variation in local labor markets, infrastructure, and average wages.Consequently, each coefficient in this specification captures how wages differ, on average, *within* the same district, rather than merely within the same state.In practice, this design helps net out unobserved factors such as district-specific cost-of-living differences or historical development patterns that might otherwise inflate or mask the college premium at broader geographies.

Despite this stricter adjustment, the key coefficient on remains positive and statistically significant across all three rounds, with estimates broadly comparable to those in the baseline. Indeed, for both the younger and older age groups, the estimated premium continues to rise over time. This consistency suggests that the strong correlation between college completion and higher wages is not solely an artifact of state-level heterogeneity but also appears within narrower local labor-market comparisons. Overall, the district fixed-effects results reinforce our main conclusion that college graduates enjoy a substantially higher wage level, even after accounting for more localized unobserved differences.

Additional specifications for the dependent variables were examined for the baseline regressions reported in Tables A.13 and A.14. The findings align with the original baseline results, showing rising wage premiums across the five NSS round.

5.2 Two-Part (Cragg) Model

A large share of individuals in each National Sample Survey (NSS) round report zero weekly earnings (see Tables A.7 and A.8), we estimate a two-part model following Cragg (1971) as described by equations (3) and (4). The first hurdle is a Probit for the probability that weekly earnings are positive; the second hurdle is an OLS regression for the *log* of real weekly wages among earners. The unconditional mean wage for each schooling group equals the product of its fitted participation probability and its predicted conditional wage, thereby capturing both extensive and intensive margins.

Table 9.4 presents expected weekly earnings and the *college-to-high-school earnings ratio*. Panel A (ages 20–40) shows that high-school graduates move from about INR 8.8 in 1987-88 to roughly INR 95 by 2011-12, whereas college graduates increase from INR 17.5 to INR 366. The implied ratio of expected earnings rises from 2.0 to 3.9, peaking at 4.1 in 2004-05; a young college graduate therefore earned twice a high-school graduate's wage in the late-1980s, four times as much at mid-decade, and just under four times as much by 2011-12.

Panel B (ages 41–60) exhibits the same pattern on a larger scale: the ratio climbs from 2.3 in 1987-88 to 7.0 in 2004-05, before easing to 6.5 in 2011-12. Although the multiple is always higher for older workers, the late-period decline suggests a modest narrowing of the unconditional gap, consistent with an expanding supply of college graduates moving into occupations that pay below the earlier college-graduate norm.

Baseline Mincer regressions, estimated only on positive earners, yield log-wage gaps of 0.35-0.55, equivalent to earning ratios of roughly 1.4-1.7. Introducing the extensive margin through the two-part model more than doubles these ratios in every NSS round and magnifies them at the 2004-05 peak. Similar findings—where conditional regressions understate overall differentials once zero earners are considered—appear in Elek et al. (2011), Alam et al. (2015), and Madden (2006).

Three mechanisms explain the sizeable multiples. First, college graduates are *more likely* than high-school graduates to record any positive wage. Also seen in Table A.11 Second, among those who work, their conditional wages are much higher. Third, the real-wage base for high-school graduates remains very low, so even moderate absolute gaps produce large ratios. The post-2004 fall in the earnings multiple may reflect both a slight rise in high-school pay and the absorption of new college graduates into lower-paying white-collar and service jobs.

In sum, the two-part estimates reveal a substantially larger unconditional earnings gap than standard wage regressions, while the recent tapering of the ratio may hint at early evidence of college-educated unemployment as tertiary-education supply expands.

5.3 Cohort and Age-Bin Analyses

The 2009–2012 surge in university enrollment was followed by large wave of new graduates into India's labor market. Workers born between 1980 and 1994—who were 20–30 years old in 2011–12—entered the workforce at the height of this expansion, when competition among degree-holders appeared strongest. Figure 9.1 plots the estimated college wage premium for each five-year birth cohort across the NSS survey rounds.

For cohorts born before 1945 the premium remains small—never exceeding 0.30—and even turns slightly negative for the 1927–1931 group in 1987–88. By contrast, cohorts born between 1950 and 1966 exhibit a steady rise: the 1957–1961 group moves from 0.14 in 1993–94 to 0.56 in 2011–12, while the 1962–1966 group rises from -0.01 to 0.61. During the years when these groups studied, college places were relatively scarce, and their wage advantage persists well into mid-career.

The largest premium is still posted by the 1972-1981 cohort (0.73 in 2011-12). However, among the youngest workers, the 1987-1991 cohort in 2011-12 enjoys a higher premium than the youngest cohorts in any earlier survey wave—around 0.40 versus considerably lower figures in previous rounds—showing that new graduates retain a notable earnings edge even if it is smaller than that of their slightly older peers. [2] Across successive rounds the college premium tends to be largest for the youngest cohorts, with the 2011-12 survey showing the widest gap yet. This elevated premium could stem more from slower wage growth (or outright declines) among high-school graduates than from higher pay for degree-holders, so the pattern should be interpreted with caution.

In addition to our cohort-based regressions, we disaggregated individuals by *age bins* (e.g. 20–24, 25–29, 30–34, and so on) within each survey year. In other words, for each NSS round, we collapse the data into five-year age groups and estimate the log-wage premium for each bin. Figure 9.2 shows how the college premium β varies as a function of age within each survey year. This differs conceptually from our cohort approach: the age-bin specification captures a *cross-sectional* snapshot of different age groups at a given time, whereas the cohort figure follows the *same* birth cohorts as they age across multiple survey rounds. Viewed together, the two perspectives paint a richer picture. The age-bin figure highlights

[2] All cohort-survey cells satisfy a reliability threshold of at least 100 observations and a combined sampling weight of 100,000, except those born 1947–1951 in 2011–12 ($N = 83$, weight 139,600); omitting this cell leaves the pattern unchanged.

the distinct life-cycle pattern of college payoffs in *each year's cross-section* (e.g. the 2011 line shows which ages then enjoyed the highest returns), while the cohort figure isolates how the premium evolves for a single cohort as it moves from, say, age 25–29 in one survey to 30–34 in the next. By combining these approaches, we gain a more comprehensive understanding of both *within-year cross-sectional* differences by age and *across-year changes* for each birth cohort.

Across rounds, the curves in Figure effig:age-stateFE share a common bell shape: college premiums rise from the early 20s, peak around the early-to-mid 30s, and then taper off as workers move into their late 40s and 50s—an inverted-U life-cycle profile. Later survey years crest slightly later and at substantially higher levels, suggesting that the post-2000 expansion of tertiary education translated into larger and more persistent returns for younger cohorts before gradually converging toward those of older workers.

5.4 Alternative Reference Groups for College

Figure 9.3 traces how the college wage premium evolves as we broaden the comparison group—from only high-school graduates (top-left) to all non-degree holders (bottom-right). Two distinct patterns emerge. First, the wage premium gap between the 1987 survey and subsequent surveys widens as the reference group expands to include less-educated workers. This reflects an economy in 1987 characterized by fewer graduate-level opportunities and more compressed wages, contrasting sharply with later decades where graduates enjoyed significantly higher pay while wages for less-educated groups stagnated. Second, each survey shows an inverted-U shape: cohorts born in the late 1950s and 1960s, who completed college before degrees became common, consistently experience the largest premiums. Conversely, the premium shrinks among older cohorts and falls notably for those born after 1980, coinciding with the early-2000s enrollment surge that significantly increased the supply of new graduates. The pronounced drop in the 2011 curve for these younger cohorts likely indicates declining premiums due to excess supply rather than technological shifts favoring lower skills. Furthermore, as the reference group broadens, the youngest cohorts' wage premiums decrease further, suggesting that the observed rise in college premiums may reflect declining wages among high-school graduates rather than genuine increases in college graduate pay. This trend implies that the college degree has become akin to the previous status of a high-school diploma in earlier decades.

The age-bin panels in Figure 9.4 shift the focus from cohorts to work experience. For all surveys except 1987, the college premium begins modestly among workers aged 20-24, increases steadily through their thirties and early forties, and either levels off or slightly declines after age 50. The trajectory is steeper and reaches a higher peak as the reference group expands to include lower educational levels, primarily because older workers without degrees or with minimal education see limited wage growth in their later careers. The relative positioning of the survey years across the four age-bin panels shifts as the reference group broadens, revealing periods and conditions under which the college premium expanded.

With a narrow comparison group of only high-school graduates (top-left panel), the college premium consistently rises across survey waves, with each newer survey's curve positioned above the previous one. Expanding the group to include dropouts (top-right panel) diminishes the gap between the 2004 and 2011 surveys, especially for workers older than

45, indicating that much of the premium growth had already occurred by 2004. Including literate workers with below-primary education (bottom-left panel) further reduces the gap between 2004 and 2011 for mid-career workers, suggesting minimal subsequent growth for this group. Finally, when the broadest group—comprising all non-college workers, including illiterates—is used as the reference (bottom-right panel), the 2004 curve surpasses the 2011 curve at every age, highlighting the substantial increase in premiums that occurred primarily between 1999 and 2004, with limited growth afterward. Throughout, the 1987 curve remains consistently low and flat, underscoring a compressed wage structure prior to economic reforms. Collectively, these panels emphasize the 1999-2004 period as the critical window of significant growth in college wage differentials, followed by stabilization for younger cohorts and only modest increases for older workers in subsequent years.

5.5 Sub-sample Analysis: Gender Heterogeneity

In the backdrop of declining female labor-force participation and rising educational attainment discussed before, Table 9.5 examines how the college wage premium has evolved across sexes and age groups. As younger cohorts have increasingly pursued higher education, particularly since 2000, we observe a marked rise in the college wage premium over time. For younger women (ages 20-40), the premium grew significantly from a modest and statistically insignificant 0.032 log points in 1987-88 to a highly significant 0.690 log points by 2011-12. Similarly, young men's premiums increased steadily from 0.132 to 0.620 log points over the same period.

For older cohorts (ages 41-60), a similar but less dramatic rise occurred. Male premiums grew from an insignificant 0.062 log points in 1987-88 to a robust 0.577 log points in 2011-12, while female premiums experienced fluctuations but ultimately increased significantly from 0.013 to 0.447 log points. These findings align with the previously documented expansion in college attainment among younger individuals, highlighting how the surge in educational investments—potentially driven by greater demand for college degrees—has translated into higher wage premiums. The increase in women's premiums appears especially pronounced among younger cohorts, suggesting that despite declining participation rates, women who remain employed and hold college degrees have benefited substantially in India's evolving labor market.

6 A Spatial Decomposition of the Premium Gap

The cohort and age-bin patterns documented above point to a widening premium for younger workers relative to older workers, with the gap most pronounced in the post-2000 rounds. This section reports a decomposition of the young–old college premium gap across India's six regional zones, alongside a parallel measure of zone-level college supply.

A standard interpretation of the young–old premium gap follows Card and Lemieux (2001): when the supply of college graduates expands rapidly, the relative supply of *young* college-educated workers rises while the stock of *older* college-educated workers (whose schooling decisions were made in an earlier regime) is essentially fixed. If young and old

workers are imperfect substitutes within skill groups, the premium for young college graduates should compress relative to that of older college graduates wherever supply expanded most aggressively. The zone-level decomposition examines whether this prediction holds in the cross-section.

We aggregate states into the six zones using 1987 administrative boundaries; states created by post-1987 administrative splits are folded back into their parent states. Table 6.1 lists the full classification. Following the age-band convention of Keng et al. (2017), we redefine young workers as ages 26–35 and old workers as ages 46–55. For each zone–round–age-band cell satisfying a reliability threshold of at least 50 college and 50 high-school observations, we compute the cell-specific premium as the difference in mean log weekly wages between college graduates and individuals whose highest qualification is Grade 12.

Table 6.1: NSS Zone Classification of Indian States and Union Territories

Zone	**States and Union Territories**
Central	Madhya Pradesh,[a] Rajasthan, Uttar Pradesh[a]
South	Andhra Pradesh,[a] Goa, Karnataka, Kerala, Lakshadweep, Pondicherry, Tamil Nadu
West	Dadra & Nagar Haveli (and Daman & Diu), Gujarat, Maharashtra
East	Andaman & Nicobar Islands, Assam, Bihar,[a] Orissa, Tripura, West Bengal
North	Chandigarh, Delhi, Haryana, Himachal Pradesh, Jammu & Kashmir,[a] Punjab
North-East	Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim

[a] States created by post-1987 administrative splits are pooled into their parent states throughout the analysis to maintain constant boundaries: Chhattisgarh is included in Madhya Pradesh (split 2000); Jharkhand in Bihar (split 2000); Uttarakhand in Uttar Pradesh (split 2000); Telangana in Andhra Pradesh (split 2014); Ladakh in Jammu & Kashmir (split 2019).

Figure 6.1 plots the national young and old college premiums separately across the five NSS rounds. The two series begin essentially overlapping in 1987 (young at 0.15, old at 0.17) and diverge steadily thereafter, with the young premium rising from 0.15 to 0.67 and the old premium rising from 0.17 to 0.51. Both series move upward over the period, but the young premium rises faster, opening up a gap that reaches roughly 0.16 log points by 2011. This decomposition makes clear that the rising premium documented in Section 5.1 reflects movement on both margins, with the young margin rising more steeply. The widening gap between young and old premiums after 1993 is consistent with demand for college-educated labor rising faster than the relative supply of young college graduates over this period.

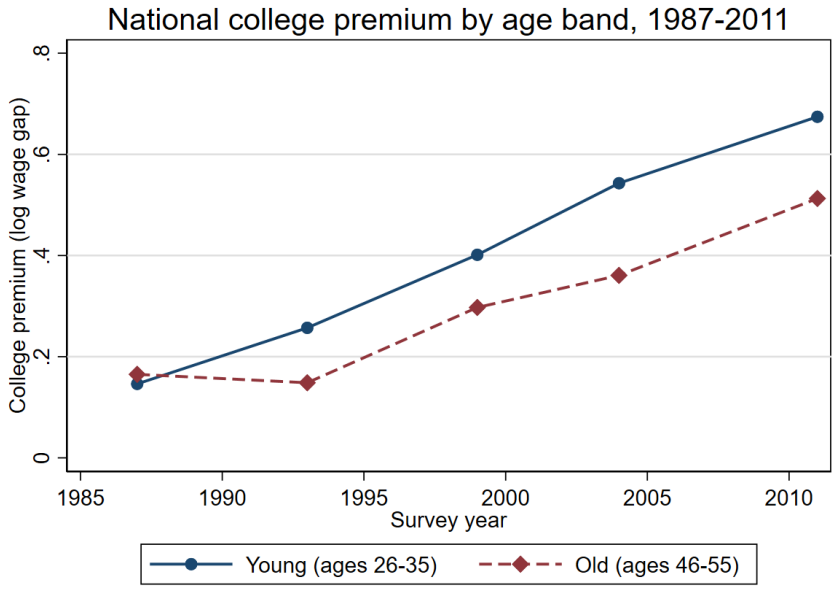


Figure 6.1: National college wage premium by age band, 1987–2011. Young workers are aged 26–35; old workers are aged 46–55. Cells require at least 100 college and 100 high-school observations.

Figure 6.2 disaggregates the same exercise across India's six regional zones. Three patterns emerge. First, young premiums rise across all six zones, consistent with the national trend, though the magnitude varies substantially: from 0.25 to 0.78 in North and from 0.27 to 0.71 in East, but with much smaller movement in North-East. Second, the trajectories of the old premium are markedly more heterogeneous. In North and East the old premium remains relatively flat (rising only modestly), opening a substantial young–old gap. In Central and South the old premium rises in parallel with the young series and converges by 2011, leaving little gap. Third, North-East shows young and old premiums that move closely together throughout, both at low levels.

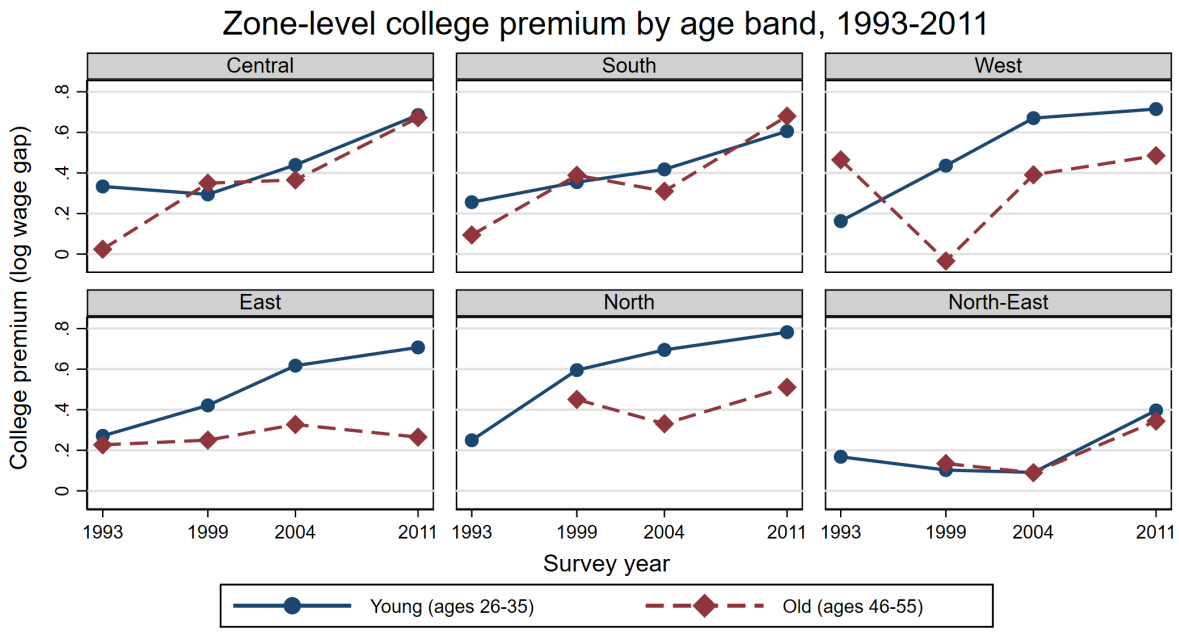


Figure 6.2: Zone-level college wage premium by age band, 1993–2011. Zones are arranged in descending order of post-1995 college-stock growth (top row: Central, South, West; bottom row: East, North, North-East). Cells require at least 50 college and 50 high-school observations.

The zone where young and old converge to the same premium by 2011 (Central) is also the zone with the largest absolute increase in college supply, while the zone where the young premium pulls farthest ahead of the old premium (East) experienced more modest supply growth. To examine the relative-supply mechanism directly, we construct a parallel zone-year panel of college supply using the All India Survey on Higher Education (AISHE) institution-level file. AISHE reports the year of establishment for each affiliated college as of the survey snapshot date; we aggregate to zone-year cumulative stocks separately by management category (private, government, university-managed). State names are pooled to 1987 boundaries to align with the NSS panel.

Figure 6.3 shows cumulative college stocks by zone and management type from 1950 to 2024, with a reference line at 1995 to mark the inflection point of India's private-college expansion. The zones are arranged in the same top-to-bottom and left-to-right ordering as Figure 6.2, so the two figures can be read as paired panels. Three observations follow. First, the post-1995 acceleration is largest in absolute terms in Central (rising to roughly 16,000 colleges by 2024) and South (roughly 13,500), moderate in West (around 8,000), and modest in East and North (each around 3,000–5,000). The North-East stock remains under 500 throughout the period. Second, in every zone with a substantial post-1995 increase, the expansion is dominated by private institutions; the government-managed stock remains nearly flat across the entire 1950–2024 window in all six zones. Third, the timing of the acceleration is similar across the four high- and medium-expansion zones, beginning in the late 1990s and steepening through the 2000s, suggesting a common policy trigger rather than

zone-specific drivers.

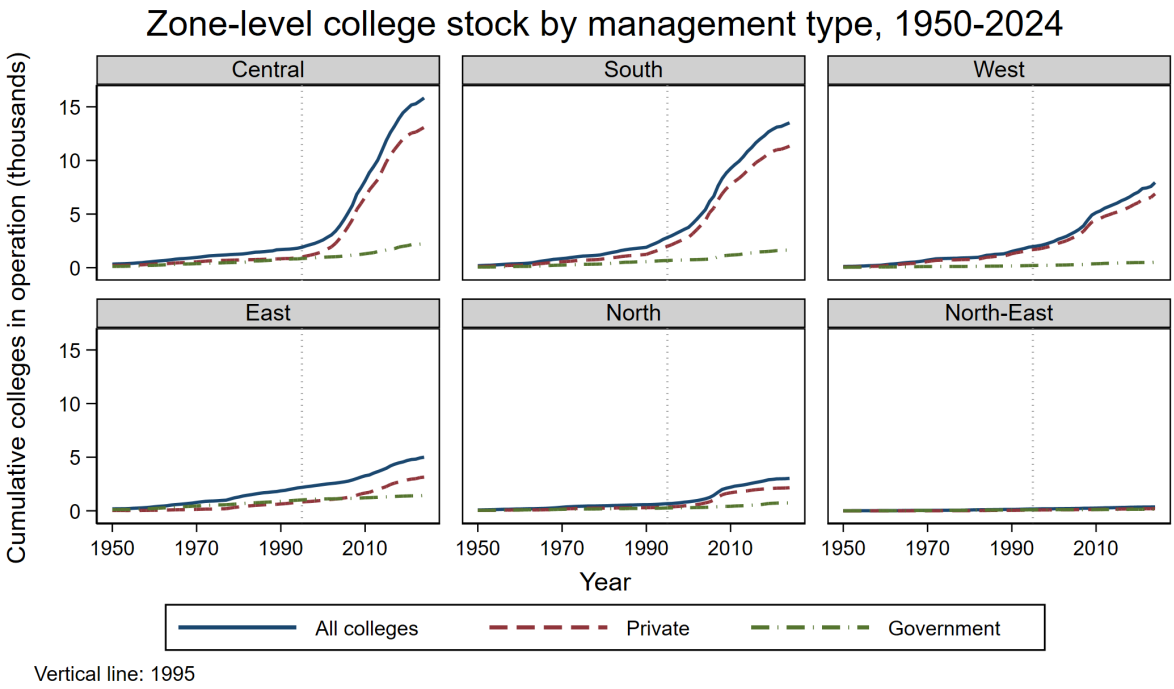


Figure 6.3: Zone-level cumulative college stock by management type, 1950–2024. Zone ordering matches Figure 6.2: top row contains the three zones with the largest post-1995 expansion, bottom row contains the three with smaller expansion. Vertical dotted line marks 1995. Source: AISHE, college-level establishment dates aggregated to zone-year using NSS-1987 state boundaries.

Comparing Figures 6.2 and 6.3 yields the following cross-zone correspondence. Central and South—the zones with the largest absolute post-1995 supply expansion—exhibit converging young and old premiums by 2011, with old premiums rising to meet or exceed young premiums by the end of the sample. East and North—with smaller supply expansion—retain substantial young–old gaps through 2011. West shows an intermediate pattern on both dimensions. The single zone inconsistent with the relative-supply prediction is North-East, where minimal supply expansion is paired with a small young–old gap; both young and old premiums in North-East are low and move together throughout. The North-East cells are also the smallest in the data, and we are cautious about reading much into this panel.

The cross-zone pattern is consistent with a relative-supply mechanism for five of six zones. We emphasize that this is a descriptive correspondence, not an identified effect: zones differ in many dimensions other than college supply (industrial composition, urbanization, demand-side shocks), and the comparison uses raw stocks rather than per-capita supply. The exercise nonetheless suggests that the spatial variation in college expansion is a substantively meaningful.

7 Limitations and Future Work

This study analyses the 1987–2011 period in order to capture the post-1992 privatisation era and the subsequent college boom. One limitation is that the survey rounds examined only partly overlap with the graduates most affected by the expansion. Extending the window to include more recent Periodic labor Force Survey (PLFS) rounds would illuminate longer-run effects. However, harmonising PLFS and NSS data remains a methodological challenge that future research must address.

A second limitation arises from how wages are calculated. Weekly wages are constructed by summing daily wages over seven days without accounting for hours worked, thereby conflating part-time jobs with full-time low pay. Constructing hourly earnings—following Hnatkovska et al. (2012)—would sharpen the distinction between wage rates and work intensity. Moreover, candidates preparing for competitive exams outside formal institutions are recorded as non-employed, which may understate effective labor-force participation.

A further measurement issue concerns the aggregation of bachelor's and master's degrees. Early NSS rounds do not distinguish between these two levels of qualification, so if the share of master's graduates increases at the same time bachelor's degrees become more widespread, estimates of the college premium may be biased upward. Separating these qualifications whenever survey detail permits would clarify how undergraduate expansion feeds into postgraduate enrollment and subsequent earnings.

Finally, employing more advanced econometric techniques and models could yield further insights. Principal-component methods, for instance, can concisely capture variations in household background and local economic conditions, thereby enhancing precision while mitigating over-fitting. Beyond reduced-form approaches as mentioned above (similar studies (Ou and Zhao, 2016)), structural or general-equilibrium frameworks—such as the approach outlined in Li et al. (2025)—could illuminate the ways in which greater graduate supply reverberates through firm behavior, educational choices, and occupational sorting.

In short, broader data coverage, refined outcome measures, and identification strategies that exploit cross-state variation in college supply would greatly enhance our understanding of how India's tertiary boom has affected wage differentials.

8 Conclusion

The NSS data suggest that India's post-1990s expansion of tertiary enrollment has been accompanied by a widening college-to-higher-secondary wage gap, increasing from roughly 0.10 in 1987-88 to about 0.63 in 2011-12. Two-part Cragg model estimates reveal even larger unconditional differentials. Notably, for the youngest cohorts entering the labor market after the mid-2000s boom, the premium appears to plateau, indicating that the rapid increase in graduate supply might be moderating the wage premium. The spatial decomposition in Section 6 reinforces this reading: the young–old premium gap is smallest in the regions with the largest post-1995 college expansion (Central and South) and largest in regions where supply growth was modest (East, North), a cross-zone correspondence consistent with a relative-supply mechanism in five of six zones.

It is crucial to recognize that the growing wage gap between college and high-school-educated workers does not necessarily indicate that college graduates are inherently better off. Instead, it might reflect worsening outcomes for high school graduates or a shifting baseline where college education is increasingly becoming the minimum requirement. As more young individuals in India pursue higher education, understanding the opportunity costs associated with educational investments remains essential.

9 Main Tables and Figures

Table 9.2: College Wage Premium by Survey—Baseline Specification with Urban and Landholding Controls

	1987–1988	1993–1994	1999–2000	2004–2005	2011–2012
Age 20–40 (Young)	0.113*	0.271***	0.304***	0.470***	0.533***
Age 41–60 (Old)	0.057	0.267***	0.351***	0.426***	0.530***

Notes. Entries are coefficients on the college-degree indicator from the baseline log-wage specification that additionally controls for sector (urban/rural) and household landholding quartiles (see Appendix Tables A.9 and A.10 for full results). Robust standard errors in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9.1: Baseline Regression: College Wage Premium by Survey Round

	Survey Years				
Variable	1987-1988	1993-1994	1999-2000	2004-2005	2011-2012
Panel A: Age 20-40 (Young Workers)					
College degree	0.097**	0.268***	0.355***	0.497***	0.632***
	(0.042)	(0.032)	(0.029)	(0.025)	(0.027)
Age	-0.022	0.107***	0.170***	0.200***	0.157***
	(0.041)	(0.030)	(0.029)	(0.023)	(0.024)
Age squared	0.000	-0.001**	-0.002***	-0.002***	-0.002***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Female	0.090**	0.371***	-0.165***	-0.149***	-0.244***
	(0.041)	(0.032)	(0.036)	(0.030)	(0.031)
Muslim	-0.037	-0.017	-0.192***	-0.048	-0.080*
	(0.064)	(0.056)	(0.053)	(0.038)	(0.041)
Other religions	0.142*	-0.075	0.088*	-0.017	0.044
	(0.074)	(0.065)	(0.051)	(0.051)	(0.048)
SC/ST	-0.105*	-0.029	-0.068*	-0.141***	-0.179***
	(0.058)	(0.045)	(0.039)	(0.030)	(0.036)
Constant	5.037***	3.518***	2.000***	2.160***	3.180***
	(0.620)	(0.452)	(0.434)	(0.353)	(0.357)
R-squared	0.069	0.116	0.241	0.252	0.249
Observations	5,113	6,962	9,823	11,448	13,474
Panel B: Age 41-60 (Older Workers)					
College degree	0.061	0.156***	0.324***	0.404***	0.565***
	(0.072)	(0.060)	(0.035)	(0.032)	(0.036)
Age	-0.201*	0.571***	0.054	0.131**	0.249***
	(0.119)	(0.134)	(0.058)	(0.056)	(0.064)
Age squared	0.002*	-0.006***	-0.000	-0.001*	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female	0.121*	0.502***	-0.217***	-0.209***	-0.089**
	(0.071)	(0.066)	(0.049)	(0.039)	(0.040)
Muslim	-0.148	-0.277**	-0.173***	-0.073	-0.129**
	(0.107)	(0.127)	(0.050)	(0.048)	(0.062)
Other religions	0.167	-0.232**	-0.052	0.042	-0.058
	(0.107)	(0.110)	(0.049)	(0.046)	(0.059)
SC/ST	-0.089	0.092	-0.103**	-0.084**	-0.093**
	(0.087)	(0.083)	(0.049)	(0.036)	(0.047)
Constant	9.724***	-7.779**	4.079***	3.014**	0.344
	(2.954)	(3.237)	(1.427)	(1.367)	(1.553)
R-squared	0.098	0.125	0.141	0.150	0.157
Observations	1,711	2,246	5,763	6,924	7,894

Note. Dependent variable is the log of real weekly wages. Reference categories: high school (education), male (gender), Hindu (religion), and Non-SC/ST (caste). Robust standard errors in parentheses. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 9.3: College Wage Premium Across Last Three Rounds with District Fixed Effects

Variable	**1999-2000**	**2004-2005**	**2011-2012**
Panel A: Age 20-40 (Young Group)			
College	0.351***	0.460***	0.545***
	(0.028)	(0.025)	(0.024)
Age	0.173***	0.191***	0.147***
	(0.027)	(0.021)	(0.021)
Age squared	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)
Female	-0.183***	-0.141***	-0.229***
	(0.035)	(0.029)	(0.029)
Muslim	-0.148***	-0.038	-0.028
	(0.049)	(0.042)	(0.040)
Other religions	0.095*	-0.011	0.068
	(0.049)	(0.049)	(0.042)
SC/ST	-0.029	-0.115***	-0.151***
	(0.037)	(0.029)	(0.032)
Constant	2.054***	1.851***	3.296***
	(0.409)	(0.366)	(0.374)
R-squared	0.347	0.362	0.382
Observations	9,823	11,448	13,474
Panel B: Age 41-60 (Old Group)			
College	0.321***	0.398***	0.513***
	(0.029)	(0.031)	(0.034)
Age	0.108*	0.137**	0.253***
	(0.055)	(0.057)	(0.054)
Age squared	-0.001	-0.001**	-0.002***
	(0.001)	(0.001)	(0.001)
Female	-0.207***	-0.206***	-0.112***
	(0.044)	(0.039)	(0.039)
Muslim	-0.124**	-0.095**	-0.113*
	(0.051)	(0.046)	(0.060)
Other religions	-0.069	0.020	-0.103*
	(0.045)	(0.048)	(0.062)
SC/ST	-0.047	-0.116***	-0.065
	(0.044)	(0.036)	(0.042)
Constant	2.511*	3.452**	0.349
	(1.353)	(1.361)	(1.309)
R-squared	0.309	0.281	0.318
Observations	5,763	6,924	7,894

Note. Dependent variable is log of real weekly wages. Models include district fixed effects. Reference categories: high school (education), male (gender), Hindu (religion), and Non-SC/ST (caste). Robust Standard errors in parentheses; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9.4: Expected Weekly Earnings and College/High-School Earnings Ratio: Two-Part Cragg Model

	HS (Rs./week)	College (Rs./week)	College/HS Ratio
Panel A: Age 20–40			
1987–88	8.81***	17.47***	1.98
	(0.69)	(1.55)	
1993–94	46.81***	129.66***	2.77
	(0.34)	(3.12)	
1999–00	26.79***	96.34***	3.60
	(0.33)	(3.10)	
2004–05	62.36***	253.72***	4.07
	(0.55)	(6.29)	
2011–12	94.94***	365.60***	3.85
	(0.90)	(7.83)	
Panel B: Age 41–60			
1987–88	7.79***	17.67***	2.27
	(0.54)	(1.92)	
1993–94	44.71***	148.80***	3.33
	(0.56)	(6.33)	
1999–00	33.88***	217.31***	6.42
	(0.52)	(6.36)	
2004–05	83.65***	585.94***	7.01
	(1.15)	(12.53)	
2011–12	110.66***	723.58***	6.54
	(2.41)	(18.57)	

Notes: Expected weekly wages equal $\Pr(w > 0) \times \mathbb{E}(\ln w \mid w > 0)$, obtained from a two-part Cragg model (probit first stage; truncated OLS second stage). The final column is the ratio of expected college to expected high-school wages; e.g. a ratio of 2.0 means college graduates earn twice as much as high-school graduates. Robust standard errors are shown in parentheses beneath each wage estimate. All specifications control for age, age^2, gender, caste, religion, and state fixed effects. $^{***}p < 0.001$.

Table 9.5: College Wage Premium by Survey—Sex and Age Group

	1987–1988	1993–1994	1999–2000	2004–2005	2011–2012
Age 20–40 (Young)					
Female	0.032	0.207***	0.380***	0.594***	0.690***
	(0.071)	(0.040)	(0.069)	(0.072)	(0.066)
Male	0.132**	0.332***	0.354***	0.480***	0.620***
	(0.052)	(0.051)	(0.032)	(0.027)	(0.030)
Age 41–60 (Old)					
Female	0.013	0.236**	0.366***	0.180*	0.447***
	(0.129)	(0.095)	(0.104)	(0.101)	(0.100)
Male	0.062	0.075	0.320***	0.427***	0.577***
	(0.088)	(0.071)	(0.036)	(0.033)	(0.038)

Notes. Entries are coefficients on the college-degree indicator in log-wage regressions that control for quadratic age, religion, caste, and survey-year fixed effects. Robust standard errors are shown in parentheses. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

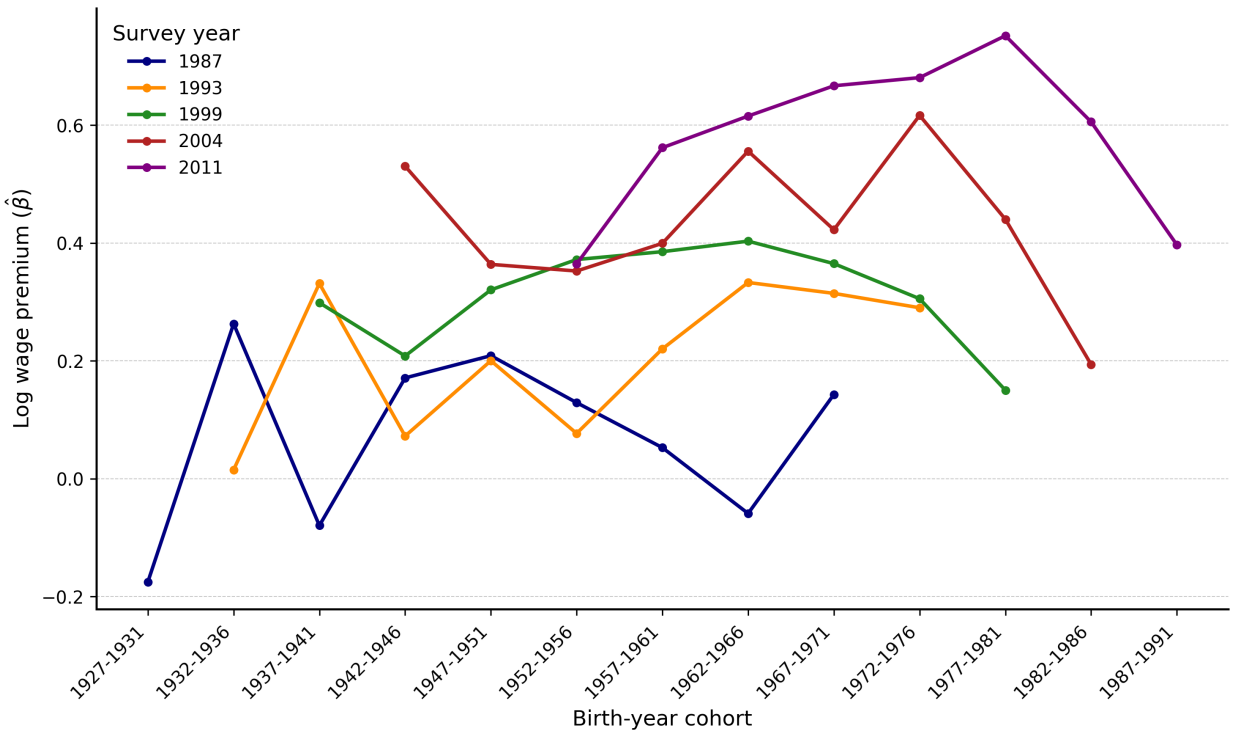


Figure 9.1: College wage premium by birth-year cohort— regression-adjusted with state fixed effects.

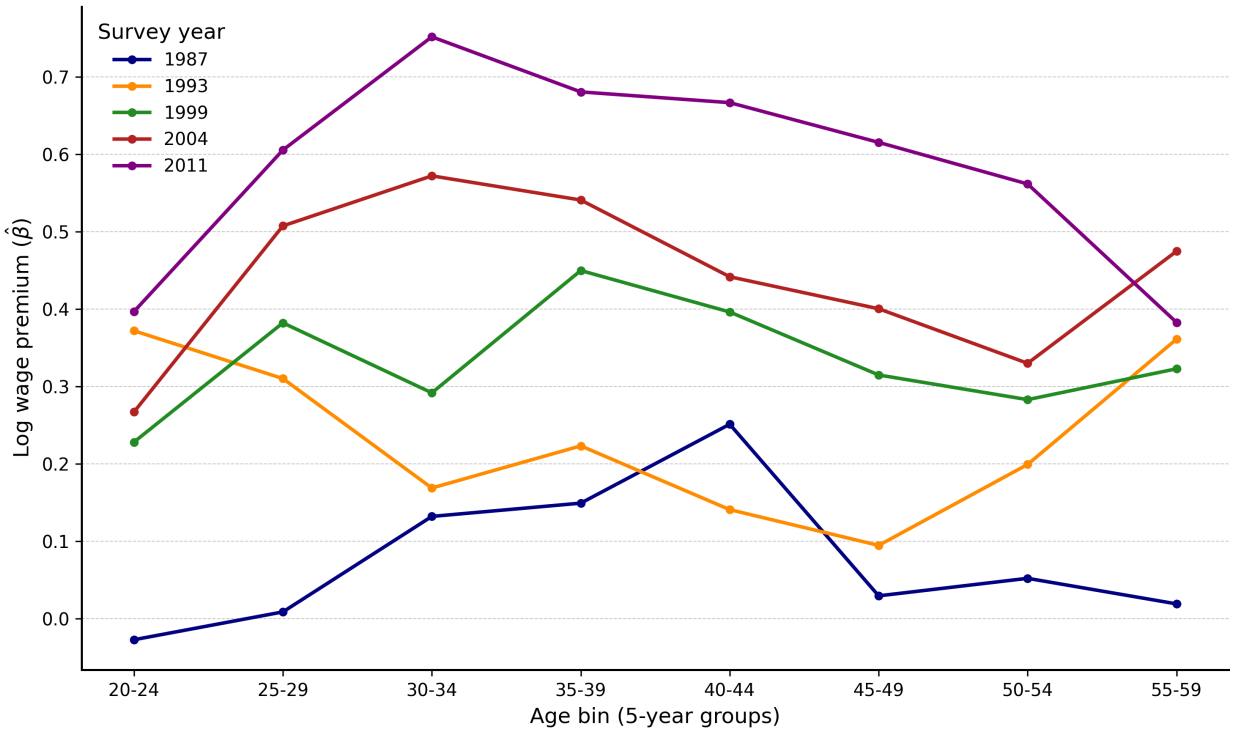


Figure 9.2: College wage premium by age bins — regression-adjusted with state fixed effects.

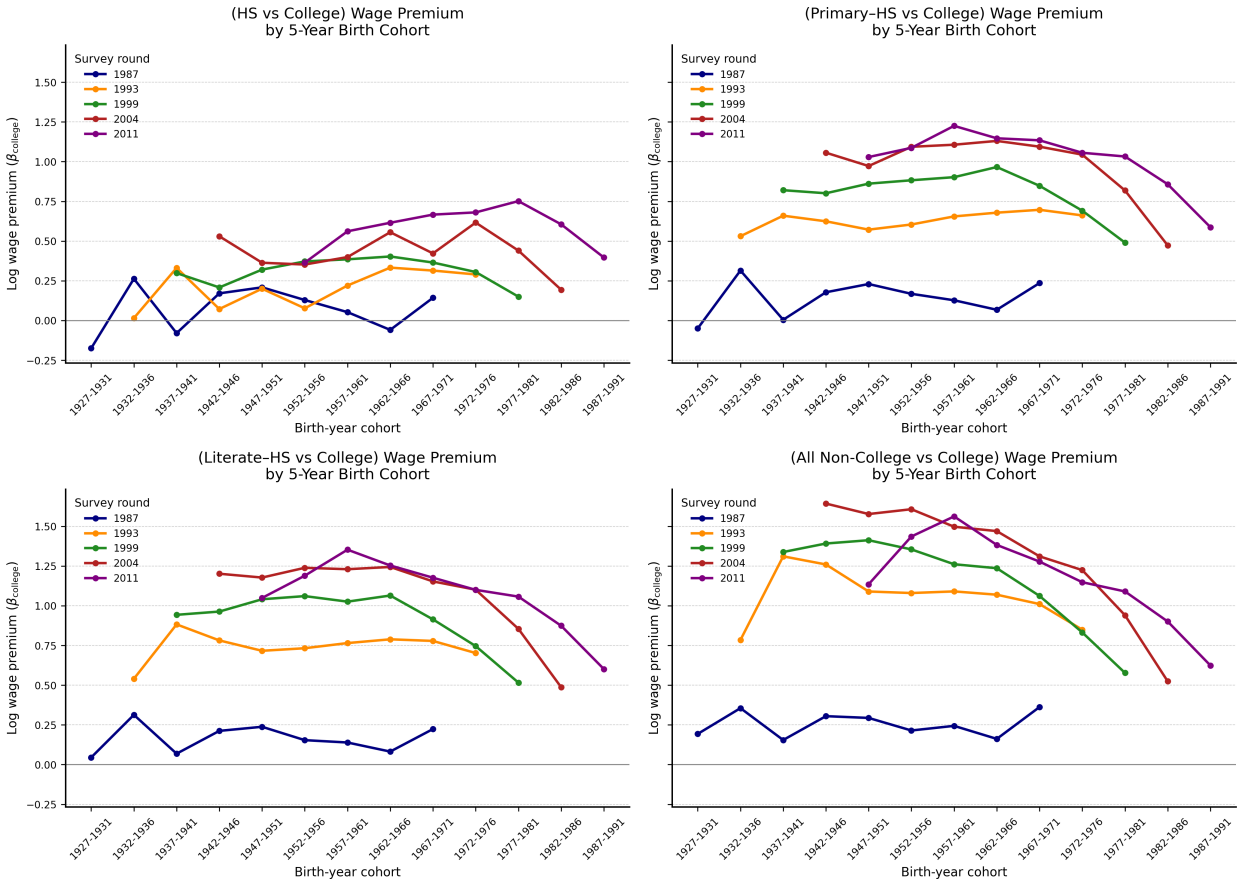


Figure 9.3: Log real wage premiums by 5-year birth cohort, across five NSS survey rounds. Panels show:(i) High-school vs. college; (ii) Primary to HS vs. college; (iii) Literate to HS vs. college; (iv) All non-college vs. college.

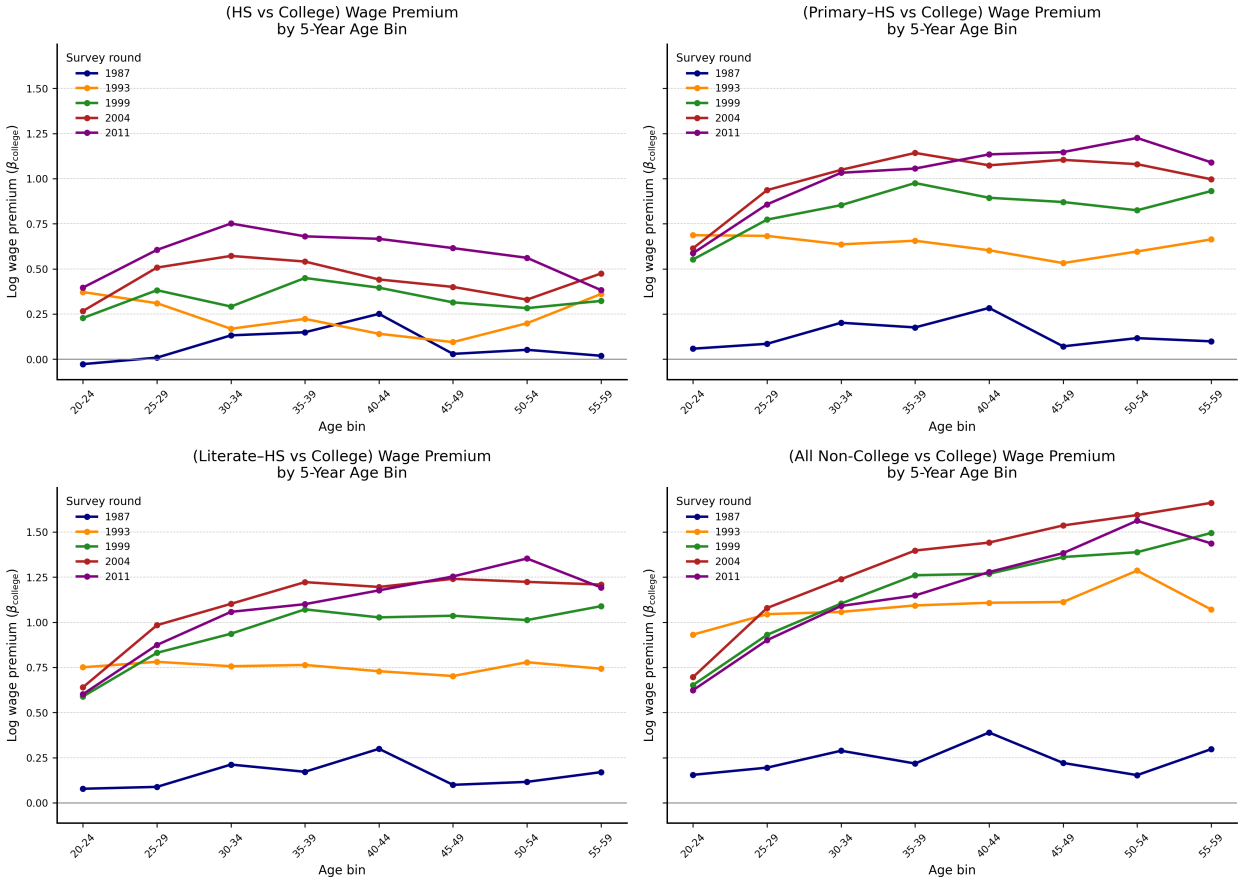


Figure 9.4: Log real wage premiums by 5-year age bin, across five NSS survey rounds. Panels show:(i) High-school vs. college; (ii) Primary to HS vs. college; (iii) Literate to HS vs. college; (iv) All non-college vs. college.

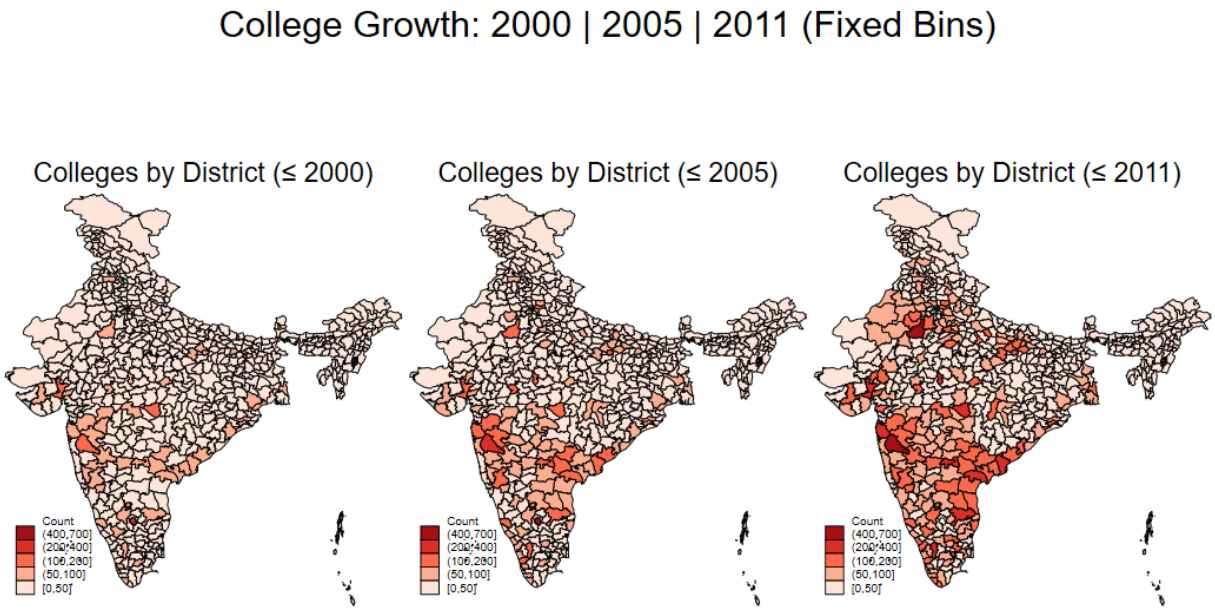


Figure 9.5: College Growth Across Districts: 2000, 2005, and 2011. District-level counts(cumulative) of colleges to each cut-off year. Common fixed bins (0–50, 50–100, 100–200, 200–400, 400–700) make the three maps directly comparable.

A Appendix Tables

Table A.1: Sample by Survey Years: Unweighted Counts and Sample Share

Year	**Unweighted** n	**Sample Share (%)**
1987–1988	25,012	8.05
1993–1994	66,664	21.46
1999–2000	76,602	24.66
2004–2005	75,601	24.34
2011–2012	66,756	21.49

Note. Sample restricted to individuals aged 20-60 who were not enrolled in any educational institution at the time of the survey, who provided their highest educational attainment, and who reported positive wages.

Table A.2: Sample by Survey Years (Weighted): Population Counts and Shares

Year	**Weighted N**	**Weighted Share (%)**
1987-1988	22,670,296	4.17
1993-1994	95,751,208	17.61
1999-2000	126,189,552	23.21
2004-2005	130,065,056	23.92
2011-2012	169,008,608	31.09

Note. Sample restricted to individuals aged 20-60 who were not enrolled in any educational institution at the time of the survey, who provided their highest educational attainment, reported positive wages, and to whom population weights have been applied.

Table A.3: Unweighted Distribution and Mean Wages by Education Group

Year	**College (%)**	**High School (%)**	**Mean Wage (College)**	**Mean Wage (HS)**
1987-1988	36.09	63.91	147.26	127.44
1993-1994	60.26	39.74	524.68	399.36
1999-2000	64.99	35.01	360.25	234.56
2004-2005	70.51	29.49	1013.75	619.27
2011-2012	70.65	29.35	1216.99	611.97

Note. "College" and "High School" shares are computed from the unweighted sample counts for individuals aged 20-60, not enrolled in education at survey, who provided their highest educational attainment and reported positive wages. Mean wages are weekly real earnings in Indian rupees.

Table A.4: Weighted Distribution and Weekly Mean Wages by Education Group

Year	**College (%)**	**High School (%)**	**Mean Wage (College)**	**Mean Wage (HS)**
1987-1988	35.33	64.67	147.26	127.44
1993-1994	59.14	40.86	524.68	399.36
1999-2000	63.28	36.72	360.25	234.56
2004-2005	71.39	28.61	1 013.75	619.27
2011-2012	70.95	29.05	1 216.99	611.97

Note. Shares and weekly real mean wages (in Indian rupees) are computed using population weights for individuals aged 20-60, not enrolled in education at survey, who provided their highest educational attainment and reported positive wages.

Table A.5: Baseline Estimates of College Wage Premium, with five-year age dummies.(Ages 20-40)

	Survey Round				
Variable	1987-1988	1993-1994	1999-2000	2004-2005	2011-2012
Panel A: Age 20-40 (Young Workers)					
College degree	0.098**	0.275***	0.361***	0.516***	0.644***
	(0.042)	(0.032)	(0.028)	(0.025)	(0.027)
Age 25-29 (vs. 20-24)	-0.054	0.247***	0.378***	0.397***	0.296***
	(0.057)	(0.046)	(0.047)	(0.036)	(0.037)
Age 30-34	-0.026	0.422***	0.648***	0.641***	0.466***
	(0.060)	(0.049)	(0.040)	(0.039)	(0.037)
Age 35-39	-0.071	0.504***	0.852***	0.841***	0.587***
	(0.064)	(0.050)	(0.039)	(0.039)	(0.039)
Age 40-44	0.009	0.643***	0.914***	0.933***	0.691***
	(0.092)	(0.063)	(0.061)	(0.048)	(0.058)
Female	0.087**	0.374***	-0.161***	-0.153***	-0.250***
	(0.041)	(0.032)	(0.035)	(0.031)	(0.031)
Muslim	-0.036	-0.014	-0.205***	-0.055	-0.090**
	(0.064)	(0.056)	(0.054)	(0.038)	(0.041)
Other religions	0.139*	-0.076	0.089*	-0.008	0.048
	(0.074)	(0.065)	(0.051)	(0.052)	(0.048)
SC/ST	-0.105*	-0.027	-0.075**	-0.144***	-0.187***
	(0.058)	(0.045)	(0.038)	(0.031)	(0.036)
Constant	4.734***	5.287***	4.844***	5.448***	5.732***
	(0.088)	(0.092)	(0.082)	(0.066)	(0.064)
R-squared	0.070	0.114	0.236	0.238	0.242
Observations	5,113	6,962	9,823	11,448	13,474

Note. Dependent variable is the log of weekly real wages (approx. % changes). Age group dummies use five-year bins, with 20-24 omitted. Reference categories: high school (education), male (gender), Hindu (religion), Non-SC/ST (caste), rural (sector). Robust SEs in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Baseline Estimates of College Wage Premium, with five-year age dummies.(Ages 41-60)

	Survey Round				
Variable	1987-1988	1993-1994	1999-2000	2004-2005	2011-2012
Panel B: Age 41-60 (Older Workers)					
College degree	0.064	0.149**	0.319***	0.407***	0.564***
	(0.071)	(0.060)	(0.035)	(0.032)	(0.036)
Age 45-49 (vs. 40-44)	-0.070	0.062	0.093***	0.080**	0.119***
	(0.091)	(0.070)	(0.034)	(0.035)	(0.042)
Age 50-54	-0.070	0.198**	0.194***	0.267***	0.340***
	(0.089)	(0.081)	(0.048)	(0.032)	(0.042)
Age 55-59	-0.019	-0.054	0.331***	0.302***	0.359***
	(0.104)	(0.110)	(0.040)	(0.037)	(0.049)
Age 60-64	0.303*	-0.921**	-0.304	-0.237	-0.353
	(0.160)	(0.397)	(0.278)	(0.302)	(0.220)
Female	0.113	0.506***	-0.218***	-0.209***	-0.083**
	(0.070)	(0.065)	(0.049)	(0.039)	(0.041)
Muslim	-0.149	-0.285**	-0.181***	-0.070	-0.119*
	(0.107)	(0.129)	(0.050)	(0.048)	(0.061)
Other religions	0.171	-0.253**	-0.038	0.046	-0.036
	(0.108)	(0.109)	(0.050)	(0.045)	(0.060)
SC/ST	-0.091	0.074	-0.111**	-0.085**	-0.104**
	(0.087)	(0.080)	(0.049)	(0.036)	(0.047)
Constant	4.914***	6.041***	5.788***	6.652***	6.850***
	(0.176)	(0.132)	(0.059)	(0.077)	(0.075)
R-squared	0.099	0.124	0.143	0.150	0.167
Observations	1,711	2,246	5,763	6,924	7,894

Note. Dependent variable is the log of weekly real wages (approx. % changes). Age group dummies use five-year bins, with 40-44 omitted. Reference categories: high school (education), male (gender), Hindu (religion), Non-SC/ST (caste), rural (sector). Robust SEs in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: High-School Graduates: Weekly Wage Participation by Survey Round(Year)

Round (Year)	Zero Wage	Positive Wage	Total	% Zero	% Positive
43 (1987-88)	31 668	4 405	36 073	87.81%	12.19%
50 (1993-94)	12 274	3 660	15 934	77.05%	22.95%
55 (1999-00)	11 999	5 457	17 456	68.75%	31.25%
61 (2004-05)	13 269	5 421	18 690	70.99%	29.01%
68 (2011-12)	16 108	6 271	22 379	71.98%	28.02%

Note. Sample restricted to individuals aged 20-60 who were not enrolled in any educational institution at the time of the survey, who provided their highest educational attainment.

Table A.8: College Graduates: Weekly Wage Participation by Survey Round(Year)

Round (Year)	Zero Wage	Positive Wage	Total	% Zero	% Positive
43 (1987-88)	14 363	2 487	16 850	85.26%	14.74%
50 (1993-94)	15 117	5 551	20 668	73.15%	26.85%
55 (1999-00)	14 097	10 132	24 229	58.19%	41.81%
61 (2004-05)	16 743	12 964	29 707	56.37%	43.63%
68 (2011-12)	18 250	15 097	33 347	54.73%	45.27%

Note. Sample restricted to individuals aged 20-60 who were not enrolled in any educational institution at the time of the survey, who provided their highest educational attainment.

Table A.9: Full Specification of Table 8.2 -College Wage Premium by Survey Round (Panel A: Age 20-40)

	Survey Round Years				
Variable	1987-1988	1993-1994	1999-2000	2004-2005	2011-2012
College degree	0.113*	0.271***	0.304***	0.470***	0.533***
	(0.062)	(0.041)	(0.034)	(0.029)	(0.031)
Age	−0.058	0.146***	0.106***	0.178***	0.104***
	(0.058)	(0.038)	(0.032)	(0.026)	(0.026)
Age squared	0.001	−0.002***	−0.001	−0.002***	−0.001**
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Female	0.083	0.400***	−0.165***	−0.136***	−0.237***
	(0.061)	(0.047)	(0.047)	(0.034)	(0.033)
Muslim	−0.128	0.036	−0.202***	−0.034	−0.087*
	(0.093)	(0.073)	(0.075)	(0.043)	(0.046)
Other religions	0.048	−0.062	0.079	−0.055	0.044
	(0.111)	(0.080)	(0.068)	(0.057)	(0.057)
SC/ST	−0.022	−0.036	−0.125***	−0.110***	−0.138***
	(0.087)	(0.059)	(0.045)	(0.034)	(0.039)
Urban	0.146	−0.126**	−0.001	0.065**	0.329***
	(0.103)	(0.056)	(0.038)	(0.032)	(0.034)
Landholding, 2nd quartile	0.069	0.032	−0.028	−0.052	0.071*
	(0.075)	(0.054)	(0.044)	(0.037)	(0.037)
Landholding, 3rd quartile	−0.298***	−0.118	−0.178***	−0.116***	0.049
	(0.105)	(0.073)	(0.056)	(0.044)	(0.049)
Landholding, 4th quartile	−0.316***	−0.042	0.071	−0.102**	0.160***
	(0.113)	(0.067)	(0.054)	(0.047)	(0.050)
Constant	5.575***	2.992***	2.975***	2.524***	3.815***
	(0.884)	(0.563)	(0.492)	(0.390)	(0.396)
R^2	0.123	0.123	0.254	0.257	0.270
Observations	2,420	4,456	5,850	8,679	10,466

Notes. Dependent variable is ln(weekly real wage). Controls: sector (urban/rural), landholding quartiles. Reference categories: high school, male, Hindu, non-SC/ST, rural, 1st land quartile. Robust SEs in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Full Specification of Table 8.2 -College Wage Premium by Survey Round (Panel B: Age 41-60)

	Survey Round Years				
Variable	1987-1988	1993-1994	1999-2000	2004-2005	2011-2012
College degree	0.057	0.267***	0.351***	0.426***	0.530***
	(0.123)	(0.070)	(0.036)	(0.038)	(0.040)
Age	−0.222	0.520***	0.004	0.135**	0.262***
	(0.192)	(0.184)	(0.060)	(0.068)	(0.070)
Age squared	0.002	−0.005***	0.000	−0.001	−0.002***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Female	0.201*	0.814***	−0.168**	−0.182***	−0.087**
	(0.113)	(0.099)	(0.067)	(0.047)	(0.042)
Muslim	0.098	−0.253	−0.153***	−0.087	−0.053
	(0.211)	(0.174)	(0.054)	(0.056)	(0.069)
Other religions	0.151	−0.151	−0.029	0.070	0.009
	(0.149)	(0.098)	(0.065)	(0.045)	(0.055)
SC/ST	−0.080	0.095	−0.134**	−0.115***	−0.102**
	(0.124)	(0.083)	(0.060)	(0.042)	(0.050)
Urban	0.145	−0.802***	−0.054	−0.038	0.137***
	(0.161)	(0.103)	(0.042)	(0.033)	(0.040)
Landholding, 2nd quartile	0.087	0.012	0.116***	−0.049	0.142***
	(0.121)	(0.095)	(0.044)	(0.040)	(0.041)
Landholding, 3rd quartile	−0.012	−0.315**	0.001	−0.102*	0.109*
	(0.196)	(0.127)	(0.053)	(0.053)	(0.056)
Landholding, 4th quartile	−0.042	−0.128	−0.000	0.010	0.122*
	(0.171)	(0.101)	(0.055)	(0.046)	(0.071)
Constant	9.772**	−6.528	5.094***	2.924*	−0.184
	(4.710)	(4.442)	(1.500)	(1.652)	(1.708)
R^2	0.171	0.210	0.171	0.160	0.163
Observations	764	1,501	3,770	5,320	6,776

Notes. Dependent variable is ln(weekly real wage). Controls: sector (urban/rural), landholding quartiles. Reference categories: high school, male, Hindu, non-SC/ST, rural, 1st land quartile. Robust SEs in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Probability of Reporting Zero Cash Earnings by Education and NSS Round

NSS round	**College**	**High school**
43 (1987–88)	0.867	0.936
50 (1993–94)	0.736	0.752
55 (1999–00)	0.573	0.720
61 (2004–05)	0.574	0.745
68 (2011–12)	0.553	0.716

Table A.12: Occupation Mix of Zero-Wage Reporters (%) by Education and NSS Round[†]

	College graduates			**High-school graduates**		
NSS round	White-collar	Blue-collar	Agriculture	White-collar	Blue-collar	Agriculture
43 (1987-88)	62.7	25.0	12.3	8.3	28.3	63.4
50 (1993-94)	61.9	20.9	17.2	6.3	29.8	63.9
55 (1999-00)	38.5	32.9	28.6	5.4	27.4	67.2
61 (2004-05)	33.6	44.7	21.7	10.9	35.1	54.0
68 (2011-12)	50.5	23.8	25.7	13.5	32.7	53.8

[†] Rows show the percentage distribution of zero-wage reporters across three broad occupation groups within each education category.

Table A.13: Robustness—Alternative Specification Using log(Wage + 1) as Dependent Variable

	Survey year / NSS round				
	1987-88 (43)	1993-94 (50)	1999-2000 (55)	2004-05 (61)	2011-12 (68)
Panel A: Young (20-40)					
College	0.109***	0.204***	0.518***	0.839***	1.073***
	(0.020)	(0.043)	(0.046)	(0.049)	(0.056)
Panel B: Old (41-60)					
College	0.160***	−0.074	0.863***	1.191***	1.442***
	(0.041)	(0.087)	(0.087)	(0.095)	(0.090)

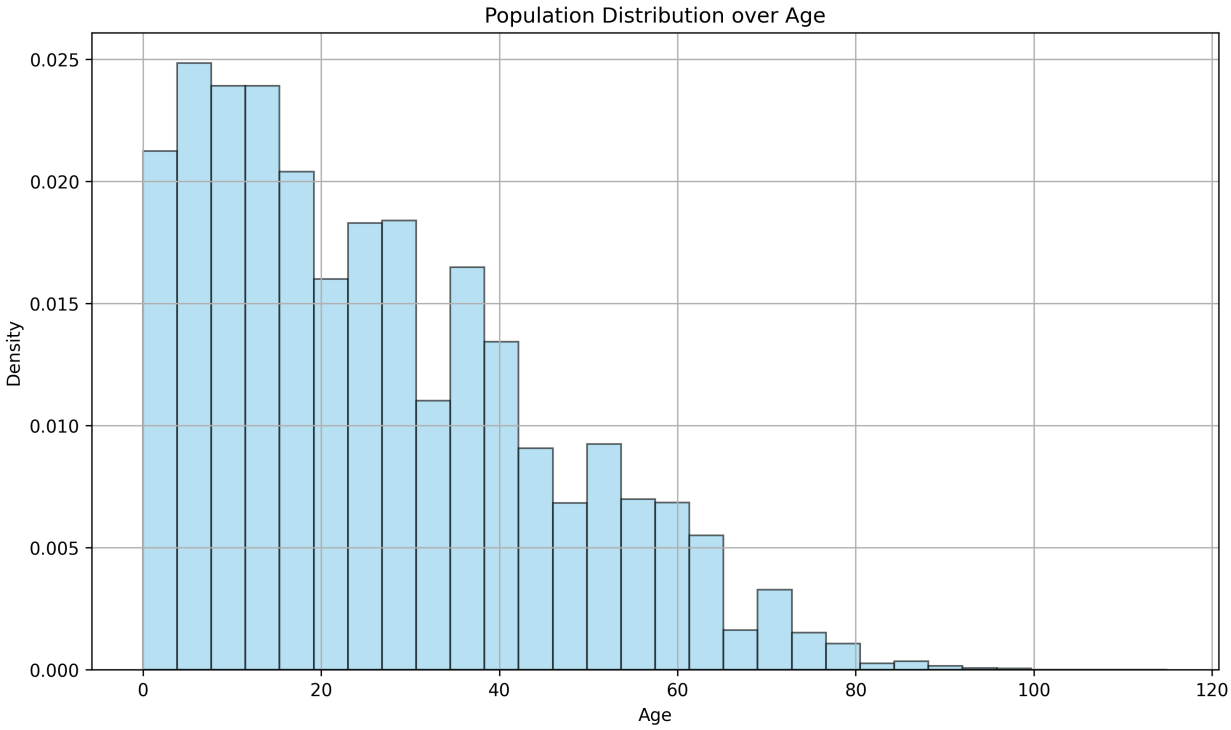
Standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

Table A.14: Robustness—Alternative Specification Using Inverse Hyperbolic Sine of Wage as Dependent Variable

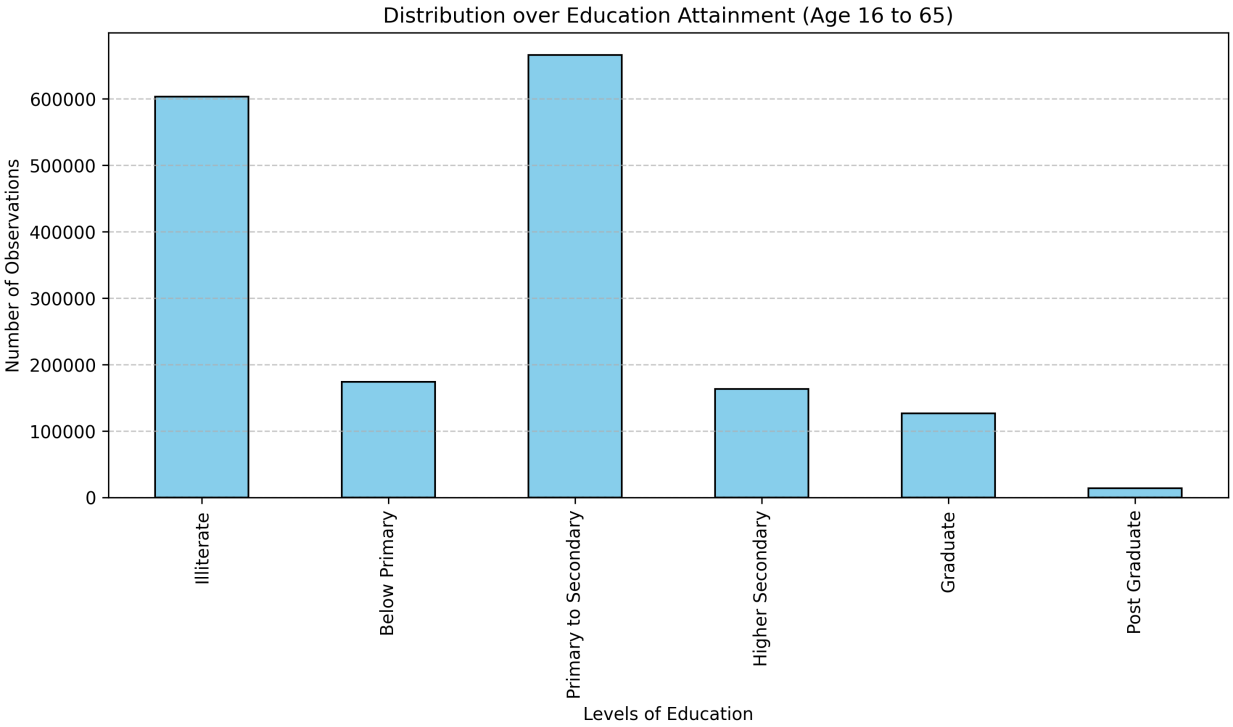
	Survey year / NSS round				
	1987-88 (43)	1993-94 (50)	1999-2000 (55)	2004-05 (61)	2011-12 (68)
Panel A: Young (20-40)					
College	0.124***	0.222***	0.572***	0.916***	1.169***
	(0.023)	(0.048)	(0.052)	(0.054)	(0.062)
Panel B: Old (41-60)					
College	0.182***	−0.087	0.947***	1.293***	1.564***
	(0.047)	(0.096)	(0.097)	(0.104)	(0.099)

Standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

B Appendix Figures



(a) Population distribution over age.



(b) Distribution over education attainment.

Figure B.1: Basic sample distributions for age and education.

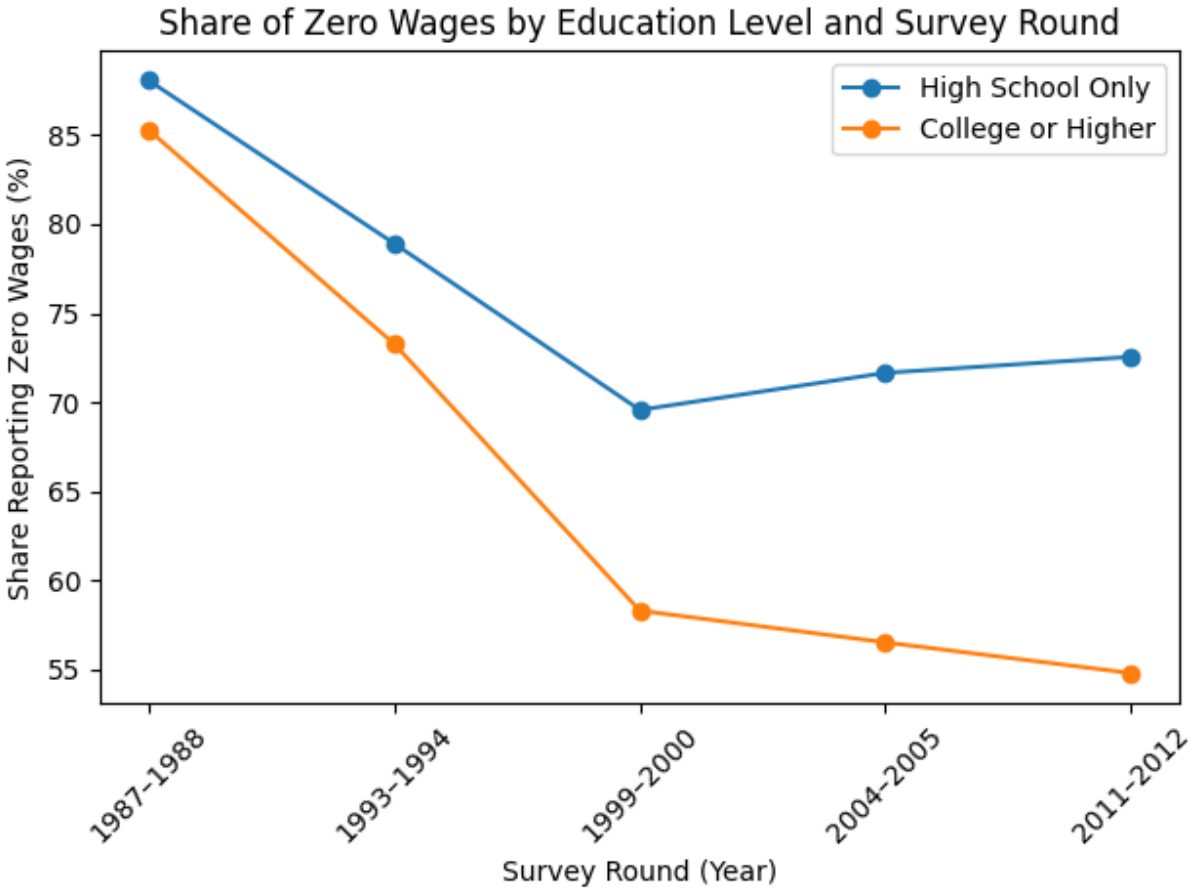
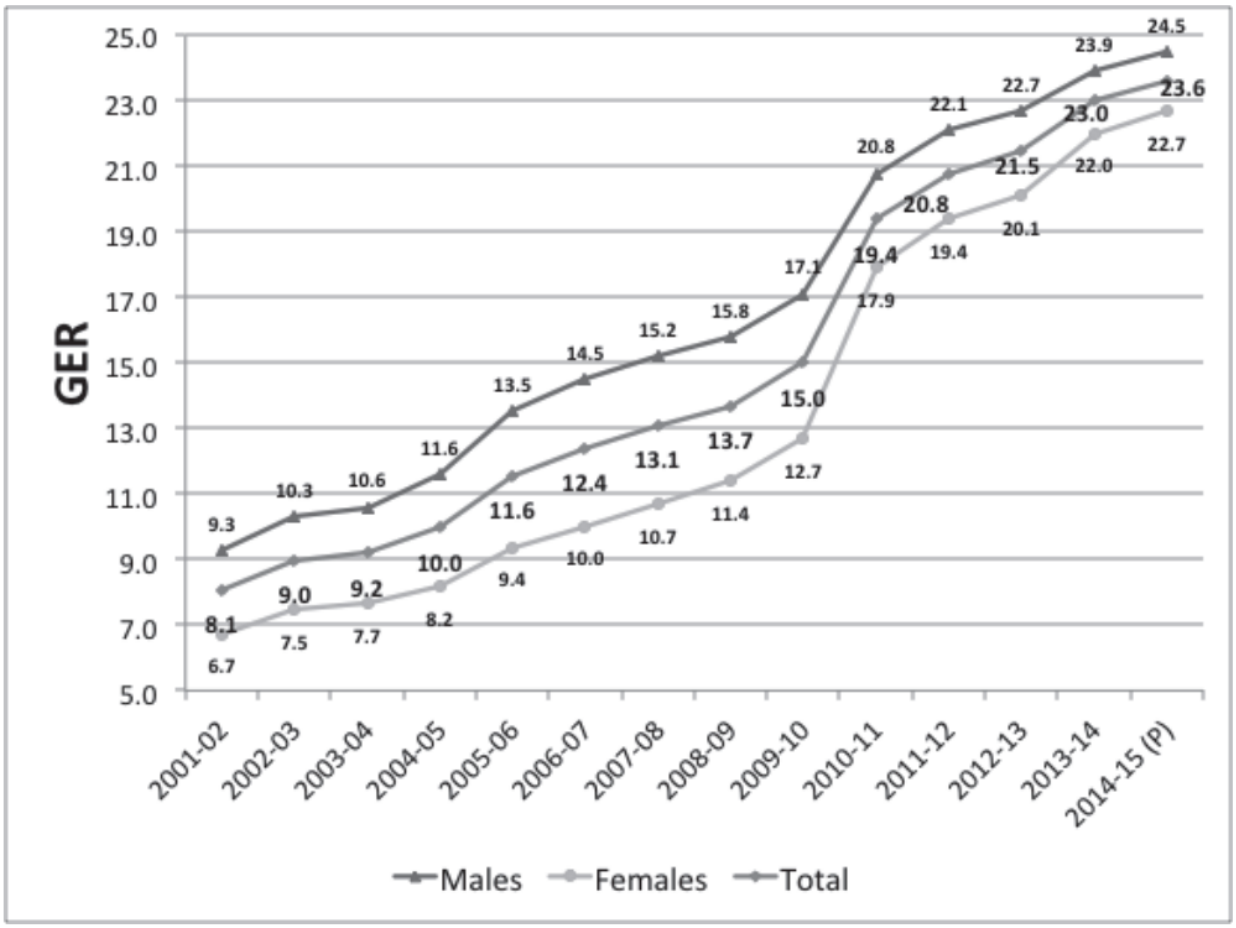


Figure B.2: Share of Zero Wages by Education Level and Survey Round



Source: MHRD (2014a, 2015a, 2015b, 2015c)

Figure B.3: Gross Enrollment Ratio as presented in Joshi and Ahir (2016)

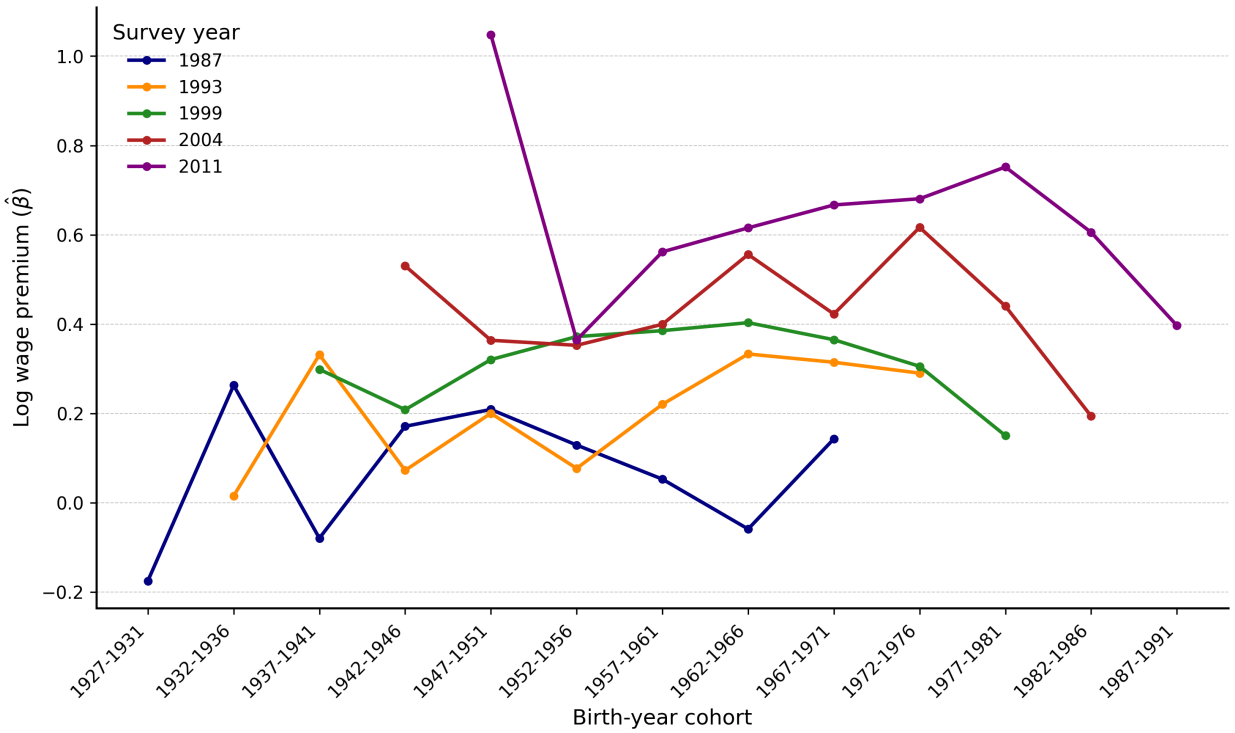


Figure B.4: College wage premium by birth cohort (same as Figure 8.1) without weight and observation restrictions

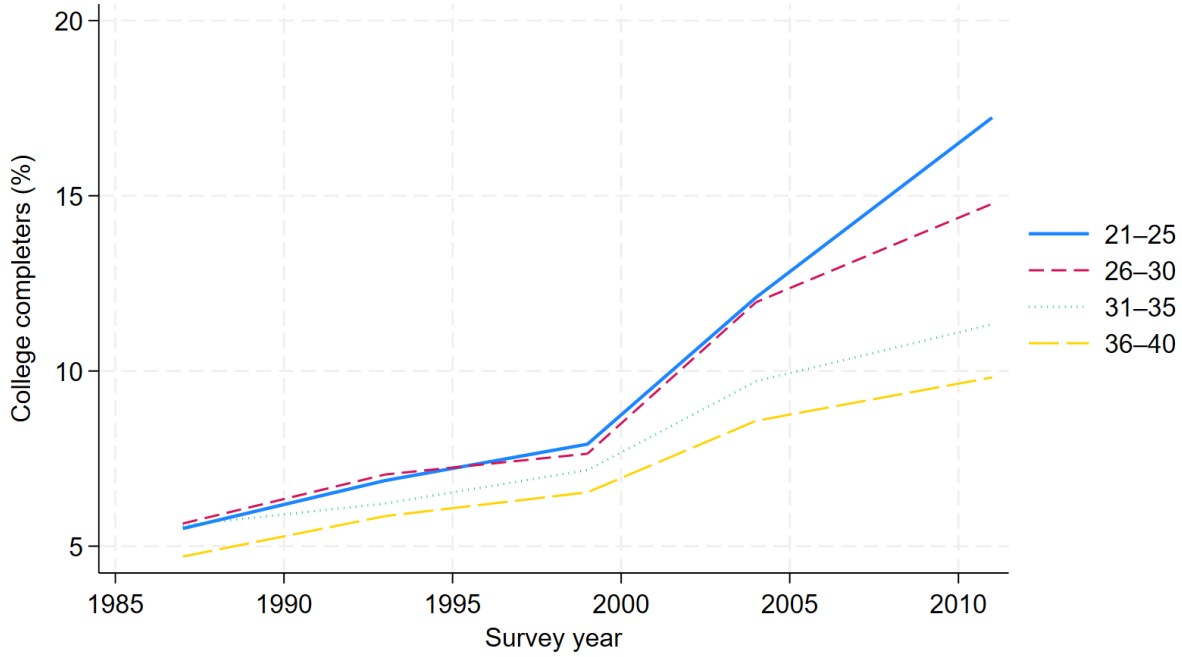


Figure B.5: Weighted-percentage bands of college completers

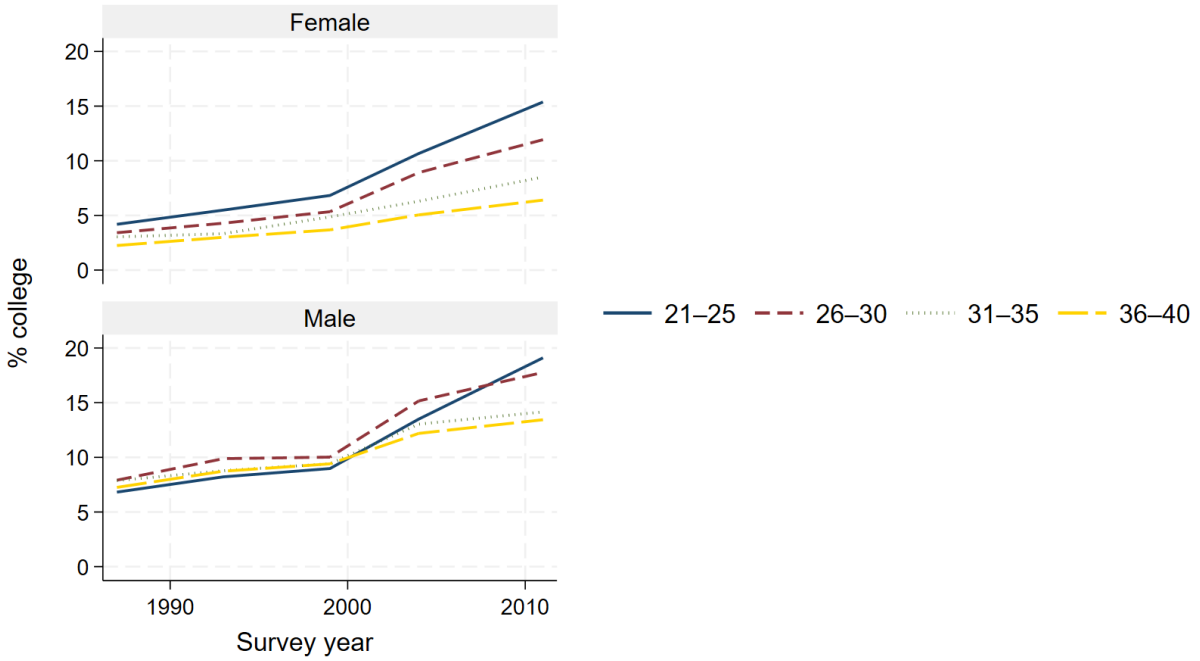


Figure B.6: Weighted-percentage bands of college completers by sex

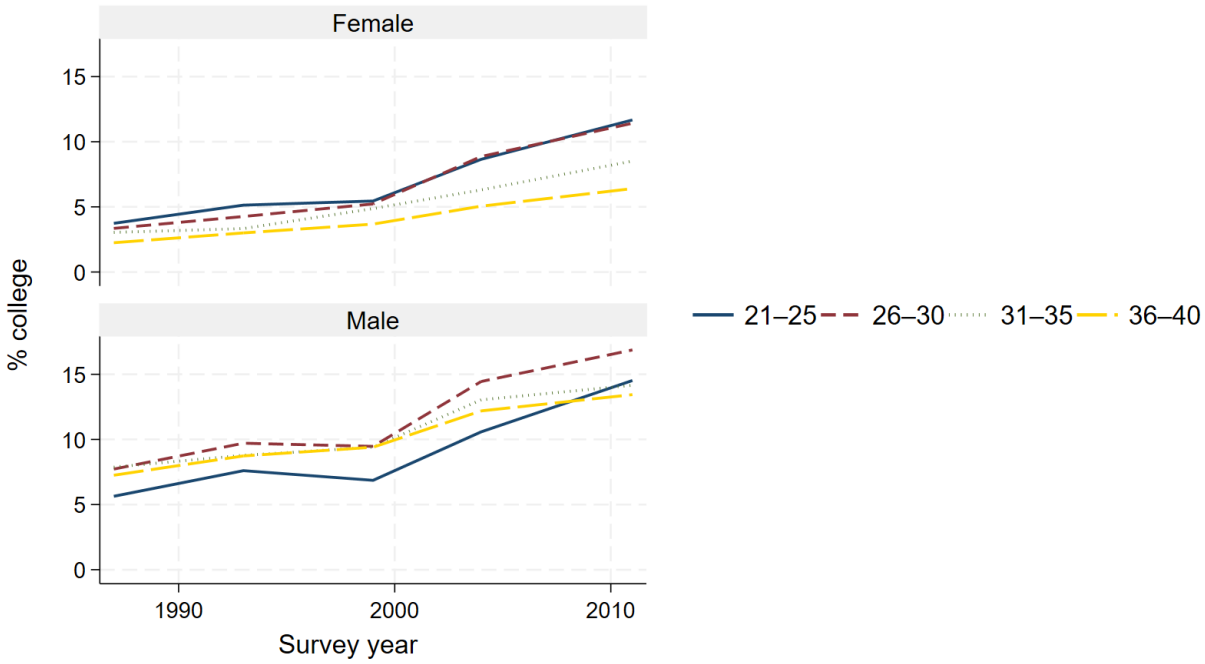


Figure B.7: Weighted-percentage bands of college completers by sex (currently not enrolled in education institution)

References

Acemoglu, D. (2002): "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 40, 7–72, invited survey.

Agarwal, P. (2006): "Higher Education in India: The Need for Change," ICRIER Working Paper 180, Indian Council for Research on International Economic Relations (ICRIER), New Delhi.

Alam, I. M., S. Amin, and J. M. Rives (2015): "A Double-Hurdle Analysis of the Gender Earnings Gap for Working Children in Bangladesh," *The Journal of Developing Areas*, 49, 41–58.

Autor, D. H. (2014): "Skills, Education, and the Rise of Earnings Inequality Among the "Other 99 Percent"," *Science*, 344, 843–851.

Azam, M. (2010): "India's Increasing Skill Premium: Role of Demand and Supply," *The B.E. Journal of Economic Analysis Policy*, 10, Article 94, uses NSS Rounds 38-61; age-group analysis in Figure 3.

Balestra, S. and U. Backes-Gellner (2017): "Heterogeneous Returns to Education over the Wage Distribution: Who Profits the Most?" *Labour Economics*, 44, 89–105.

Becker, G. S. (1964): *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, New York: Columbia University Press for the National Bureau of Economic Research, 1 ed.

Bhandari, P. and S. Thakur (2013): "Field of Study and Earnings Differentials in India's Graduate Labour Market," *Indian Journal of Labour Economics*, 56, 543–560.

Bharti, N. K. and L. Yang (2021): "Human Capital Accumulation in China and India in the 20th Century," Working paper, Paris School of Economics Berlin School of Economics, September 2021.

Bhat, S. A., S. Singh, and A. Omrane (2021): "Trends and Patterns of Women Employment in Organized and Unorganized Sector in India," *The Empirical Economics Letters*, 20, 107–123.

Card, D. and T. Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 116, 705–746.

Chakraborty, D. (2019): "Quality Matters: College Quality and Labor-Market Returns in India," *Economic and Political Weekly*, 54, 61–69.

Chen, J., S. Kanjilal-Bhaduri, and F. Pastore (2022): "Updates on Returns to Education in India: Analysis using PLFS 2018-19 Data," GLO Discussion Paper 1016, Global Labor Organization (GLO).

Cragg, J. G. (1971): "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods," *Econometrica*, 39, 829–844.

Duraisamy, P. (2002): "Changes in Returns to Education in India, 1983-94: By Gender, Age-Cohort and Location," *Economics of Education Review*, 21, 609–622.

Elek, P., J. Köllő, B. Reizer, and P. A. Szabó (2011): "Detecting Wage Underreporting Using a Double Hurdle Model," IZA Discussion Paper 6224, Institute of Labor Economics (IZA).

Emont, J. (2024): "Big Dreams Built on Higher Education Sour Worldwide for Jobless Graduates," *The Wall Street Journal*, accessed via PDF archive.

Goldin, C. and L. F. Katz (2008): *The Race Between Education and Technology*, Cambridge, MA: Harvard University Press.

Hanushek, E. A. and L. Woessmann (2008): "The Role of Cognitive Skills in Economic Development," *Journal of Economic Literature*, 46, 607–668.

Heckman, J. J., L. Lochner, and C. Taber (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1, 1–58.

Hnatkovska, V., A. Lahiri, and S. Paul (2012): "Castes and labor mobility," *American Economic Journal: Applied Economics*, 4, 274–307.

Joshi, K. M. and K. V. Ahir (2016): "Higher Education Growth in India: Is Growth Appreciable and Comparable?" *Higher Education Forum*, 13, 57–74.

Keng, S.-H., C.-H. Lin, and P. F. Orazem (2017): "Expanding College Access in Taiwan, 1978-2014: Effects on Graduate Quality and Income Inequality," *Journal of Human Capital*, 11, 1–34.

Kijima, Y. (2006): "Why Did Wage Inequality Increase? Evidence from Urban India 1983–99," *Journal of Development Economics*, 81, 97–117.

Kingdon, G. G. and J. Unni (2001): "Education and Women's Labour Market Outcomes in India," *Education Economics*, 9, 173–195.

Li, H., J. Liang, and B. Wu (2025): "Labor Market Experience and Returns to Education in Fast Growing Economies," *American Economic Review*, 115, 1123–1169.

Madden, D. (2006): "Sample Selection Versus Two-Part Models Revisited: The Case of Female Smoking and Drinking," HEDG Working Paper 06/12, Health Econometrics and Data Group, University of York.

Mincer, J. (1974): *Schooling, Experience, and Earnings*, New York: Columbia University Press for the National Bureau of Economic Research.

Mitra, S. (2021): "Inequality and Imbalances in the Education System of India: Insights from NSSO 71st Round," in *Vision of Education in India*, ed. by M. Dubey and S. Mitra, Abingdon and New York: Routledge, 173–214.

NSSO (1988): "Employment and Unemployment: National Sample Survey 43rd Round (1987-1988), Schedule 10 — All Blocks," Data set accessed 2025-04-29.

——— (1994): "Employment and Unemployment: National Sample Survey 50th Round (1993-1994), Schedule 10 — All Blocks," Data set accessed 2025-04-29.

——— (2000): "Employment and Unemployment: National Sample Survey 55th Round (1999-2000), Schedule 10 — All Blocks," Data set accessed 2025-04-29.

——— (2005): "Employment and Unemployment: National Sample Survey 61st Round (2004-2005), Schedule 10 — All Blocks," Data set accessed 2025-04-29.

——— (2012): "Employment and Unemployment: National Sample Survey 68th Round (2011-2012), Schedule 10 — All Blocks," Data set accessed 2025-04-29.

Ou, D. and Z. Zhao (2016): "Higher Education Expansion and Labor Market Outcomes for Young College Graduates," IZA Discussion Paper 9643, IZA - Institute of Labor Economics, available at IZA Discussion Paper Series.

Spence, M. (1973): "Job Market Signaling," *The Quarterly Journal of Economics*, 87, 355–374.

Tilak, J. B. G. and P. K. Choudhury (2021): "Inequality in Access to Higher Education in India between the Poor and the Rich: Empirical Evidence from NSSO Data," Paper presented at the 36th IARIW General Conference, August 2021.

Willis, R. J. and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87, S7–S36, earlier version circulated as NBER Working Paper No. 249, 1978.

World Bank (2018): "China—Systematic Country Diagnostic: Towards a More Inclusive and Sustainable Development," *World Bank Group: Washington, DC, USA*.

——— (2024):